MONASH

IP-73

CENTRE of POLICY STUDIES and the IMPACT PROJECT

Eleventh Floor
Menzies Building
Monash University Wellington Road
CLAYTON Vic 3168 AUSTRALIA

Telephone:
(03) 9905 2398, (03) 9905 5112

from overseas: 61 3 990 52398 or
61 3 9905 5112

Fax: (03) 9905 2426

from overseas: 61 3 9905 2426

e-mail

impact@vaxc.cc.monash.edu.au

Decomposing Simulation Results with Respect to Exogenous Shocks

by

W. Jill HARRISON
J. Mark HORRIDGE
and
K. R. PEARSON

Centre of Policy Studies
Monash University

Preliminary Working Paper No. IP-73 May 1999

ISSN 1 031 9034

ISBN 0 7326 1514 3

The Centre of Policy Studies and Impact Project is a research centre at Monash University devoted to quantitative analysis of issues relevant to Australian economic policy.

Decomposing Simulation Results with respect to Exogenous Shocks

W.Jill Harrison, J.Mark Horridge and K.R. Pearson

Centre of Policy Studies, Monash University, Australia

Email: Jill.Harrison@buseco.monash.ed.au

Email: Mark.Horridge@buseco.monash.ed.au

Email: Ken.Pearson@buseco.monash.ed.au

Post: Centre of Policy Studies, Monash University, Clayton 3168, Australia

Phone: +61-3-9905-5483 [Harrison] +61-3-9905-2464 [Horridge] +61-3-9905-5484 [Pearson]

Abstract

When a general equilibrium model is solved, there are often a large number of exogenous shocks. The change in each endogenous variable obviously depends on these different shocks.

We point out a natural way of decomposing the changes (or percentage changes) in the endogenous variables as sums of the contributions made by the change in each exogenous variable. The change in any endogenous variable is exactly equal to the sum of the contributions to this change attributed to each of the exogenous variables.

The contribution of a group of exogenous variables to the change (or percentage change) in any endogenous variable is defined to be the sum of the contributions of the individual exogenous variables in the group. If all the exogenous variables are partitioned into several groups that are mutually exclusive and exhaustive, the change (or percentage change) in any endogenous variable is just the sum of the contributions made by these groups.

We introduce, and motivate, these decompositions in the context of a published GTAP application in which 10 regions remove import tariffs and non-tariff barriers to imports. We use the methods given in this paper to report numerical values for the contributions to the welfare gains of various regions due to tariff reductions by particular regions or groups of regions in this simulation. We show how the values obtained via the decomposition are related to the estimates in the published study of the contributions to welfare gain due to certain groups of tariff reductions.

We describe a practical procedure for calculating the contributions of individual exogenous variables or groups of exogenous variables to the changes (or the percentage changes) in all of the endogenous variables. This procedure, which applies to a wide range of general equilibrium models, is now automated in GEMPACK in a version that will be made publicly available in the future.

The contributions that make up the decomposition are defined as integrals. As such, they depend on the path by which the exogenous values move from their pre-simulation to post-simulation values. We propose one natural path, namely a straight line between these two points. Along this path, the ordinary rate of change is constant for each variable.

JEL classification: C63, C68

Contents

Decomposing Simulation Results with Respect to Exogenous Shocks

W.Jill HARRISON, J.Mark HORRIDGE and K.R. PEARSON

1. Introduction

Policy simulations using computable general equilibrium (CGE) models often involve a multiplicity of shocks. It is frequently of interest to partition the total effect of a package of shocks between individual shocks or groups of shocks. For example we might ask, how much of the welfare change that North America derives from a multilateral trade liberalization is due to trade liberalization by Japan? If the shocks are partitioned into several groups, and if the contributions attributed to these groups add exactly to the overall simulation result, we call this a **decomposition** of the simulation results. Analysts have used various methods to calculate decompositions; in certain cases, as we shall see, the different methods yield quite different numerical results. We present yet another method of decomposing total endogenous changes into the effects of individual shocks. We argue that our method has a natural interpretation, yields sensible values, and is in addition rather cheap and easy to compute.

We begin, in section 2, by applying our proposed decomposition method to a GTAP simulation from the literature. We compare our results to those yielded by other decomposition approaches.

Sections 3 and 4 are devoted to the mathematical details of our decomposition. In section 3 we explain our method under the simplifying (but rarely true) assumption that endogenous variables are explicit functions of exogenous variables.

We delay until section 4 the general case, where we explain how the decomposition introduced in sections 2 and 3 can be calculated for models expressed as a simultaneous system of implicit (nonlinear) equations. Section 4 includes a brief discussion of the GEMPACK implementation of this decomposition.

In section 5 we describe briefly how this decomposition extends over a sequence of simulations, where each simulation begins from the post-simulation data arising from the previous simulation (such as in a year-to-year forecast).

We are not sure if this decomposition is a new result. If any reader knows of an existing derivation, we will be most grateful to hear of it.

We are grateful to Peter Dixon for encouraging us to try to find a decomposition with respect to exogenous variables.

2. Reducing Import Barriers: Who Gains from What?

In this section we revisit the simulations reported by Linda Young and Karen Huff in Chapter 9 of the book *Global Trade Analysis: Modeling and Applications* (Hertel ed., 1997). We have chosen this particular application because:

(a) we can build on the careful analysis of Young and Huff;

(b) he chapter illustrates the strengths and limitations of a traditional method of decomposing simulation results;

(c) the GTAP project has widely distributed the computer files needed to replicate the experiment; and

(d) in these simulations the GTAP model exhibits some rather non-linear behaviour – this increases the contrast between results derived from alternative decomposition methods.

The Young-Huff chapter, entitled *Free trade in the Pacific Rim: On what basis?*, examines the effect of eliminating import barriers both within members of the APEC (Asia Pacific Economic Cooperation) group and between the APEC group and the rest of the world (ROW). Two important features of the simulations are (i) that they start from a synthetic database that incorporates the effects of NAFTA, and (ii) that they show only the effects of eliminating import restrictions – export subsidies and taxes are not altered.

The simulations show the effects of three packages of shocks:

1. Removal of import barriers between the 9 APEC regions (81 shocks).
2. Removal of barriers on imports to APEC from the ROW (9 shocks).
3. Removal of barriers on imports to the ROW from APEC (9 shocks).

Some of the results are summarized in the following table[1]:

Table 1: Effects of trade liberalization on welfare ($US million)

	A	B	C	D	E
	Preferential: effect of [1]	effect of [2]	MFN: ROW does not reciprocate	effect of [3]	MFN: with ROW reciprocating
effect of shock:	[1]	[2|1]	[12]	[3|12]	[123]
North America	-6611	-10108	-16720	14467	-2252
Japan	72289	-14489	57800	37494	95294
Australia & New Zealand	-202	-929	-1131	1871	740
China & Hong Kong	5940	-3147	2792	4633	7426
Taiwan	5091	-1726	3366	3143	6508
South Korea	8471	-1555	6916	4203	11119
Malaysia & Singapore	2147	-1246	901	1947	2848
Thailand & Philippines	-4509	125	-4384	2858	-1526
Indonesia	-202	-978	-1180	1751	571
ROW	-31631	49241	17610	-64698	-47088

Columns A, C and E show results from 3 central experiments, namely:

[1] Compare this with Table 9.6 of the Young-Huff chapter. The numbers differ slightly because we have computed the results more accurately: the same data, model, and shocks were used. Column D is our own addition.

