GEMPACK: History, How it Works and an Application to New Quantitative Trade Modelling

CoPS Working Paper No. G-357, August 2025

Peter B. Dixon

And

Maureen T. Rimmer

Centre of Policy Studies,
Victoria University

ISSN 1 921654 02 3 ISBN 978-1-921654-66-4

The Centre of Policy Studies (CoPS), incorporating the IMPACT project, is a research centre at Victoria University devoted to quantitative analysis of issues relevant to economic policy.
Address: Centre of Policy Studies, Victoria University, PO Box 14428, Melbourne, Victoria, 8001
home page: www.vu.edu.au/CoPS/ email: copsinfo@vu.edu.au Telephone +61 3 9919 1877

About us

Researchers at the Centre of Policy Studies have a 50-year history of continuous achievement in the development, application and dissemination of large-scale economic models. Our models and software are used around the world to analyse a diverse range of economic issues. CoPS' funders include: Australian federal and state government departments; private firms and universities in many parts of the world; central government agencies such as finance and trade ministries in many countries; and international development organisations. The Centre's GEMPACK software, used for solving large economic models, is used at more than 700 sites in over 95 countries.

Citation

Dixon, Peter, B. and Maureen T. Rimmer (2025), "GEMPACK: history, how it works and an application to New Quantitative Trade modelling" Centre of Policy Studies Working Paper No. G-357, Victoria University, August 2025.

GEMPACK: history, how it works and an application to New Quantitative Trade modelling

***Presented at the 5th Annual Workshop on Bridging the Gap Between CGE and NQT Models,* Beijing, August 13, 2025**

By Peter B. Dixon and Maureen T. Rimmer
Centre of Policy Studies, Victoria University, Melbourne

Abstract

Computable General Equilibrium (CGE) modelling started with the publication in 1960 of Johansen's model of Norway. It continues to the present time as an active research and policy field. In a recent count, there were 33,000 people in the GTAP CGE modelling network alone. This paper identifies GEMPACK software, developed in Australia for solving large scale CGE models in the Johansen school, as one of the factors contributing to the enduring popularity of CGE. The paper tells the story of how GEMPACK came into existence, how it works, and how it relates to Johansen. The paper was prepared for a workshop on bridging the gap between CGE and New Quantitative Trade (NQT) models. In illustrating GEMPACK and showing connections between CGE and NQT, we present a GEMPACK solution and analysis of Eaton and Kortum's seminal NQT model, published in *Econometrica* in 2002.

JEL codes: C68, F11, F13 and C63

Key words: GEMPACK CGE software; GEMPACK and GAMS; New Quantitative trade modelling; Eaton and Kortum

Acknowledgements: We thank Anton Yang and Bob Koopman for inviting us to the 5th Annual Workshop on Bridging the Gap Between CGE and NQT Models. We thank the China University of Political Science and Law, and the University of International Business and Economics for sponsoring the workshop and making our participation financially possible. The workshop was a lively affair, and we thank participants for discussion.

Contents

1. Introduction

CGE started with the publication in 1960 of Johansen's 20-industry model of Norway. This was the first model that encompassed what are now recognized as the essential CGE characteristics: computability in a model identifying multiple optimizing agents whose behaviour is coordinated in markets. The field developed slowly at first. It took off in Australia in the 1970s and has blossomed in the rest of the world since the 1990s. At a recent count, there were 33,000 people in the GTAP CGE modelling network alone.

There are 3 factors that explain the endurance and acceleration of CGE modelling. First, it is an adaptable tool for analysing many issues in trade, environment, disasters, labour markets, population, immigration, income distribution, technology, resources and public finance. Second, it has been brilliantly served by the international data and model-building effort lead and coordinated by Tom Hertel and his colleagues at the Global Trade Analysis Project (GTAP). Third, it has been democratized by the development of two CGE software packages: GEMPACK and GAMS. These packages have made it possible for economists without specialist computer expertise to apply state-of-the-art models, and to incorporate new theory for tackling new problems.

This paper tells the story of GEMPACK. Section 2 explains how GEMPACK came into existence, how it works, and how it relates to Johansen (1960). Section 3 sets out an illustrative GEMPACK application. For the illustration, we chose Eaton and Kortum's 2002 model which is highly sighted in NQT literature. This is a suitable choice for a conference devoted to bridging the gap between CGE and NQT. It was also a chance for us to further our own education. As well as concluding remarks on GEMPACK, section 4 exhibits our current level of understanding of the EK model. We set out an interpretation of the relationship between the EK and Armington (1969) specifications of trade.

2. GEMPACK software: brief history, theory and how it works

GEMPACK was designed to solve models that can be written as

$$F(V) = 0 \tag{2.1}$$

where F is a vector of M differentiable functions and V is a vector of R variables, with R > M.

In most GEMPACK applications the R variables all refer to single year, t. The variables include prices and quantities for year t, start-of-year stocks for year t, and end-of-year stocks for year t. The stock variables can include values of physical capital and values of financial assets and liabilities. The M equations specify relationships between year-t variables such as prices equal costs, demands equal supplies, demands reflect optimizing decisions, and end-of-year stocks equal start-of-year stocks plus relevant flow variables.

A critical point about (2.1) is that we have an initial solution, that is a vector V(0) that satisfies

$$F(V(0)) = 0 \tag{2.2}$$

GEMPACK operates by calculating new solutions as deviations from an initial solution.

In moving from an initial solution to a new solution, we use a version of the equation

$$F_V(V(0)) * dV = 0 \tag{2.3}$$

where $F_V(V(0))$ is the M-by-R matrix of partial derivatives of the F functions evaluated at the initial solution and dV is an R-by-1 vector of deviations in the variables away from their initial values. The idea underlying (2.3) is that if we have an initial solution, then in going to a new solution we should continue to satisfy (2.1) by keeping the value of the F vector on zero.

In most GEMPACK applications, (2.3) is transformed largely into a percentage deviation format:

$$A(V(0)) * v = 0 \tag{2.4}$$

where $A(V(0))$ is an M-by-R matrix composed mainly of cost shares, sales shares, elasticity parameters and 1's and 0's.[1] The R-by-1 vector v is composed mainly of percentage deviations in variables away from their initial values, but for some variables, the entry in the v vector continues to be a change, rather than a percentage change. Unit-free percentage deviations are convenient, but not appropriate for variables such as the balance of trade that can pass through zero.

To solve the model, we need to specify a closure, that is, we need to nominate M variables to be endogenous and R-M variables to be exogenous. Then GEMPACK partitions the $A(V(0))$ matrix into two submatrices. The first [$A_1(V(0))$] is the M-by-M matrix formed from the columns of $A(V(0))$ corresponding to the endogenous variables and the second [$A_2(V(0))$] is the M-by-(R-M) matrix formed from the columns of $A(V(0))$ corresponding to the exogenous variables. With this partition, (2.4) can be written as:

$$A_1(V(0)) * v_1 + A_2(V(0)) * v_2 = 0 \tag{2.5}$$

where v_1 and v_2 are the vectors of percentage deviations (or deviations) in the endogenous and exogenous variables.

From (2.5) GEMPACK derives the solution

$$v_1 = B(V(0)) * v_2 \tag{2.6}$$

where

$$B(V(0)) = -A_1^{-1}(V(0)) * A_2(V(0)) \tag{2.7}$$

$B(V(0))$ has the dimensions M-by-(R-M). The i,jth element shows the sensitivity (usually an elasticity) of the ith endogenous variable with respect to variations in the jth exogenous variable.