A: *Preferential* incorporating shock package 1 above, labelled [1].

C: *MFN: ROW does not reciprocate* incorporating shock packages 1 and 2 above, labelled [12].

E: *MFN: with ROW reciprocating* incorporating all 3 shock packages above, labelled [123].

In the spirit of our question from the Introduction, it is natural to ask how much of the total benefit that North America derives in column E of Table 1 is due to each of shock packages [1], [2] and [3]?

Columns A, B and D supply one set of answers to these questions. Column A shows the effect of applying shock package [1] to the initial equilibrium. Column B shows the effect of applying shock package [2] to an equilibrium that has already absorbed the effects of [1]; it is labelled [2|1] (2 given 1). Similarly column D shows the effect of applying shock package [3] after both [1] and [2]; it is labelled [3|12] (3 given 1 and 2).

Columns A (effect of [1]), B (effect of [2]) and D (effect of [3]) exactly sum to the total results in column E. In short, A, B and D are a meaningful decomposition of E[2].

Table 2: Different estimates of the contribution of the 3 shock packages to North American welfare change ($US million)

	Method	Effect of [1]	Effect of [2]	Effect of [3]	Total
1	order 123	-6611	-10108	14467	-2252
2	order 132	-6611	-13577	17936	-2252
3	order 213	-2725	-13994	14467	-2252
4	order 231	2762	-13994	8980	-2252
5	order 312	-1809	-13577	13134	-2252
6	order 321	2762	-18148	13134	-2252
7	average	-2039	-13900	13686	-2252
8	new way	-973	-14887	13608	-2252

In Table 1, the decomposition of the total effect into columns A, B and D depends on our applying shock packages [1], [2] and [3] *in that order*. We might ask, how sensitive are our estimates of the contributions to the order in which shocks are applied? That question is addressed by the first 6 rows of Table 2, which show some alternative methods of decomposing the North American welfare change. Row 1 shows the effect of applying the shocks in the order 123, and is thus the same as columns A, B and D in the North America row of Table 1. Since there are 6 (=3!) way of ordering 3 shocks, the next five rows show how the same three contributions could be computed[3] using different shock orderings. The variation between the first six rows shows that these sequential estimates of the contributions of the 3 shock packages are rather sensitive to the

[2] Simulations [1], [12], and [123] all began from the same starting point. To compute column B of Table 1, [2|1], we apply shocks [2] to the post-simulation database produced by simulation [1]. Another way to compute the same result would simply be to take the difference between columns A and C, that is, [2|1] = [12]-[1]. Similarly, we could also compute column D, [3|12], as the difference between columns C and E ([3|12] = [123]-[12]).

[3] Although for n shocks there are n! decompositions corresponding to different shock orders, there are only 2^(n-1) ways to calculate each particular contribution. Thus some numbers appear twice in rows 1 to 6 of Table 2.

order in which shocks are applied. The variation arises because the effects of each shock depend to some degree on the database to which it is applied – and this in turn is affected by previously applied shocks.

Similarly, because the model is nonlinear, the obvious strategy of individually applying each of packages [1], [2] and [3] to the **original** database yields estimates of contributions that are not a decomposition of the welfare results since they do **not** add up to the total effect in column E.[4]

Sensitivity of contribution estimates to shock ordering could be a problem if similar techniques were used to decompose the effects of other simulations, in which shocks could not be ordered in an obvious or natural way[5].

If the shocks were divided into N groups, there would be N! ways of decomposing the total result. In Table 2, N=3, giving rise to rows 1 to 6. For larger N, we could not compute or compare so many combinations.

The large number of possible shock orderings, combined with the potential sensitivity of decomposition results to the order used to calculate contributions, are disadvantages of the sequential method of result decomposition. An order-independent way of calculating contributions seems desirable.

Returning to Table 2, row 7 shows the average of the preceding 6, while row 8 shows a decomposition computed by our proposed new method, which we will now explain.

2.1. Preview of the New Method

Suppose that one endogenous variable Z can be expressed as a function F of n exogenous variables $X_1, X_2, \ldots, X_n$ via the equation

$$Z = F(X_1,\ldots, X_n) \quad .$$

Suppose that the vector of exogenous variables $\underset{\sim}{X} = (X_1,\ldots, X_n)$ moves along some path

beginning at $\quad \underset{\sim}{X}_0 = (X_{10},\ldots, X_{n0})$

and ending at $\quad \underset{\sim}{X}_1 = (X_{11}, \ldots, X_{n1}) = (X_{10}+\Delta X_1,\ldots, X_{n0}+\Delta X_n)$.

Now suppose that we divided all the shocks into 100 equal instalments. The effect of applying the first instalment (ie, one hundredth part of all the shocks) could be accurately approximated as:

$$dZ = F_1 dX_1 + F_2 dX_2 \ldots + F_n dX_n \quad \text{where} \quad F_i = \partial F/\partial X_i \quad \text{and} \quad dX_i = \Delta X_i/100$$

Provided the dX_i were small enough the approximation would be exact, and the terms on the right hand would unambiguously distribute the total change dZ between the n exogenous variables[6].

[4] Moving along the diagonal of Table 2 we see that the effects on North America of applying packages [1], [2] and [3] to the original database are respectively -6611, -13994, and 13134. These sum to -7471, not -2252.

[5] In the Young and Huff chapter, the order [1]-[2]-[3] arose naturally from their discussion of the political and eco nomic background of trade issues. It seemed unlikely, for example, that ROW would admit APEC imports, if APEC did not first admit ROW imports. In the general case, such a natural ordering may not be apparent.

We could go on to apply the other 99 instalments in just the same way (the F_i, which depend on X and Z, would change as we progressed), and by adding up the contributions obtained at each step, obtain the final contribution, ΔZ_i, of each shock ΔX_i to the total change ΔZ, *along this path.*

Row 8 of Table 2 was computed in essentially this way[7].

We can visualize exogenous space as an n-dimensional cube with the starting point $\underset{\sim}{X}_0$ at one vertex and the ending point $\underset{\sim}{X}_1$ at the diagonally opposite vertex. Under the method just described, the exogenous variables move *together* towards their final value along a *straight line* (through the interior of the cube) between these diagonally opposite vertices (since $dX_i = \Delta X_i/100$ in each instalment). By contrast, rows 1 to 6 of Table 2 could be computed by adding over different paths in which only one group of variables at a time was changing. These paths also lead from vertex $\underset{\sim}{X}_0$ to $\underset{\sim}{X}_1$ diagonally opposite, *travelling only along the edges.*

Since the straight line path is an average of all possible edgewise routes, it would be natural to suppose that each of our straight line estimates, ΔZ_i, might lie in the middle of the range of the corresponding estimates derived from the various edgewise routes. This would seem particularly likely if the partial derivatives $F_i = \partial F/\partial X_i$ were monotonic functions of the X_i over the relevant range. And indeed we do see that each value in row 8 of Table 2 lies within the range of rows 1 to 6 in the same column.

Some support for this intuition comes from the special case where the function F is quadratic, that is,

$$Z = F(X_1,..., X_n) = \sum_{i=1,n} \sum_{j=i,n} \beta_{ij} X_i X_j .$$

Then it turns out that our estimate of ΔZ_i (the part of ΔZ due to ΔX_i), which we compute as an integral, is just equal to the arithmetic mean of the various estimates of ΔZ_i which might be obtained by shocking one variable at a time[8]. This mean is shown in row 7 of Table 2. If welfare in the GTAP model were a quadratic function of import tariffs, we should expect rows 7 and 8 of Table 2 to be identical.

Table 3, which is computed using our new method, is the analogue of Table 1. Columns [1], [2] and [3] show the contributions of shock groups [1], [2] and [3] towards the total change in each region's welfare: their sum is the same as column E of Table 1. Unlike Table 1, Table 3 does not require us to impose any particular order of application for shocks.

Table 4 is a more ambitious application of our new method. This time, each country's welfare change is decomposed into the parts due to the lifting of restrictions on imports to each to the 10 regions. This enables us to answer the question posed at the beginning of the Introduction: actions[9] by Japan contributed 12995 to towards North America's total welfare change of -2252.

---

[6] Luckily, the values for F_i, which our decomposition requires, arise as a by-product of the normal GEMPACK solution algorithm (see section 4.2 for details).

[7] The adding of the 100 instalments is replaced by an integral (which is the precise way of adding infinitely small instalments), as explained in section 3.

[8] We prove this in Section 3.9.

[9] The 'actions' cover the abolition of barriers to all imports to Japan, not just imports from North America. This suggests a further decomposition, which would also be easy to compute.

Table 3: Contributions of 3 groups of import liberalizing shocks to total welfare changes ($US million)

Effect of facilitating imports from:	APEC to APEC	ROW to APEC	APEC to ROW	Total
	[1]	[2]	[3]	[1]+[2]+[3]
North America	-973	-14887	13608	-2252
Japan	70843	-15934	40385	95294
Australia & New Zealand	987	-2132	1884	740
China & Hong Kong	7547	-4774	4653	7426
Taiwan	5582	-2208	3134	6508
South Korea	10171	-3125	4073	11119
Malaysia & Singapore	2408	-1473	1913	2848
Thailand & Philippines	-1699	-2628	2801	-1526
Indonesia	340	-1498	1729	571
ROW	-35745	53977	-65321	-47088

Table 4: Contribution of each region's liberalization to welfare change in each region ($US million)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	NAM	Jpn	ANZ	Chn	Twn	SKor	MySg	ThaPh	Idn	Row	Total
NAM: North America	-34542	12995	392	544	2222	1610	-129	866	183	13608	-2252
Jpn: Japan	38165	3342	1837	2735	575	3202	349	3767	936	40385	95294
ANZ: Australia,NewZealand	396	1055	-2683	474	-184	-206	12	1	-10	1884	740
Chn: China, Hong Kong	1659	2108	350	-3053	256	1230	-78	165	135	4653	7426
Twn: Taiwan	1271	128	80	543	252	124	157	631	189	3134	6508
SKor: South Korea	796	1060	175	-160	48	4293	46	439	348	4073	11119
MySg: Malaysia, Singapore	-426	262	214	-121	106	346	-59	522	90	1913	2848
ThaPh: Thailand, Philippines	-173	353	5	-36	122	476	80	-5163	9	2801	-1526
Idn: Indonesia	140	186	-30	233	-22	68	-48	97	-1783	1729	571
ROW	3483	6078	971	3333	1252	655	-201	1936	726	-65321	-47089

As in our previous example, we could have produced a decomposition similar to that of Table 4, by assuming that regions acted in a particular order, and measuring the effect of each successive liberalization. With 10 regions, the different possible orders allow more than 3 million (=10!) different decompositions – no one of which is obviously more plausible than the rest. Our decomposition, which treats all regions equally, is, so far as we are aware, the only practical way to produce more complex decompositions such as that of Table 4.

3. The Decomposition

In this section we introduce the decomposition in the case where the endogenous variables can be written explicitly as functions of the exogenous variables. Experienced modellers will realise that this assumption is almost never true in practice. However it is the easiest way to understand the decomposition.

In section 4 we explain how the decomposition can be calculated in the more realistic case where the exogenous and endogenous variables are linked by implicit functions.

3.1. The Decomposition for One Endogenous Variable

In this section we suppose that there is just one endogenous variable Z. [We address the case where there are several endogenous variables in subsection 3.3 below.] We suppose that Z can be expressed as a function F of the n exogenous variables $X_1, X_2, \ldots, X_n$ via the equation

$$Z = F(X_1,\ldots, X_n) \quad . \tag{1}$$

Suppose $\underset{\sim}{X} = (X_1,\ldots, X_n)$ moves from

$$\underset{\sim}{X}_0 = (X_{10},\ldots, X_{n0}) \text{ to } \underset{\sim}{X}_1 = (X_{11}, \ldots, X_{n1}).$$

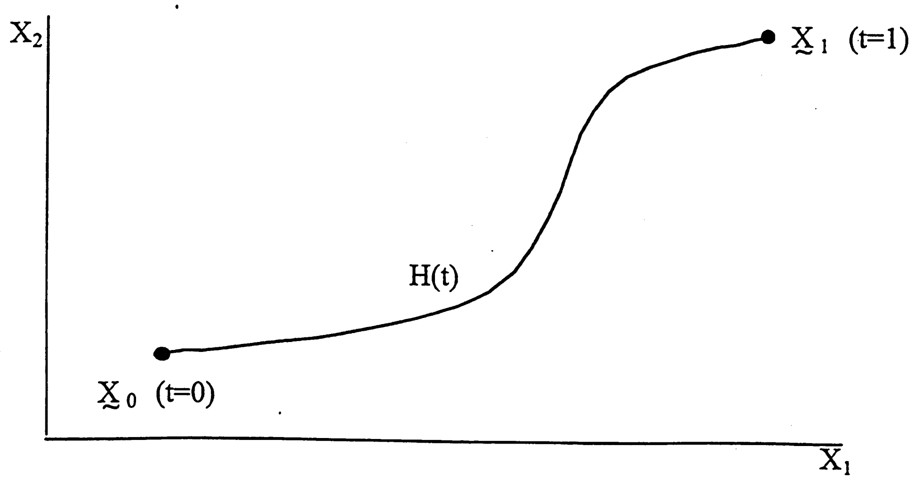

Diagram A: Path of 2 exogenous variables (n=2)

Here we are thinking of a simulation where $\underset{\sim}{X}_0$ is the vector of exogenous values at the pre-simulation solution of the model and $\underset{\sim}{X}_1$ is the vector at the post-simulation solution of the model. Suppose further that $\underset{\sim}{X} = (X_1,\ldots, X_n)$ moves from $\underset{\sim}{X}_0$ to $\underset{\sim}{X}_1$ along some path parameterized by t where t moves from 0 to 1. That is,

$$\underset{\sim}{X} = \underset{\sim}{H}(t)$$

where $\underset{\sim}{X}_0 = \underset{\sim}{H}(0)$ and $\underset{\sim}{X}_1 = \underset{\sim}{H}(1)$. Then

$$Z = F(X_1,\ldots, X_n) = F(\underset{\sim}{X}) = F(H(t)) = Q(t)$$

for some function Q. Suppose Z has pre-simulation value Z_0 (that is, when $\underset{\sim}{X} = \underset{\sim}{X}_0$) and that Z has post-simulation value Z_1 (when $\underset{\sim}{X} = \underset{\sim}{X}_1$). Then

$$Z_0 = F(\underset{\sim}{X}_0) = F(H(0)) = Q(0) \quad \text{and}$$

$$Z_1 = F(\underset{\sim}{X}_1) = F(H(1)) = Q(1)\,.$$

Then, under the assumption that F and H are differentiable functions, it follows from the Chain Rule [see, for example, Theorem 6-14 of Apostol (1957)] that

$$dZ/dt = \sum_{i=1,n} (\partial F/\partial X_i)(dX_i/dt) \quad . \tag{2}$$

Integration of (2) with respect to t shows that

$$\int_{t=0}^{1} (dZ/dt)\, dt = \sum_{i=1,n} \int_{t=0}^{1} (\partial F/\partial X_i)\,(dX_i/dt)\, dt \quad .$$

The left-hand side of the above equals

$$Z(\text{when } t{=}1) - Z\,(\text{when } t{=}0) = Z_1 - Z_0$$

and so is equal to the change in Z. Thus we have

$$c\_Z = \sum_{i=1,n} \int_{t=0}^{1} (\partial F/\partial X_i)\,(dX_i/dt)\, dt \tag{3}$$

where c_Z denotes the change in Z.