Equations (2.6) - (2.7) are the solution presented by Johansen in his 1960 book. In Johansen's model, there are 86 endogenous variables and 46 exogenous variables. He displayed his B matrix in a foldout and devoted much of the book to interpreting it. He regarded its 3,956 entries (= 86*46) as the model's fundamental results and used them in discussing issues of importance such as whether growing demands for food in Norway, combined with land scarcity, would require increased employment in agriculture, or whether technical progress in agriculture would dominate leading to reduced employment in agriculture.

---

[1] The transformation from a levels form such as (2.2) to a deviation form in percentage change variables is discussed in section 2.3.

Another fascinating aspect of Johansen's B matrix was what it said about multiplier effects. In the 1950s when Johansen was creating his model, Leontief's input-output model (Leontief, 1936) was the dominant vehicle for working out the effects of exogenous demand increases in one industry on outputs and employment in other industries. In the Leontief model these effects were always positive. But Johansen's B matrix indicated that the results were predominantly negative. In Leontief's model there were no resource constraints. Labour and capital were under-used, a reasonable assumption for most industries in the 1930s when Leontief formulated his model. Johansen's model incorporated resource constraints. This was appropriate for the booming 1950s. It meant that exogenous expansion of one industry could have negative effects on other industries by depriving them of scarce labour and capital.

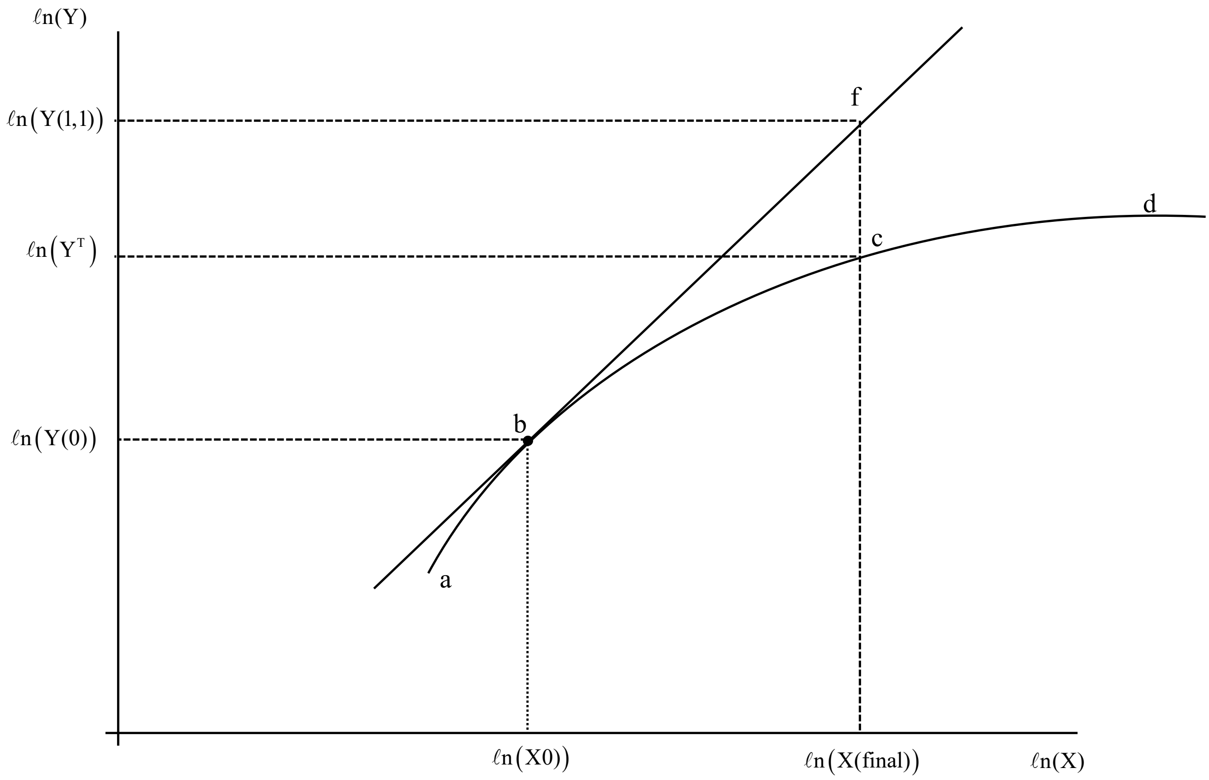
Johansen's techniques were simple and effective, but they came at a cost. His solution equations, (2.6) – (2.7), give only an approximation to the solution of the underlying non-linear model. This is illustrated in Figure 2.1 for a 1-equation 2-variable model. In the figure we denote the exogenous variable as X and the endogenous variable as Y. The logarithms of the exogenous and endogenous variables, $\ell n(X)$ and $\ell n(Y)$, are on the horizontal and vertical axes. The curve *ad* shows the relationship between the endogenous and exogenous variables implied by the non-linear model. The initial solution is at point *b*. The relationship implied by Johansen's equations (2.6) - (2.7) is the straight line *ef*. This line is tangent to the non-linear relationship at the initial solution. On Johansen's line, the elasticity of the endogenous variable with respect to the exogenous variable is fixed at its value at the initial solution. In a Johansen computation, a movement in the logarithm of the exogenous variable from $\ell n(X(0))$ to $\ell n(X(\text{final}))$ causes a movement in the endogenous variable from $\ell n(Y(0))$ to $\ell n(Y(1,1))$ where (1,1) denotes value at the end of 1 step of a 1-step computation. The true movement, that is the movement implied by the non-linear model is from $\ell n(Y(0))$ to $\ell n(Y^T)$. The linearization error is the gap between $\ell n(Y(1,1))$ and $\ell n(Y^T)$.

Johansen's work was largely ignored for 15 years. CGE modellers were keen to avoid linearization errors and perhaps this caused them to overlook the strength of Johansen's approach. Sustained development of Johansen's style of CGE modelling was not undertaken until work was started in 1975 on the ORANI model of Australia.[2]

ORANI was built for the Australian government's Industries Assistance Commission at the IMPACT Project directed by Alan Powell. He recruited a team of modellers led by Peter Dixon to build the model. The ORANI team was presented with a sharp policy question: *where would the jobs come from to re-employ people who would lose their jobs if Australia transitioned from high protection to low protection*? The government wanted an answer with lots of detail, lots of industries, lots of regions and lots of occupations.

[2] Starting in the mid-1970s, there were some one-off applications of Johansen's technique, see Taylor and Black (1974), Staelin (1976), Bergman (1978), and Keller (1980).

Figure 2.1. A Johansen solution



The ORANI model had an unprecedented amount of detail: over 100 industries 6 regions and 9 occupations. It distinguished between purchasers' prices and producer prices and modelled the taxes, and transport, retail and wholesale margins in between. It included technology and preference variables associated with each commodity flow. It allowed for multi-product industries and incorporated many econometrically estimated transformation and substitution elasticities. In raw form the dimensions of the model were huge: several million equations and variables. This was manageable because the model was formulated in the linear form (2.4).