This is the decomposition. We see that the change in Z is equal to the sum of the contributions due to each of $X_1, \ldots, X_n$, where these contributions are as defined below.

3.1.1. Definition (Contribution due to One Exogenous Variable)

The *contribution to the change in Z due to the change in X_i as $\underset{\sim}{X}$ moves along the path H* is defined to be

$$\int_{t=0}^{1} (\partial F/\partial X_i)\,(dX_i/dt)\, dt \quad .$$

3.1.2. Proposition (The Decomposition for One Variable)

Under the assumptions set out above (including those that F and H are differentiable functions), the change in the endogenous variable Z is equal to the sum of the contributions due to each of $X_1, \ldots, X_n$ as $\underset{\sim}{X}$ moves along the path H.

3.1.3. Example

Consider the example shown in Diagram B, in which

$$Z = X_1 X_2, \quad \underset{\sim}{X}_0 = (1, 1), \quad \underset{\sim}{X}_1 = (2,3)\,.$$

Suppose that $\underset{\sim}{X} = (X_1, X_2)$ moves along a straight line from $\underset{\sim}{X}_0$ to $\underset{\sim}{X}_1$. Then

$$X_1 = 1 + t, \quad X_2 = 1 + 2t, \quad \text{for } 0 \le t \le 1,$$

$$c\_Z = 6 - 1 = 5\,.$$

Now $\partial F/\partial X_1 = X_2$ so that the contribution to the change in Z due to X_1 as $\underset{\sim}{X}$ moves along this path is

$$\int_{t=0}^{1} X_2.1dt = \int_{t=0}^{1} (1+2t)\, dt = 2 \;.$$

Similarly, the contribution due to X_2 is

$$\int_{t=0}^{1} X_1.2dt = \int_{t=0}^{1} (1+t).2dt = 3 \;.$$

Note that the sum of these two contributions is 5, as expected from Proposition 3.1.2.

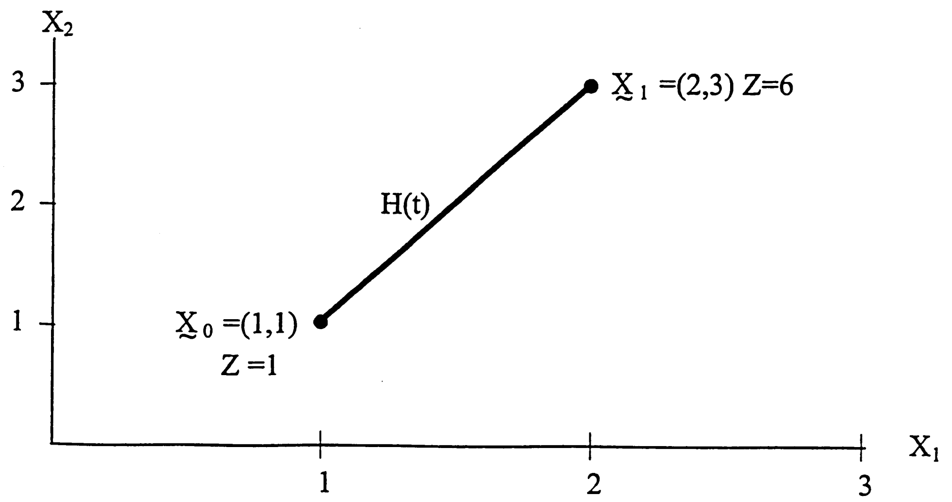

Diagram B: Path for example 3.1.3

3.2. Contributions Made by Groups of Exogenous Variables

Suppose we group the exogenous variables into two groups, say $X_1,\ldots,X_p$ and $X_{p+1},\ldots,X_n$. Then

$$c\_Z = \sum_{i=1}^{p} (\text{cont. due to } X_i) + \sum_{i=p+1}^{n} (\text{cont. due to } X_i) \;. \tag{4}$$

The first term on the right-hand side is defined to be the contribution due to the first group of exogenous variables, and similarly for the second.

In general, the contribution due to any set of exogenous variables is defined to be the sum of their individual contributions.

This is stated a little more formally in the next definition. The proposition below follows easily from Proposition 3.1.2 and equation (4).

3.2.1. Definition (Contribution of Group)

The *contribution to the change in Z as $\underset{\sim}{X}$ moves along the path H due to any set of exogenous variables* is defined to be the sum of their individual contributions to the change in Z as $\underset{\sim}{X}$ moves along the path H.

3.2.2. Proposition (The Decomposition for Groups of Shocks)

If the set of exogenous variables is partitioned into several mutually exclusive and exhaustive subsets then, under the assumptions in Proposition 3.1.2, the change in Z is equal to the sum of the contributions of these sets of exogenous variables as $\underset{\sim}{X}$ moves along the path H.

3.3. The Decomposition for Several Endogenous Variables

As a generalisation of section 3.2 above, now suppose that there are several endogenous variables (not just one as in section 3.1), say $Z_1, Z_2, \ldots, Z_m$. Suppose also that each of these endogenous variables can be expressed as an explicit function of the exogenous variables via the m equations

$$Z_j = F_j (X_1,\ldots, X_n) \qquad j = 1,\ldots, m , \tag{5}$$

where each function F_j is differentiable. Then

$$dZ_j/dt = \sum_{i=1}^{n} (\partial F_j/\partial X_i).(dX_i/dt) \qquad j = 1,\ldots, m . \tag{6}$$

Then we can define the contribution of any X_i (or groups of X_i's) to the change in any one of the Z_j's.

The following proposition is clear.

3.3.1. Proposition (The Decomposition)

Consider (5) above. If the set of exogenous variables is partitioned into several mutually exclusive and exhaustive subsets then, for $1 \le i \le m$, the change in Z_i is equal to the sum of the contributions of these sets of exogenous variables to the change in Z_i as $\underset{\sim}{X}$ moves along the path H (provided $F_1,\ldots,F_m$, H are differentiable).

3.3.2. Example

Suppose that all is as in Example 3.1.3 above and that there is a second endogenous variable W given by the equation

$$W = X_1 / X_2 .$$

The change in W is equal to $(2/3) - (1/1) = -1/3$.

The contribution to the change in W due to X_1 as $\underset{\sim}{X}$ moves along the straight-line path in Diagram B is equal to

$$\int_{t=0}^{1} (1/X_2).1dt = \int_{t=0}^{1} dt/(1+2t) = (ln3)/2 .$$

Similarly, the contribution to the change in W due to X_2 as $\underset{\sim}{X}$ moves along this path is equal to

$$\int_{t=0}^{1} (-X_1/X_2^2).2dt = \int_{t=0}^{1} -[2(1+t)/(1+2t)^2]\ dt = -(1/3) - (ln3)/2 .$$

Note that these two contributions do indeed add to -1/3, as expected from Proposition 3.3.1. [Of course the contributions to the change in Z due to X_1 and X_2 are as calculated in Example 3.1.3.]

3.4. The Decomposition Depends on the Path of the Exogenous Variables

Although the integral of the left-hand side of (2) does not depend on the path H (since its value is equal to the change c_Z in Z whatever the path), the integrals which are defined to be the contributions of the different exogenous variables X_i to the change in Z do depend on the path H by which $\underset{\sim}{X}$ goes from $\underset{\sim}{X}_0$ to $\underset{\sim}{X}_1$. This can be seen from the next example.

3.4.1. Example

Consider again Example 3.1.3, but with a different path for moving from (1,1) to (2,3). See Diagram C. This time first go along to point A = (2,1) in a straight line and then go up from there to (2,3) in a straight line.

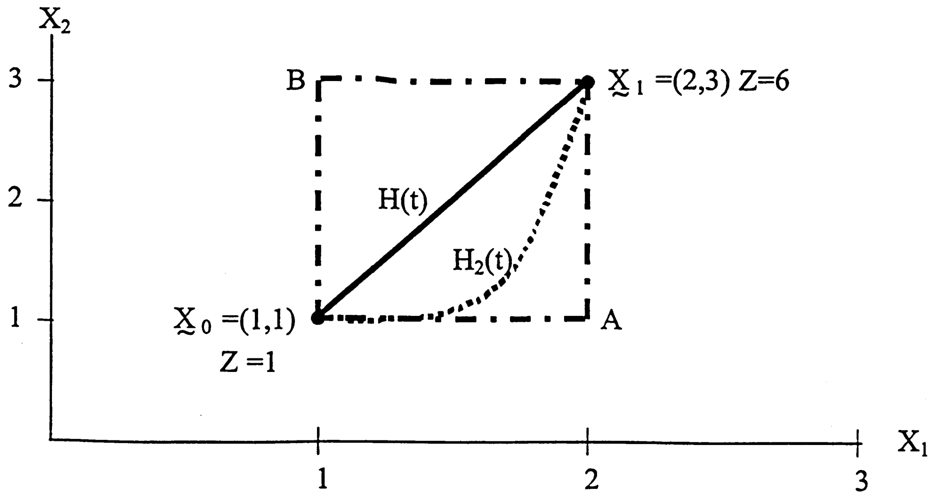

Diagram C: Evaluating Contributions along Alternative Paths

Clearly the first part of this path produces zero contribution from X_2 (which does not change on this part of the path), while the second part produces zero contribution from X_1. Equally clearly the contribution to c_Z from X_1 along the first part of this path is equal to all the difference in Z along this part of the path, namely 2-1=1. And the contribution to c_Z from X_2 along the second part of this path is equal to all the difference in Z along this part of the path, namely 6-2=4.

Alternatively, start by going up to point B: only X_2 is changing and Z increases by 2. Then go along to (2,3): now only X_1 is changing and Z increases by 3. So, going via B, X_1 contributes 3 and X_2 contributes 2.

Table 5 summarizes the contributions due to X_1 and X_2 along all 3 paths: the straight line path and the edge-wise routes visiting A and B. Note that:

- the contribution to c_Z due to a particular variable depends on the path taken,
- the row sum of the contributions (total change in Z) does not depend on the path taken,
- the straight line estimates lie between the estimates arising from the two edge-wise paths. Indeed, because Z is a quadratic function of X_1 and X_2, the straight line estimates are just the arithmetic mean of the estimates from the other two paths (proved in section 3.9 below).

Table 5: Contributions to c_Z along 4 different paths

Path	contribution due to X_1	contribution due to X_2	Total
$\underset{\sim}{X}_0$ to $\underset{\sim}{X}_1$ via A	1	4	5
$\underset{\sim}{X}_0$ to $\underset{\sim}{X}_1$ direct	2	3	5
$\underset{\sim}{X}_0$ to $\underset{\sim}{X}_1$ via B	3	2	5
$\underset{\sim}{X}_0$ to $\underset{\sim}{X}_1$ via $H_2(t)$	1.67	3.33	5

The last row of Table 5 shows the contributions calculated along the curve $H_2(t)$ in Diagram C. That curve is defined by the parametric equations $X_1 = 1 + t$ and $X_2 = 1 + 2t^2$. Notice that the curve lies between the direct route and the route via A; the contributions from the curve also lie between those computed along the other two routes (see first 2 rows of Table 5).

3.5. Is There a Natural Path?

Because the contributions due to the different exogenous variables may depend on the path chosen from the pre-simulation exogenous values to the post-simulation values, it is natural to ask if there is a preferred path to take. Probably the answer to this is somewhat model-specific or even simulation-specific.

However, we think that there is one natural path, namely the straight line between the pre- and post-simulation values (as, for example, used in Example 3.1.3 above). Note that,

a) the rate of change in any exogenous variable is constant along this path,

b) the value of each exogenous variable remains between its pre- and post-simulation values at every point along this path.

Of course, many different paths have property (b) but only the straight line has property (a).

It really does not matter whether you judge this to be the most natural path (or one of the natural paths). The Decomposition, namely Proposition 3.3.1, remains valid for whatever path you may prefer.

Nevertheless, even broad-minded people may agree that some paths are not natural. Some of these are shown, for 2 dimensions, in Diagram D on the next page.

We can exclude paths such as these by requiring that as we move along a path parameterized by t ranging from 0 to 1, dX_i/dt does not change sign. That is, each exogenous variable either increases continuously or decreases continuously. As well as precluding kinky and loopy behaviour, this restricts paths to within the zone bounded by dotted lines in Diagram D. From this point of view, sequential decompositions, represented by traverses along the dotted lines, tread the boundaries of acceptable behaviour.

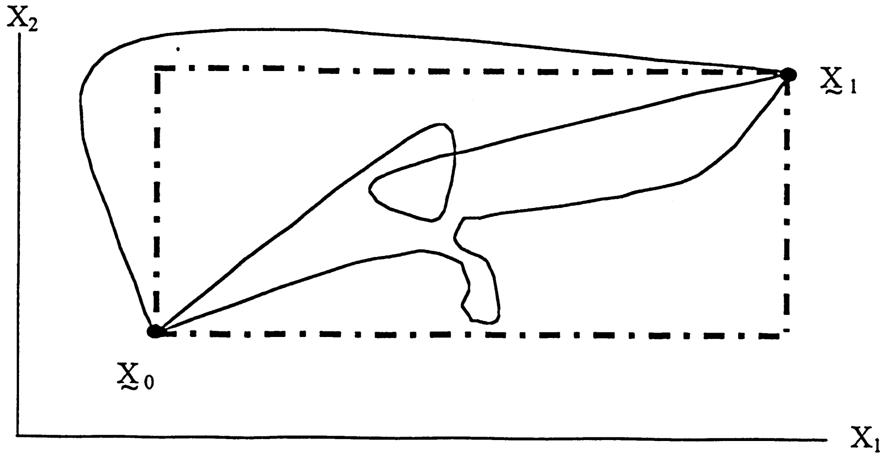

Diagram D: Some Unnatural or Perverse Paths

3.6. Decomposition of Percentage-Change Results

Many modellers prefer to report mainly percentage-change results (rather than changes). The decomposition here is easily adapted to that. It is easy to define the contribution to the percentage change in any endogenous variable due to any group of exogenous variables along the relevant path. Again the contributions to this percentage change due to the different groups of mutually exclusive and exhaustive sets of exogenous variables add up to the total percentage change. We show this in this section.