Using the linear form, the dimensions of the computational model could be sharply reduced by substituting out equations and variables. For example, consider the equation

$$x(i,s,j,k,m) = z(j,k) + a(i,s,j,k,m) \qquad (2.8)$$

where

x(i,s,j,k,m) is the percentage change in the use of margin m to facilitate the flow of commodity i from source s (domestic, imported) to industry j for purpose k (current production or capital creation);

z(j,k) is the percentage change in j's production or capital creation; and

$a(i,s,j,k,m)$ is the percentage change in technology associated with the use of margin m to facilitate flow (i,s,j,k).

With 100 commodities, 2 sources, 100 industries, 2 purposes and 10 margin commodities, (2.8) is a block of 400,000 equations. In a linear system these equations and the LHS variable can be eliminated easily by replacing x(i,s,j,k,m) wherever it appears in the model by the RHS of (2.8). No detail is lost in this procedure. The specification of margin demands in

(2.8) continues to influence model results as intended and if results are required for the eliminated high-dimensional variables, then these can be recovered post simulation by back-solving. Substitution, back-solving and omission of inactive exogenous variables together with the use of Johansen's solution technique allowed the ORANI modellers to produce policy relevant results in less than two years, Dixon *et al.* (1977).

While substitution and back-solving were tedious procedures in the 1970s and 80s, they have subsequently become automated and streamlined. This means that CGE models in the Johansen tradition can be very large before computing considerations become a limiting factor.

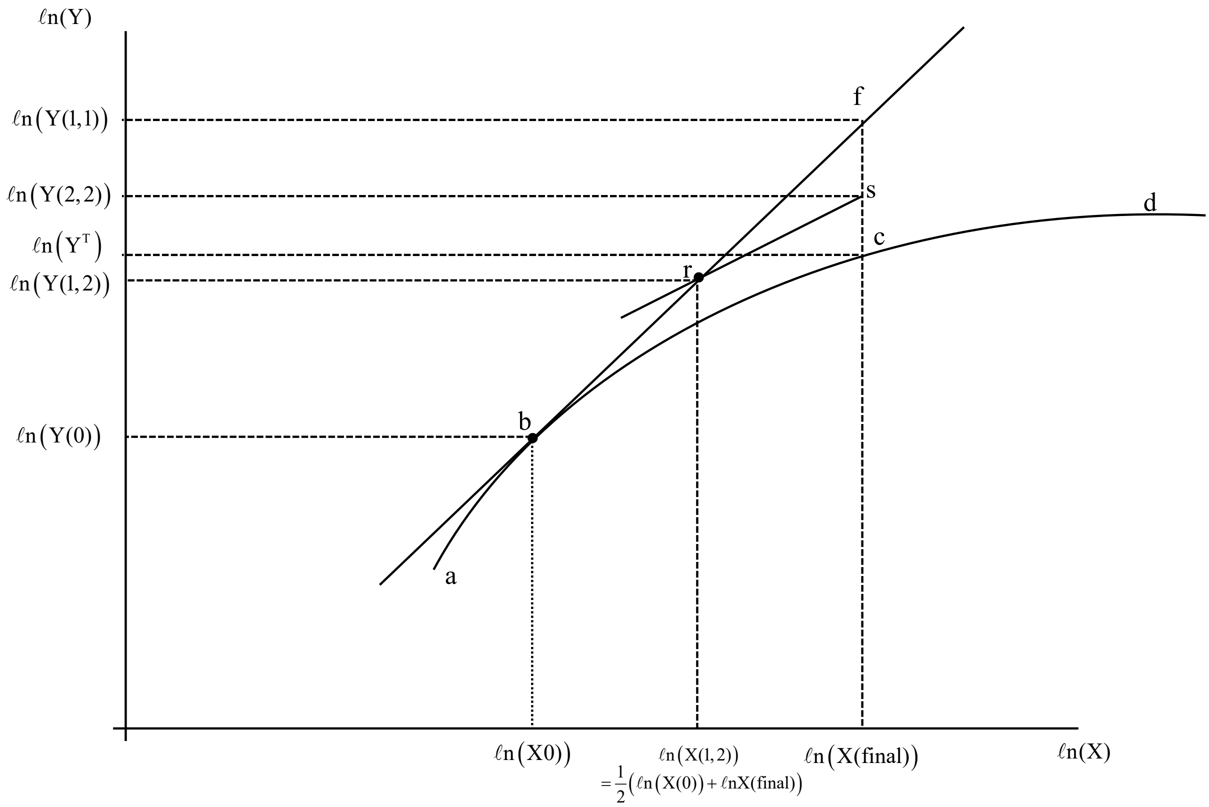
A bonus of the Johansen technique was flexible closures. Johansen used just one closure. But it became apparent that by regarding the model as the rectangular system (2.4), the closure decision, and the consequent partition of the A matrix, need not be hard-wired into the model. It could be tailor-made to each application. For example, in simulations concerned with the short run, capital stocks by industry and wage rates could be exogenous with zero movements. For the long run, the closure could be switched with capital stocks and wage rates becoming endogenous and rates of return and employment going exogenous. The ORANI model could be used to work out the effects on employment and the trade balance (endogenous variables) of given reductions in real wage rates and stimulation of government expenditure (exogenous variables). Alternatively, in the spirit of Tinbergen's instruments and targets, the model could be used to calculate the wage reductions and expenditure stimulation required (endogenous variables) to achieve targets for employment and the trade balance (exogenous variables). This instrument-target approach was an input to the Australian government's Crawford Inquiry into the industrial problems associated with a depressed economy, see Dixon *et al.* (1979).

However, the problem of linearization errors remained. In policy circles it didn't matter, but it adversely affected the academic acceptability of the model.

Solving the linearization problem while retaining the simplicity of the Johansen approach turned out to be surprisingly simple. What was required was the implementation of a sequence of Johansen solutions. To simulate the effects of a 50 per cent tariff increase, the ORANI modellers started by simulating the effects of a 22.47 per cent increase. They then paused and updated the entire database for the model to reflect the new situation with the 22.47 per cent tariff increase in place. The main task in the update was to move the values of all the input-output flows to their new positions taking account of the price and quantity movement for each cell. Starting from the new database the modellers imposed another 22.47 per cent tariff increase. By comparing the solution after two applications of 22.47 per cent tariff increases with the initial solution, the modellers were able to deduce the effects of a 50 per cent tariff increase ($1.2247^2 = 1.50$) calculated in a 2-step Johansen calculation.

Figure 2.2 shows how the introduction of a second step can sharply reduce the linearization error. In the first step of the 2-step procedure, the solution moves from (X(0),Y(0)) to (X(1,2),Y(1,2)), that is, from *b* to *r*. At the end of the first step, the database is updated and the elasticity of Y with respect to X is evaluated at point *r*. This is the slope of the line *rs* and is an approximation to the slope of the curve *ad* with X at X(1,2). Using the approximate slope, the second step calculates the effect of moving X from X(1,2) to X(final). In this step Y moves from Y(1,2) to Y(2,2). This is the 2-step solution.

Figure 2.2. A 2-step Johansen-Euler solution



Dixon *et al.* (1982, ch. 5) proved that as the number of steps increase in a multi-step Johansen computation, the solution in the final step converges to the exact solution for the non-linear model.[3]

Each step in a multistep Johansen computation requires inversion of a large matrix, and the creation of a B matrix, see (2.6) and (2.7). Convergence would not have been of practical value in the 1980s if it had required a large number of steps. However, the ORANI modellers soon stumbled onto Richardson's extrapolation. This was a well-known idea in books on numerical methods.[4] These books showed that what the ORANI modellers were doing was applying the Euler method to solve a system of differential equations and that in applications of this method, doubling the number of steps often approximately halves the linearization error, that is

$$\left(y(2,2) - y^T\right) \approx 0.5 * \left(y(1,1) - y^T\right) \tag{2.9}$$

where y(2,2) and y(1,1) are the vectors of deviations in the logarithms of the endogenous variables in 2-step and 1-step Johansen computations, and y^T is the vector of true log deviations. In terms of Figure 2.2, the 2-step error, *sc*, is approximately half the 1-step error, *fc*. Applying (2.9), the ORANI modellers found that highly accurate solutions could be computed with 1-step and 2-step Johansen computations supplemented by an extrapolation

$$y(1,1\,;2,2)^{extrap} = 2 * y(2,2) - y(1,1) \tag{2.10}$$

[3] Reassuringly for sceptics, Hertel *et al,* (1992) confirmed that the multistep Johansen procedure produces the same solutions as GAMS working with levels equations.
[4] See for example, Dahlquist *et al.* (1974, p.269).

where $y(1,1;2,2)^{extrap}$ is the solution derived by extrapolation from the 1-step and 2-step solutions.

Ex-post, it was not too surprising that extrapolations from low-step Johansen solutions produced highly accurate solutions to CGE models. Apart from small errors in the evaluation of the B matrices through the multistep procedure, it can be shown that the extrapolation in (2.10) is exact if the underlying reduced form (relationship between the endogenous and exogenous variables) is quadratic. Since linear in log variables is often a reasonable approximation to the reduced form, quadratic in log variables is often a very good approximation.

In recognition of Johansen and Euler, the ORANI modellers subsequently referred to their multistep approach to solving CGE models as the Johansen-Euler method.

2.1. The emergence of GEMPACK

The initial program using the Johansen-Euler method was tailor-made for the ORANI model, see Dixon *et al.* (1982, ch. 5). Recognition that the program could be generalized and made useful to other modellers required the entry of Ken Pearson to the story.[5]

Ken was a member of the Mathematics department at La Trobe university. As well as being a highly creative researcher, he was a conscientious teacher. In 1982 he was looking for applications of mathematics with which to interest his undergraduate students beyond those usually drawn from physics. He heard about the ORANI modelling team, part of which was located in the Economics department at La Trobe. He invited Peter Dixon to present a seminar in the Mathematics department.

Ken sensed that what was going on with ORANI was important. ORANI was gaining political traction and was being used effectively at the Industries Assistance Commission in calculating the effects of lower tariffs. Ken knew that the modelling team was keen to facilitate wider use of ORANI but that this was inhibited by computational complexity. While there was no problem with the Johansen-Euler strategy, Ken's insight was to see that dissemination required user-friendly software at the front and back ends, that is, software that didn't require specialist computing skills. For the front end, he set about writing the TABLO language through which a model could be presented in deviation form to the computer in a text file which looks like normal algebra. For the back end, Ken started the process of creating what has turned into an ever-improving suite of auxiliary programs (discussed below) for assisting modellers in interrogating their results.

Ken was enthusiastically backed by IMPACT's director, Alan Powell. Large-scale, policy-relevant economic models contain many thousands of variables and non-linear equations. Making them computationally efficient and widely accessible must have seemed a monumental task. Embarking on it was an intellectually courageous decision. For Ken, it required a break from his familiar world of mathematics at considerable risk to his burgeoning career. But he succeeded. The outcome was the GEMPACK software.

The first version of GEMPACK was unveiled at a training course on the ORANI model for public servants and academics held in 1984. Over the next 30 years, Ken continuously developed and improved GEMPACK, working with several collaborators, most notably Mark Horridge and Jill Harrison. In recent years, key contributors to GEMPACK have been Michael Jerie, Florian Schiffmann and Dean Mustakinov.

[5] Ken died in 2015. The next four paragraphs are drawn from Ken's obituary, see Dixon and Rimmer (2015).

GEMPACK is now used in about 700 sites in 95 countries. It was the original choice for the GTAP model[6] and remains popular in GTAP's worldwide network of CGE modellers. In chapter 20 of Elsevier's Handbook of Computable General Equilibrium Modeling published in 2013, the proprietors of GEMPACK and its main competitor, GAMS, conducted a computational comparison between the two packages.[7] GEMPACK was the overwhelming winner, see Figure 2.3.

Since 2013, there has been a reduction of about 70% in solution times for GEMPACK, see https://www.copsmodels.com/gp121.htm . The essentials of the multi-step Johansen-Euler procedure remain, but Richardson's extrapolation methods have been supplemented with other integration techniques. The Runge-Kutta technique is proving particularly effective.[8] With faster solution times, CGE projects that were previously unthinkable have become feasible. For example, GEMPACK was used in a recent project that required 150,000 solutions of a 23 industry, 30 region model, see Dixon *et al.* (2024).

Just as important as lightening fast solving times, GEMPACK provides an ever-improving suite of auxiliary programs. These include:

AnalyzeGE. After performing a simulation, this program allows the modeller to evaluate any variable or coefficient appearing in the algebraic (TABLO) representation of the model via *point and click* access. The evaluation can be pre- or post-simulation.

Viewsol. This program allows modellers to view several solutions at once, even solutions from different models. This is invaluable in comparing solutions in the process of developing models and displaying their properties.

RunDynam. This program assists in building multi-period baseline and policy simulations. The program can display the policy-induced deviations between the two. The program includes convenient zip facilities for storing and sharing multi-period baseline/policy simulations.

Viewhar. This is a program for browsing, extracting and modifying the contents of GEMPACK data in Header Array files

Splitcom. This is one of several GEMPACK tools designed to help modellers disaggregate commodities and industries, e.g. split the motor vehicle commodity/industry into assembly, engines, tires, etc.

2.2. Converting from levels to deviation form

When presented with the evidence, many CGE modellers who do not use GEMPACK are willing to concede that GEMPACK is a wonderful suite of programs. But they are worried about the process of converting a model expressed in levels into a deviation form suitable for GEMPACK. In most cases, conversion can be achieved by application of the simple rules in Table 2.1. Nevertheless, the developers of GEMPACK have included a feature for symbolic differentiation. This allows users to submit equations in levels form and have the program