Suppose we have just one endogenous variable Z, as in section 3.1 above. Then

$$c\_Z = Z_0 \,.\, p\_Z/100$$

where Z_0 is the initial (pre-simulation) value of Z, and p_Z is the percentage change in Z. From (2) we see that

$$Z_0.\ p\_Z/100 = \sum_{i=1}^{n} (\text{cont. due to } X_i) \quad .$$

Thus

$$p\_Z = \sum_{i=1}^{n} [100.(\text{cont. due to } X_i)]/Z_0 \,. \tag{7}$$

The terms on the right-hand side are defined to be the contributions to the percentage change in Z due to the different X_i. Thus the contribution of X_i to the percentage change in Z is simply related to its contribution to the change in Z. Formally we have the following definition and proposition.

3.6.1. Definition (Contribution to Percentage Change)

The *contribution to the percentage change in Z due to X_i along the path H* is defined to be

$$100.C_i\,/Z_0$$

where Z_0 is the pre-simulation value of Z and C_i is the contribution to the change in Z due to X_i along the path H.

3.6.2. Proposition (The Decomposition for a Percentage-change Variable)

Consider (5) above. If the set of exogenous variables is partitioned into several mutually exclusive and exhaustive subsets then the percentage change in Z is equal to the sum of the contributions to percentage change in Z due to these sets of exogenous variables as X moves along the path H (provided $F_1,\ldots,F_m$, H are differentiable).

Thus the decomposition can be given equally well for the change or the percentage change in each endogenous variable.

3.7. Implicit Relations between Endogenous and Exogenous Variables

Above the decomposition was explained and derived only in the case where each of the endogenous variables can be expressed analytically (or algebraically) as functions of the exogenous variables. Usually such algebraic expressions are not available for a general equilibrium model. Rather the relation between the exogenous and endogenous variables is implicit as in the system of equations

$$G_j(Z_1, \ldots, Z_m, X_1,\ldots, X_n) = 0 \qquad j = 1, \ldots, m \ . \tag{8}$$

While it is not usually possible to write down algebraic expressions for the partial derivatives which appear in the definition of the contributions due to each X_i , it is usually possible to calculate numerical values for these at all relevant points. This can be done by solving a system of numerical linear equations, as we explain in detail in section 4. Then it is possible to calculate arbitrarily accurate numerical approximations to the integrals used to define the contributions due to the different exogenous variables. Hence the decomposition stated above also makes sense when the model consists of a set of equations as in (8) above. The contributions due to any group of exogenous variables can be calculated numerically; we explain this in more detail in section 4.2.

Hence

> ***the Decomposition in Proposition 3.3.1 applies to models expressed as a system of non-linear equations, as in (8) above.***

The GTAP model used in section 2 can be represented as a system of equations like (8) above. The numerical decompositions presented in section 2 were calculated using a new version of the GEMPACK software (see Harrison and Pearson (1996)) which includes an implementation of the algorithm set out in section 4.

3.8. Is the Decomposition a New Result?

When we found the decomposition set out above, our first thought was that it was rather simple and was surely known before. This has been the first reaction of most of the colleagues to whom we have described the decomposition. But, so far, no one has been able to point to an earlier exposition of this decomposition. Further, we (and several others) feel that if the decomposition were known, probably there would be software facilitating the reporting of the contributions of different groups of exogenous variables. Perhaps one of the readers on this paper can point us to an earlier exposition and/or software which does these calculations; if so, we will be most grateful.

3.9. When the Model is Quadratic

Here we prove the claim made in section 2.1, that where an endogenous variable Z can be expressed as as a quadratic function of the exogenous variables, namely,

$$Z = F(X_1,..., X_n) = \sum_{i=1,n} \sum_{j=i,n} \beta_{ij}X_iX_j \qquad \text{(here } \beta \text{ is upper triangular)},$$

our straight-line estimate of the contribution ΔZ_i (the part of ΔZ due to ΔX_i) is just equal to the arithmetic mean of the various estimates of ΔZ_i which might be obtained by shocking one variable at a time.

To compute one of the latter estimates, we start by assuming, without loss of generality, that we are evaluating the contribution of X_1, that is, i=1. The change in Z as X_1 moves from its initial value, X_{10}, to its final value of X_{11} (= $X_{10}+\Delta X_1$), is given by:

$$\text{(one sequential) } \Delta Z_1 = \beta_{11} \Delta X_1(2X_{10} + \Delta X_1) + \Delta X_1 \sum_{j=2,n} \beta_{1j}X_j$$

The variation in these estimates arises only from variation in the final term, which is linear in X_j. Each of these X_j (j=2..n), might be evaluated either at its initial or its final value. Each of the n! ways of ordering shocks give rise to its own set of X_j values.

If we consider one particular X_j it is clear that in half of the n! cases it will have the initial value X_{j0}; otherwise it will have the final value X_{j1} (= $X_{j0}+\Delta X_j$). Therefore the average of the n! sequential estimates is:

$$\text{(average sequential) } \Delta Z_1 = \beta_{11} \Delta X_1(2X_{10} + \Delta X_1) + \Delta X_1 \sum_{j=2,n} \beta_{1j}(X_{j0} + \Delta X_j/2)$$

Next we compute the contribution of X_1 using our integral method. From Definition 3.1.1,

$$\Delta Z_1 = \int_{t=0}^{1} (\partial F/\partial X_1)\,(dX_1/dt)\,dt = \int_{t=0}^{1} [2\beta_{11}X_1 + \sum_{j=2,n} \beta_{1j}X_j]\,(dX_1/dt)\,dt.$$

Along our preferred straight-line path, as t goes from 0 to 1,

$$X_j = X_{j0} + t.\Delta X_j \quad \text{for all j.}$$

So $\quad dX_1/dt = \Delta X_1$, giving

$$\text{(straight-line) } \Delta Z_1 = \Delta X_1 \int_{t=0}^{1} [2\beta_{11}(X_{10} + t.\Delta X_1) + \sum_{j=2,n} \beta_{1j}(X_{j0} + t.\Delta X_j)]\,dt$$

or

$$= \beta_{11}\Delta X_1(2X_{10} + \Delta X_1) + \Delta X_1 \sum_{j=2,n} \beta_{1j}(X_{j0} + \Delta X_j/2)$$

which is the same as the average of the sequential estimates. This proves the result.

Although few models are actually of quadratic form, a quadratic approximation is often reasonably accurate over the interval through which variables change in a given simulation. When that seems plausible, we have grounds for supposing that our straight-line estimate of the contribution due to a particular shock or group of shocks lies around the middle of the range of the estimates that might be obtained by measuring the same contribution along paths where only one exogenous variable (or group of variables) at a time is allowed to vary.

For future work, we hope to investigate other conjectures. One conjecture is that if F is quadratic, and if perverse paths like those of Diagram D in section 3.5 are excluded, the set of sequential estimates described above includes both the greatest and the smallest estimate of a particular contribution that could be obtained along any reasonable path. We may also be able to replace the quadratic assumption with a weaker assumption – perhaps that derivatives are monotonic within the space under consideration – and still reach conclusions of a similar flavour.