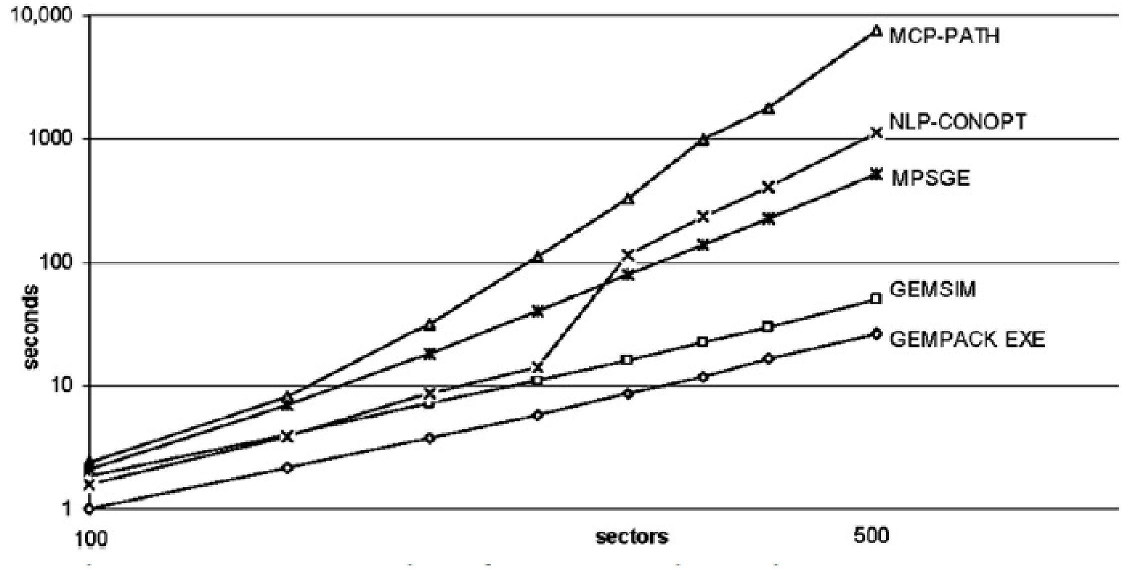
---

[6] See Hertel (1997).

[7] See Horridge, Meeraus, Pearson, and Rutherford (2013).

[8] With this method, we estimate the B matrix at a mid-point between the initial solution, (X(0),Y(0)) and the true solution (X(final),Y^T). In the simplest case, the estimate is $B(RK) = 0.5 * \left(B(X(0), Y(0) + B(X(final), Y(1,1)) \right)$. Then the solution for the effect of moving X from X(0) to X(final) is computed as $y = B(RK) * x$ where x and y are the vectors of percentage deviations in the exogenous and endogenous variables.

*Figure 2.3. Solution time versus number of sectors: Log-log scale**



* The number of sectors in a standard CGE model was increased in steps from 100 to 500. Solution times were recorded for 5 solution methods. GEMPACK.EXE and GEMSIM are GEMPACK methods. The other three are GAMS methods.

produce an algebraic deviation form in TABLO language. We don't use this feature in our own work. We think that there is a lot to be learnt from deriving and interpreting deviation forms. In any case, situations arise in our modelling in which only a deviation form can be specified. No explicit levels form is available.

To illustrate how deviation forms in GEMPACK can facilitate interpretation and cope with implicit functional forms, we work through four examples. Examples 1 and 2 are on the derivation of demands for inputs to production by industries. Examples 3 and 4 are on the derivation of consumer demand functions.

Input-demand equations implied by constrained cost minimization

The cost minimizing problems used in CGE models imply input demand functions that can be expressed conceptually in levels format as:

$$X_i = G_i(Z) * F_i(P) \tag{2.11}$$

where

X_i is an industry's demand for an input i;

Z is the production level for the industry;

P is the vector of prices faced by the industry;

F_i is a homogeneous function of degree zero (doubling all prices holding output constant does not affect demands); and

G_i is an increasing function of output (under constant returns to scale, $G_i(Z)$ is simply Z).

Table 2.1. Three rules for converting from levels equations to deviation equations[*]

	Levels	Deviation version
Multiplication rule	**X= Y*Z**	**x = y + z**
Power rule	$\mathbf{X = Y^{\alpha}}$	**x = α*y**
Addition rules	**X = Y + Z**	**X*x = Y*y + Z*z**
		or $\mathbf{x = \left(\frac{Y}{X}\right) * y + \left(\frac{Z}{X}\right) * z}$ or $\mathbf{100 * dx = Y * y + Z * z}$

* x, y and z are percentage changes in the variables denoted by the corresponding upper-case letters. α is a parameter.

As CGE modellers have striven for increasingly rich specifications of scale and price-induced substitutions, they have adopted increasingly complex forms for production functions. This has led to increasingly complex G_i and F_i functions. In some cases, F_i has no explicit form.

In a model expressed in deviation form, complexities are largely avoided. Instead of (2.11), we write:

$$x_i = E_i * z + \sum_j \eta_{ij} * p_j \tag{2.12}$$

where

x_i, z and p_j are percentage deviations from the base values in X_i, Z and P_j;

E_i is the scale elasticity; and

η_{ij} is the elasticity of demand for i with respect to changes in the price of j.

E_i and η_{ij} are not parameters. No matter how complex the production function might be, modellers can always derive formulas for E_i and η_{ij}. Using these formulas in GEMPACK, modellers can update E_i and η_{ij} through the multi-step Johansen-Euler procedure. In this way, they can achieve an accurate solution (free of linearization errors) whatever the specification of the production function. A couple of examples will clarify.

Example 1. A CES production function

The CES function, due to Arrow *et al.* (1961), is the most used functional form in CGE modelling. It takes the form

$$\left[\sum_{i=1}^{n} \left(X_i\right)^{\frac{\sigma-1}{\sigma}} * \delta_i\right]^{\frac{\sigma}{\sigma-1}} = Z \tag{2.13}$$

where

δ_i and σ (the substitution elasticity) are positive parameters.

Cost-minimization subject to (2.13) produces input demand functions whose elasticities are given for all i by :

$$E_i = 1;\ \eta_{ij} = \sigma * S_j \text{ for all } j \neq i; \text{ and } \eta_{ii} = -\sigma * (1 - S_i) \tag{2.14}$$

where

S_j is the share of costs accounted for by input j, that is

$$S_j = \frac{P_j * X_j}{\sum_k P_k * X_k} \tag{2.15}$$

We can use (2.15) to update cost shares between steps of a Johansen-Euler procedure. Then we can use (2.14) to update the elasticities in (2.12).

Instead of having (2.12) as part of the model, we could substitute from (2.14) into (2.12) to obtain the intuitively appealing form:

$$x_i = z - \sigma\left[p_i - \sum_j S_j * p_j\right] \tag{2.16}$$

In (2.16), a one per cent increase in output will cause a one per cent increase in all input demands. But an industry's demand for i will fall relative its demand for other inputs if the percentage increase in P_i is greater than a weighted average of the percentage increases in the prices of all inputs with the weights being shares in total costs. The sensitivity of input demand for i with respect to relative price changes is controlled by the substitution parameter, σ. In each Johansen-Euler step, GEMPACK records what happens to P_j and X_j. With this information, GEMPACK updates the cost shares, S_j for all j.

Example 2. A CRESH production function

This is a generalization of CES due to Hanoch (1971). Under CRESH, the production function takes the form

$$\sum_{i=1}^{n}\left(\frac{X_i}{Z}\right)^{\frac{\sigma_i - 1}{\sigma_i}} * \delta_i = 1 \tag{2.17}$$

where σ_i and δ_i are parameters. With this production function, the cost minimization problem has no explicit solution, that is the demand functions can't be written as in (2.11) with an explicit functional form for F_i. Nevertheless, it is not hard to show that for all i

$$E_i = 1;\ \eta_{ij} = \sigma_i * \hat{S}_j \text{ for all } j \neq i; \text{ and } \eta_{ii} = -\sigma_i * (1 - \hat{S}_i) \tag{2.18}$$

where $\hat{S}_j$ is a modified cost share defined by

$$\hat{S}_j = \frac{S_j \sigma_j}{\sum_k S_k \sigma_k} \tag{2.19}$$

Between steps of a Johansen-Euler procedure, (2.19) can be used to update the modified cost shares and then (2.18) can be used to update the elasticities in (2.12).

Alternatively, we could replace (2.12) by:

$$x_i = z - \sigma_i \left[p_i - \sum_j \hat{S}_j * p_j \right] \qquad (2.20)$$

The interpretation of (2.20) is similar to that of (2.16). However, (2.20) gives extra flexibility with respect to price sensitivity: the σ parameter has an i subscript. Another difference between (2.20) and (2.16) is the modification of the cost shares.

Consumer-demand equations implied by constrained utility maximization

The utility maximizing problems used in CGE models imply consumer demand functions that can be expressed conceptually as:

$$X_i = H_i\left(M, P\right) \qquad (2.21)$$

where

X_i is consumer demand for i;

M is consumer total expenditure or budget

P is the vector of prices faced by consumers;

H_i is a homogeneous function of degree zero in M and P (doubling the budget and all prices does not affect demands).

As with production functions, CGE modellers have adopted increasingly complex utility functions. This has led to increasingly complex H_i functions, sometimes without explicit form.

Instead of (2.21), in a deviation model we write:

$$x_i = E_i * m + \sum_j \eta_{ij} * p_j \qquad (2.22)$$

where

x_i, m and p_j are percentage deviations from base values in X_i, M and P_j;

E_i is the expenditure elasticity for i; and

η_{ij} is the consumer elasticity of demand for i with respect to changes in the price of j.

Modellers can always derive formulas for E_i and η_{ij}. In a multi-step Johansen-Euler procedure, they can use these formulas in GEMPACK to update E_i and η_{ij}. Whatever the specification of the utility function, they can achieve accurate solutions for consumer demands. We give two clarifying examples.

Example 3. A linear expenditure system (LES)

LES is attributed to Klein & Rubin (1948-9), Geary (1950-1) and Stone (1954). It is popular in CGE modelling because the demand functions are simple but have sufficient flexibility to allow for expenditure elasticities that differ from one. The underlying utility function takes the form:

$$U = \sum_i \beta_i * \ell n(X_i - \gamma_i) \qquad (2.23)$$

where

β_i and γ_i are parameters. β_i is positive and $\sum_i \beta_i = 1$. γ_i is usually interpreted as a subsistence level of consumption of commodity i.

Maximization of (2.23) subject to the budget constraint generates consumer demand functions of the form:

$$X_i = \gamma_i + \frac{\beta_i}{P_i} * \left(M - \sum_j P_j * \gamma_j \right) \qquad (2.24)$$

Equation (2.24) is readily interpretable. It shows that consumption of i has two parts: subsistence (γ_i) and luxury or discretionary. The parameter β_i is the marginal budget share. It determines the fraction of additions to luxury expenditure that are devoted to commodity i.

While (2.24) could be used directly in a CGE model, in GEMPACK it can be replaced by (2.22) with the elasticities determined for all i by:

$$E_i = \frac{\beta_i}{S_i}; \quad \eta_{ij} = -E_i * SS_j \text{ for all } j \neq i; \text{ and } \eta_{ii} = -E_i * SS_i - E_i * LS \qquad (2.25)$$

where

S_i is the share of commodity i in consumer expenditure defined by

$$S_i = \frac{P_i * X_i}{M} \qquad (2.26)$$

SS_j is the subsistence share of commodity j in expenditure defined by

$$SS_j = \frac{P_j * \gamma_j}{M} \qquad (2.27)$$

and LS is the luxury share in expenditure defined by

$$LS = \frac{M - \sum_j P_j * \gamma_j}{M} \qquad (2.28)$$

Between steps of a Johansen-Euler procedure, (2.26) - (2.28) can be used to update the various shares and then (2.25) can be used to update the elasticities in (2.22).

Example 4. An Implicitly Directly Additive Demand System (AIDADS)

AIDADS is due to Rimmer and Powell (1996). It generalizes LES by giving the expenditure elasticities additional flexibility to embrace variations over the wide range of per capita income levels observed in international data. The utility function for AIDADS takes the indirect form:

$$\sum_i U_i(X_i, U) = 1 \qquad (2.29)$$

where the U_i functions are given by

$$U_i = \Phi_i * \ell n\left(\frac{X_i - \gamma_i}{A * e^U} \right) \qquad (2.30)$$

$$\Phi_i = \left(\frac{\alpha_i + \beta_i * e^U}{1 + e^U} \right) \qquad (2.31)$$

and

α_i , β_i , γ_i and A are parameters with $\sum_i \alpha_i = \sum_i \beta_i = 1$. As with LES, γ_i is usually interpreted as a subsistence level of consumption of commodity i.

For any given level of utility, U, we assume that consumers choose X_i for all i to minimize total expenditure M where $M = \sum_j P_j * X_j$. This leads to a demand system that has the form

$$X_i = \gamma_i + \frac{\Phi_i}{P_i} * \left(M - \sum_j P_j * \gamma_j \right) \tag{2.32}$$

Unlike (2.24), in (2.32) the marginal budget shares, Φ_i, are variables whose values depend on utility, which in turn depends on the budget and prices. Utility can only be specified implicitly. Nevertheless, at the expense of some algebra, the demand elasticities in AIDADS can be derived for all i as:

$$E_i = \frac{\Phi_i}{S_i} * (1 + C_i); \quad \eta_{ij} = -\frac{\Phi_i}{S_i} * (SS_j + C_i * S_j) \text{ for all } j \neq i; \quad \text{and} \quad ; \tag{2.33}$$
$$\eta_{ii} = -\frac{\Phi_i}{S_i} * (SS_i + C_i * S_i) - \frac{\Phi_i}{S_i} * LS$$

where SS_j, S_j and LS continue to be defined by (2.26) – (2.28) and

$$C_i = \frac{1}{D} * \left(\frac{e^U}{1 + e^U} \right) * \left(\frac{\beta_i}{\Phi_i} - 1 \right) \tag{2.34}$$

with

$$D = 1 + \left(\frac{e^U}{1 + e^U} \right) * \left(1 - \sum_k \beta_k * \ell n \left(\frac{X_k - \gamma_k}{A * e^U} \right) \right) \tag{2.35}$$

Given the initial solutions for U, Φ_i and X_i for all i, together with the parameters β_i , γ_i and A, GEMPACK can use (2.34) and (2.35) to obtain initial values for C_i and D. After some more algebra, we can derive an equation that shows how U moves in response to changes in prices and total expenditure:

$$dU = 0.01 * \frac{1}{D} * \frac{1}{LS} * \left(m - \sum_k S_k * p_k \right) \tag{2.36}$$

where dU is the deviation in U.

We include (2.36) as part of the GEMPACK equation system. Between steps in a multistep Johansen-Euler procedure, the solution for dU can be used to update U. The updated value for U, together with the updates for the X_is can used to obtain updated values for D, C_i and Φ_i. Updated elasticity values can now be computed from (2.33). These can be used in (2.22). At this stage, GEMPACK is ready to move to the next step of the multistep procedure.

3. An EK model in GEMPACK

People wanting to become familiar with GEMPACK can find many examples of GEMPACK code for relatively simple uncluttered models[9]. Almost all these models have conventional production and utility specifications and usually use the Armington assumption for trade. This section provides GEMPACK familiarization, but with a different sort of trade model.