4. Calculating the Decomposition

In practice we don't know the functions $F_1,\ldots,F_m$, in (5) explicitly. Rather the relation between the exogenous and endogenous variables is implicit as in

$$G_j(Z_1, \ldots, Z_m, X_1,\ldots, X_n) = 0 \qquad j = 1, \ldots, m \tag{8}$$

and we must calculate the partial derivatives $(\partial F_j \partial X_i)$ numerically by solving a system of linear equations. Suppose that each function $G_1,\ldots,G_m$ is differentiable.

If we partially differentiate (8) with respect to X_i using the Chain Rule [see, for example, Theorem 6-14 of Apostol (1957)], we see that, for j=1,..., m,

$$\sum_{k=1,m} (\partial G_j / \partial Z_k)(\partial Z_k / \partial X_i) + \partial G_j / \partial X_i = 0$$

since $\partial X_s/\partial X_i = 0$ for $s \neq i$. Hence

$$A\underset{\sim}{V_i} = \underset{\sim}{W_i} \tag{9}$$

where A is the $m \times m$ matrix whose entry in row j and column k is $(\partial G_j/\partial Z_k)$, and $\underset{\sim}{V_i}$ and $\underset{\sim}{W_i}$ and the $m \times 1$ vectors whose entries in row j are $\partial Z_j/\partial X_i$ and $-\partial G_j/\partial X_i$ respectively. This means that, at any point along the curve H from $\underset{\sim}{X_0}$ to $\underset{\sim}{X_1}$, we can calculate the numerical values of the partial derivatives $\partial Z_j/\partial X_i$ (i=1,…,m; j=1,..., m) by solving the system (9) of linear equations numerically.

If we multiply both sides of (9) by dX_i/dt, we see that

$$A\underset{\sim}{v_i} = \underset{\sim}{w_i}, \tag{10}$$

where $\underset{\sim}{v_i}$ and $\underset{\sim}{w_i}$ are the $m\times 1$ vectors whose entries in row j are $(\partial Z_j/\partial X_i)(dX_i/dt)$ and $-\partial G_j/\partial X_i)(dX_i/dt)$ respectively.

This means that, at any point along the curve H from $\underset{\sim}{X_0}$ to $\underset{\sim}{X_1}$ (and for any i=1,...,m), we can calculate the numerical values of the terms

$$(\partial Z_j/\partial X_i)(dX_i/dt) \qquad j=1,\ldots, m \tag{11}$$

by solving the system (10) of m linear equations numerically. But these numbers are exactly those which are needed to calculate the contribution of X_i to the change in Z_j (along the path H) since this contribution is defined to be

$$\int_{t=0}^{1} (\partial Z_j / \partial X_i)(dX_i / dt)\, dt\ . \tag{12}$$

Thus we can obtain arbitrarily accurate approximations to this contribution by various well-known ways of approximating this integral which involve using the system (10) of linear equations N times (for some $N \geq 1$).[10] Indeed, it is possible (and practical) to produce arbitrarily accurate approximations to (12) by extrapolating from two or more less accurate (but less expensive to calculate) approximations, as is well known.[11]

Note that, when solving models represented as in (8), there is usually no *a priori* guarantee of the existence of (suitably smooth) functions $F_1,\ldots,F_m$ as in e uation (5). But, once the numerical calculations have been done, if the results converge as theory suggests, we have *a posteriori* confidence of the existence of the re uired functions and of the accuracy of the model results obtained. The same applies to the numerical decomposition calculated.

4.1. Connection with the Simulation Results

In the simulation, the vector $\underset{\sim}{X}$ of exogenous variables moves from its pre-simulation values $\underset{\sim}{X}_0$ to its post-simulation values $\underset{\sim}{X}_1$. It is common to regard the changes

$$c\_Z_1,\ c\_Z_2,\ \ldots,\ c\_Z_m \tag{13}$$

in the values of the endogenous variables (between their pre-simulation and post-simulation values) as the solution of the simulation.[12]

Here we outline one way of obtaining arbitrarily accurate approximations to the changes in (13) (that is, to the solution of the simulation) which is very similar to the way outlined above of obtaining arbitrarily accurate approximations to the contributions of each exogenous variables $X_1,\ldots,X_n$ to these changes.

Let $\underset{\sim}{v}_0$ denote the m×1 vector whose jth entry is dZ_j/dt. Now, for $1 \leq j \leq m$,

$$dZ_j / dt = \sum\nolimits_{i=1,n} (\partial Z_j / \partial X_i)(dX_i / dt) \tag{14}$$

and thus

---

[10] Consider the integral $\int_0^1 c(t)dt$. Provided we are able to calculate c(t) for any point t between 0 and 1, there are many ways of obtaining arbitrarily accurate approximations to the integral. For example, consider N e ually spread points $0=t_0, t_1,\ldots, t_{N-1}, t_N=1$ (where $t_i = i/N$); then

$$\alpha_N(c) = \sum\nolimits_{i=0}^{N-1} c(t_i)/N \quad (*)$$

converges to the integral as $N \to \infty$. [Calculating N of these approximations to (12) re uires solving the system (10) of linear e uations N times.]

The approximations $\alpha_N(c)$ to $\int_0^1 c(t)dt$ given by e uation (*) above are those obtained by applying Euler's method [see, for example, Section 6.2 of Atkinson (1989)] to the initial value problem

$$d(0) = 0,\ d\_(t) = c(t);\ \text{find } d(1)\ [\text{where } d(t) = \int_{u=0}^{t} c(u)du].$$

Other methods [for example the midpoint method or Gragg's method - see, for example Chapter 15 of Press *et al* (1986)] can be used to obtain more ac curate approximations with comparable amounts of arithmetic. The $\alpha_N(c)$ values are known to converge to the integral provided the underlying function c is suitably well-behaved (see Chapter 6 of Atkinson (1989) or Chapter 15 of Press *et al* (1986) for more details).

[11] For example, the extrapolated value obtained from 4-step, 6-step and 8-step Euler approximations (those with N=4, 6, 8 respectively) is likely to be as accurate than the 100-step Euler approximation. See, for example, Chapter 15 of Press *et al* (1986) for an introduction to extrapolation.

[12] We usually know (or can infer) the pre-simulation values of the Z_j from the pre-simulation data base. If we know the pre-simulation values Z_j and their changes c_Z_j, we can infer their post simulation values.

$$\underset{\sim}{v}_0 = \sum_{i=1,n} \underset{\sim}{v}_i \tag{15}$$

where the $\underset{\sim}{v}_i$ are defined just after equation (10) above. Thus, from (15) and (10),

$$A\underset{\sim}{v}_0 = A\sum_{i=1,n} \underset{\sim}{v}_i = \sum_{i=1,n} A\,\underset{\sim}{v}_i = \sum_{i=1,n} \underset{\sim}{w}_i = \underset{\sim}{w}_0 \tag{16}$$

say. Hence, at any point along the path H, we can calculate the value of dZ_j/dt by solving the system (16) of linear equations numerically. But, as we saw just before equation (3) in section 3.1,

$$c\_Z_j = \int_{t=0}^{1} (dZ_j / dt)\, dt \qquad\qquad j=1,\ldots,m \ .$$

Thus the numbers dZ_j/dt can be used to calculate arbitrarily accurate approximations to c_Z_j following the methods outlined in footnote 1.

Thus

one algorithm for calculating the simulation results

$c\_Z_1, \ldots, c\_Z_m$

is to use the solution to the linear equations (16) to derive arbitrarily accurate approximations to these.