[9] See for example section 60 of the GEMPACK manual, available at file://C:/GP/gpmanual.htm#models .

We use the model from Eaton and Kortum (2002). This is the famous EK model which is central in new quantitative trade modelling. Bekkers *et al.* (2023) have already demonstrated GEMPACK in a model incorporating EK features. Their model is a full-scale policy-relevant version of GTAP. Here we are less ambitious. Our aim is to illustrate GEMPACK in a form that is quickly accessible to people with an EK background.

In subsection 3.1, we list and briefly explain the EK equations in levels format. In subsection 3.2 we set out the equations in deviation format suitable for GEMPACK. We also discuss two other issues important for GEMPACK, the initial solution and closures. Subsection 3.3 contains illustrative simulations. Our GEMPACK code is in the Appendix.

3.1. The EK model in levels: column 1 of Table 3.1

Table 3.1 lists the equations for our version of the EK model. Tables 3.2 and 3.3 define the notation.

The first eleven equations in the levels column of Table 3.1 are not specific to EK. They could occur in any model in which the following simplifying assumptions are applied: tariffs are the only taxes; labour is the only primary factor; production in country i includes the delivery to users (the iceberg assumption); production is governed by Cobb-Douglas functions of labour and bundles of intermediate inputs; and the bundles of goods used in country i by industries as intermediate inputs and by final users have the same commodity composition.

Equation (T3.1.1) in Table 3.1 defines the balance of trade for country i (BT_i). The first term on the RHS is the sum over all countries n of their expenditures on goods sourced from i, excluding tariffs imposed by n. The second term is the sum over all countries n of their sales to country i excluding tariffs imposed by i. The diagonal term, X_{ii}/TAX_{ii} cancels out on the RHS of (T3.1.1) leaving the balance of trade as the cif value of i's exports less the cif value of its imports.

Equation (T3.1.2) calculates expenditure by final users in country n (Y_n). This is total expenditure in n calculated by adding n's expenditures on goods from all sources less n's expenditures on intermediate inputs. Expenditure on intermediate inputs is a fixed fraction, (1-β), of the cost of production in n (the Cobb-Douglas assumption). Introducing the zero-profit assumption, the cost of production in n is calculated as the value of expenditures on n's products, excluding tariffs, summed over all countries.

Equation (T3.1.3) again reflects the Cobb-Douglas assumption. It gives the value of labor input (wage rate times employment) in country n as a fixed fraction, (β), of the cost of production in n.

Equation (T3.1.4) defines the share (Π_{in}) of country i in expenditures in country n.

Equation (T3.1.5) defines the world price level, PWORLD, as a geometric average of price levels over all countries.

Equation (T3.1.6) defines the real trade balance for each country.

Table 3.1. Equations in the EK model: levels and deviations

	1. Levels	2. Deviation version suitable for GEMPACK	No. of equations
(T3.1.1)	$BT_i = \sum_n \frac{X_{in}}{TAX_{in}} - \sum_n \frac{X_{ni}}{TAX_{ni}}$	$100*d\_bt_i = \sum_n \frac{X_{in}}{TAX_{in}}*(x_{in} - tax_{in}) - \sum_n \frac{X_{ni}}{TAX_{ni}}*(x_{ni} - tax_{ni})$	N
(T3.1.2)	$Y_n = X_n - (1-\beta)*\sum_i \frac{X_{ni}}{TAX_{ni}}$	$Y_n * y_n = X_n * x_n - (1-\beta)*\sum_i \frac{X_{ni}}{TAX_{ni}}*(x_{ni} - tax_{ni})$	N
(T3.1.3)	$W_n * L_n = \beta * \sum_i \frac{X_{ni}}{TAX_{ni}}$	$w_n + \ell_n = \sum_i \frac{X_{ni}/TAX_{ni}}{\sum_s X_{ns}/TAX_{ns}}*(x_{ni} - tax_{ni})$	N
(T3.1.4)	$\frac{X_{in}}{X_n} = \Pi_{in}$	$x_{in} - x_n = \pi_{in}$	N^2
(T3.1.5)	$PWORLD = P_1^{SXb(1)} * P_2^{SXb(2)} \ldots * P_N^{SXb(N)}$	$pworld = \sum_s SXb(i) * p_s$	1
(T3.1.6)	$RBT_n = \frac{BT_n}{PWORLD}$	$100*d\_rbt_n = 100*\frac{1}{PWORLD}*d\_bt_n - \frac{BT_n}{PWORLD}*pworld$	N
(T3.1.7)	$GDP_n = W_n * L_n + \sum_i \frac{X_{in}}{TAX_{in}}*(TAX_{in} - 1)$	$GDP_n * gdp_n = W_n * L_n * (w_n + \ell_m) + \sum_i X_{in} * x_{in} - \sum_i \frac{X_{in}}{TAX_{in}}*(x_{in} - tax_{in})$	N
(T3.1.8)	$BT2GDP_n = \frac{BT_n}{GDP_n}$	$100*d\_bt2gdp_n = 100*\frac{1}{GDP_n}*d\_bt_n - \frac{BT_n}{GDP_n}*gdp_n$	N
(T3.1.9)	$YREAL_n = Y_n / P_n$	$yreal_n = y_n - p_n$	N

Table 3.1 continues …

Table 3.1 continued …

	1. Levels	2. Deviation version suitable for GEMPACK	No. of equations
(T3.1.10)	$WREAL_n = W_n / P_n$	$wreal_n = w_n - p_n$	N
(T3.1.11)	$TAX_{in} = \left[TAXIN_n * TAXOUT_{in}\right]^{(1-KD_{in})} * FTAX_{in}$	$tax_{in} = (1-KD_{in}) * \left[taxin_n + taxout_{in}\right] + ftax_{in}$	N^2
(T3.1.12)	$\Pi_{in} = \dfrac{T_i C_i^{-\theta} D_{in}^{-\theta}}{\sum_s T_s C_s^{-\theta} D_{sn}^{-\theta}}$	$\pi_{in} = t_i - \theta * c_i - \theta * d_{in} - \sum_s \Pi_{sn} * \left[t_s - \theta * c_s - \theta * d_{sn}\right]$	N^2
(T3.1.13)	$C_n = W_n^{\beta} * P_n^{1-\beta}$	$c_n = \beta * w_n + (1-\beta) * p_n$	N
(T3.1.14)	$P_n = \gamma * \left[\sum_{i=1}^{N} T_i \left[C_i * D_{in}\right]^{-\theta}\right]^{-1/\theta}$	$p_n = \dfrac{-1}{\theta} * \sum_i \Pi_{in} * \left(t_i - \theta * c_i - \theta * d_{in}\right)$	N
(T3.1.15)	$D_{in} = DTRAN_{in} * TAX_{in}$	$d_{in} = dtran_{in} + tax_{in}$	N^2
		Total number of equations	**$4N^2 + 10N + 1$**

Table 3.2. Variables and parameters in the EK model together with their status in the standard closure and their base period values