This is essentially the algorithm implemented in GEMPACK. (See section 4.3 of Harrison and Pearson (1996) for more details.) In the GEMPACK solution algorithm, the path H from $\underset{\sim}{X}_0$ to $\underset{\sim}{X}_1$ is taken to be a straight line.

4.2. Calculating the Decomposition in GEMPACK

As indicated previously, GEMPACK takes a straight line path H from $\underset{\sim}{X}_0$ to $\underset{\sim}{X}_1$. For suitably chosen values of N, GEMPACK calculates an approximation to the simulation results in (13) by

- calculating the matrix A at each point

  $t_0=0, t_1, \ldots, t_{N-1}, t_N=1$

  (where $t_s = s/N$),

- calculating the vector $\underset{\sim}{w}_0$ at each point t_s,

- solving the system (16) of linear equations to find the entries of $\underset{\sim}{v}_0$

Then the methods outlined in footnote 1 are used to produce suitably accurate approximations to the simulation results in (13).

To now implement the calculation of the contribution to the changes in any Z_j due to any exogenous variable X_i (or of any group of exogenous variables) is relatively easy conceptually. The steps are as follows.

(a) When calculating $\underset{\sim}{w}_0$, the new algorithm has to also keep track of the separate vectors $\underset{\sim}{w}_1, \ldots, \underset{\sim}{w}_n$ (whose sum equals $\underset{\sim}{w}_0$).

(b) As well as solving (16) at each point t_s, the new algorithm must also solve (10) for each $1 \le i \le n$. The main cost (in terms of time) in solving (16) is that of calculating the LU decomposition of the matrix A. The extra time taken to solve (10) for the extra right-hand sides $\underset{\sim}{w}_1, \ldots, \underset{\sim}{w}_n$ is relatively small compared to the time taken for the LU decomposition. (See, for example, section 8.1 of Atkinson (1989) or section 2.3 of Press *et al* (1986) for information about the cost of LU decomposition.)

(c) The new algorithm must keep track of the solutions $\underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_n$ to (10), as well as the solution $\underset{\sim}{v}_0$ to (16). These $\underset{\sim}{v}_j$ are used to calculate the contribution of each X_i to the changes in each endogenous variable in the same way as the $\underset{\sim}{v}_0$ vectors (at each point $t_0, t_1, \ldots, t_N$) are used to calculate (approximations to) the changes in each endogenous variable.

Thus the extra things GEMPACK must do in order to calculate the contribution of different exogenous variables to the changes in the endogenous variables are as follows:

(i) The software must provide the user with an opportunity to say for which groups (if any) of exogenous variables calculation of the contribution is required.[13]

(ii) Extra bookkeeping is required to keep track of the separate $\underset{\sim}{w}_1, \ldots, \underset{\sim}{w}_n$ as well as $\underset{\sim}{w}_0$.

(iii) Extra CPU time will be taken to solve (10) for each right-hand side $\underset{\sim}{w}_i$ of (10).

(iv) Extra bookkeeping is required to keep track of the (approximations to) the contributions of each group of exogenous variables to each endogenous variable.

In fact, if under (i) above the user identifies k different groups of exogenous variables, there are only k extra systems of linear equations (10) to be solved at each point. These are of the form

$$A\underset{\sim}{p}_r = \underset{\sim}{q}_r \qquad r=1,\ldots,k$$

where $\underset{\sim}{q}_r$ is the sum of the $\underset{\sim}{w}_i$ for all exogenous variables in the rth group of exogenous variables. Thus there are only k vectors to keep track of under (ii), (iii) and (iv) above. Typically the number of groups of exogenous variables the user selects will be small (say 20 or less). Thus the extra CPU time taken to calculate the contributions of these groups is usually a relatively small fraction of the CPU time for the solution without these contributions.

5. The Decomposition for Sequences of Simulations

Suppose that we have a sequence of M different simulations, each one starting from the post-simulation status of the previous one. Forecasts from the MONASH model of the Australian economy (see, for example, Dixon and Rimmer, 1998) comprise such a sequence of linked annual simulations. Suppose, for simplicity, that the exogenous variables are the same in each simulation.

Then it is natural to define the contribution of a group of exogenous variables to the total change (between the starting state of the economy and the ending state, after the final simulation) in each endogenous variable as the sum of the contributions from each simulation. Clearly the main result (Proposition 3.3.1) holds for this sequence of simulations.

[13] Readers familiar with GEMPACK may be interested to know that this information is gathered from "subtotals" statements (such as those currently allowed with the program SAGEM) in the Command file. To obtain numerical contributions of different groups of exogenous variables, users only need to add the relevant "subtotal" statements.

5.1. Decomposition of Percentage-Change Results

Contributions to percentage changes (see section 3.6 above) need to be defined and calculated a little carefully to ensure that they add up as required. The ways of calculating the contributions in this case are easily understood by converting everything back to ordinary changes.

The *contribution C of a group of exogenous variables to the total percentage change p_Z in an endogenous variable Z* is defined to be

$$C = \sum_{s=1,M} A_s C_s$$

where C_s denotes the contribution of this group of exogenous variables to the percentage change p_Z_s in Z in simulation number s, and

$$A_s = \prod_{r=1,s-1}(1 + p\_Z_r/100)\ .$$

With these definitions, it is easy to see that the contributions of a set of mutually exclusive and exhaustive sets of exogenous variables add to the total percentage change p_Z over the sequence of simulations (just as in Proposition 3.3.1).

For example, if M=3, the contribution of a group of exogenous variables to p_Z is given by

$$C_1 \;+\; (1 + p\_Z_1/100)*C_2 \;+\; (1 + p\_Z_1/100)*(1 + p\_Z_2/100)*C_3$$

where C_s denotes the contribution of the group to p_Z in simulation number s.

Thus the decomposition can be carried over to any such sequence of simulations.

6. Conclusion and Further Work

We have described a way of decomposing the endogenous changes from a general equilibrium simulation into parts attributable to each of the exogenous shocks. The decomposition

- is exact (adds up to the right total),
- can be easily generalized to groups of shocks, and percentage (rather than ordinary) changes,
- is easy to understand, and
- is cheap to compute.

We think it promises to be very useful in understanding and presenting results from experiments with multiple shocks. The procedure is now automated in GEMPACK in a version that will be made publicly available in the future. We shall see then how useful other modellers find it.

Further work includes: developing software to facilitate the calculation of decomposed results for sequences of simulations; and developing visual interfaces both to choose a decomposition, and to view its results.

References

APOSTOL, Tom M. (1957), *Mathematical Analysis: A Modern Approach to Advance Calculus*, Addison-Wesley, Reading.

ATKINSON, Kendall E. (1989), *An Introduction to Numerical Analysis*, second edition, Wiley, New York.

DIXON, Peter B. and Maureen T. RIMMER (1998), 'Forecasting and Policy Analysis with a Dynamic CGE Model of Australia', CoPS/IMPACT Working Paper Number OP-90, (available electronically from URL http://www.monash.edu.au/policy/ELECPAPR/op-90.htm).

HARRISON, W. Jill and K.R. PEARSON (1996), 'Computing Solutions for Large General Equilibrium Models using GEMPACK', *Computational Economics,* vol. 9, pp 83-127.

HERTEL, Thomas W. (ed) (1997), *Global Trade Analysis: Modeling and Applications*, Cambridge University Press.

PRESS, W.H., B.P. FLANNERY, S.A. TEUKOLSKY and W.T. VETTERLING (1986), *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press.