	Variable	Description	Standard closure	No. of variables		Base period value
				Endo	**Exo**	
a	X_{in}	Expenditure in n on goods sourced from i	Endo	N^2		Data from Table 3.4
b	TAX_{in}	Power of the tariff (1+rate) on i-to-n flow	Endo	N^2		Data: assumed to be 1 for all i,n in the illustrative simulations
c	BT_i	Balance of trade for country i	Endo	N		Calculated from (T3.1.1) using a & b
d	Y_n	Expenditure in country n by final users	Endo	N		Calculated from (T3.1.2) using a & b, and β from Table 3.3
e	W_n	Wage rate in country n	Endo	N		Assigned value of 1
f	L_n	Employment in country n	Exo		N	Calculated from (T3.1.3) using a, b and e, and β from Table 3.3
g	Π_{in}	Share of expenditure in n accounted for by goods sourced from i	Endo	N^2		Calculated from (T3.1.4) using a and the addition indicated by equation (3.1) in the text
h	X_n	Total expenditure on goods in country n	Endo	N		Calculated using a and the addition indicated by equation (3.1) in the text
i	P_n	Price level in country n	Endo	N		Set at the value of γ given in Table 3.3
j	PWORLD	World price level	Exo		1	Set at the value of γ given in Table 3.3
k	RBT_n	Real balance of trade for country n: balance of trade deflated by the world price level	Endo for USA Exo for i ≠ USA	1	N-1	Calculated from (T3.1.6) using c and j
ℓ	GDP_n	GDP in country n	Endo	N		Calculated from (T3.1.7) using e, f, a and b
m	$YREAL_n$	Real final demand in n (welfare)	Endo	N		Calculated from (T3.1.9) using d and i
n	$WREAL_n$	Real wage in n	Endo	N		Calculated from (T3.1.10) using e and i

Table 3.2 continues …

Table 3.2 continued …

	Variable	Description	Standard closure	No. of variables		Base period value
				Endo	**Exo**	
o	$TAXIN_n$	Variable for imposing uniform powers of tariffs on all imports to n	Exo		N	Assigned value of 1
p	$TAXOUT_{in}$	Variable for imposing uniform powers of tariffs on all exports out of i	Exo		N^2	Assigned value of 1
q	$FTAX_{in}$	Shift variable to retain flexibility in tariff scenarios	Exo		N^2	Calculated as 1 from (T3.1.11) using b, o and p
r	T_i	Productivity variable in i: parameter in Fréchet distribution of i's productivity draws	Exo		N	Assigned value of 1
s	C_i	Cost of labour and intermediate-input bundle used in production in i	Endo	N		Calculated as $\gamma^{1-\beta}$ from (T3.1.13) using e, i, and β and γ values from Table 3.3.
t	D_{in}	Trade-cost markup factor on flows from i to n determined by a combination of exogenous transport and tariff markups	Endo	N^2		Calculated as $\Pi_{in}^{-1/\theta}/\gamma^{1-\beta}$ from (T3.1.12) and (T3.1.14) using Π, T, C and P values from g, r, s, and i, and β and γ values from Table 3.3
u	$DTRAN_{in}$	Transport markup on flows from i to n	Exo		N^2	Calculated from (T3.1.15) using t and b
v	$BT2GDP_n$	Ratio of trade balance to GDP	Endo	N		Calculated from (T3.1.8) using c and ℓ
		Total number of variables		**$4N^2 + 10N + 1$**	**$3N^2+4N$**	

Table 3.3. Parameters values used in the illustrative simulations

	Description	Value
β	Labor share in costs in all industries in all countries	0.5
θ	Dispersion parameter in the Fréchet distribution of productivities. If θ has a low value, then values far away from the mean are more likely than if θ has a high value.	8.28, central value estimated by EK.
γ	The value of this parameter has no effect on percentage deviation results. Consequently, its value is of no importance.*	0.873744
SXb_n	Weight given to prices in n in determining world price level: set at the initial value of $X_n / \sum_s X_s$	
KD_{in}	Kronecker's delta: one if i = n, else zero	

* In deriving (T 3.1.14), Eaton and Kortum show that $\gamma = \left[\Gamma\left(\frac{\theta+1-\sigma}{\theta}\right)\right]^{1/(1-\sigma)}$, where Γ is the gamma function and σ is the elasticity of substitution by industries and final users between commodities. We assume that $\sigma = 5$, but strangely, the value of σ doesn't matter. It affects only the value of γ which doesn't matter. Nevertheless, true to EK, the value we adopt for γ reflects the assumed values for θ and σ.

Equation (T3.1.7) defines GDP for country n as payments to primary factors (only labor in this model) *plus* indirect taxes collected by country n (only tariffs).

Equation (T3.1.8) defines the ratio of the trade balance to GDP in each country.

Equations (T3.1.9) and (T3.1.10) define real final demand and the real wage rate in country n by deflating the nominal values by n's price level. Movements in real final demands can be used as welfare measures.

Equation (T3.1.11) is a convenient mechanism for simulating the effects of a tariff imposed by country n uniformly across all trade partners i, $i \neq n$. This can be done by treating $TAXIN_n$ as an exogenous and shocked variable. Tariffs imposed against i by i's trade partners can be simulated with shocks to $TAXOUT_{in}$ for all n. For example, to simulate a situation in which the US imposes a 30 per cent tariff against all trade partners and these partners retaliate with 30 per cent tariffs against the U.S., we can set $TAXIN_{USA} = 1.30$ and $TAXOUT_{USA,i} = 1.30$ for all i. No flexibility is lost by the introduction of (T3.1.11): any tariff scenario can be simulated by shocks to $FTAX_{in}$.

Equations specific to EK start with (T3.1.12).

Central to the EK model is the determination of Π_{in}. In (T3.1.12), i's share in n's market is determined by three factors: productivity in country i reflected in the variable T_i; unit costs in production in i, C_i; and trade costs including tariffs applying to sales from i to n, D_{in}. The parameter θ is greater than zero. The denominator on the RHS of (T3.1.12) ensures that the Π_{in}'s sum over i to 1. Together, (T3.1.12) and (T3.1.4) imply that

$$\sum_i X_{in} = X_n \tag{3.1}$$

To avoid a structural singularity, this equation is not included in the model's equation list.

Equation (T3.1.12) means that improvements in i's productivity relative to that of other suppliers to n (an increase in T_i relative to the T's for other countries) generates an increase in i's share in n's expenditures. A relative increase in i's unit costs (C_i) reduces i's share in n's expenditures. Country i's share in n's market is also reduced by relative increases in i-to-n trade costs (D_{in}).

In deriving equation (T3.1.12), EK assume that there is a continuum of commodities named on the interval [0,1]. These commodities can be produced in all countries. The price at which country i can offer commodity j to country n is given by

$$P_{in}(j) = \frac{C_i}{Z_i(j)} * D_{in} \tag{3.2}$$

where

$Z_i(j)$ is i's productivity in the production of j, that is output per unit of input.

Productivities in country i are determined by draws from a probability distribution. The draw that gives the value for $Z_i(j)$ is independent of i's draw for other commodities. EK assume that the probability distribution is Fréchet, defined by the cumulative distribution:

$$F_i(z) = \exp(-T_i z^{-\theta}) \tag{3.3}$$

We have already encountered T_i and θ in equation (T3.1.12). In (3.3), they are positive parameters. For an EK model, θ must be greater than 1. In our version of the EK model, we treat T_i as an exogenous variable rather than a parameter. This allows us to simulate the effects of productivity improvements in country i as increases in T_i. Notice from (3.3) that for any given value for z, the probability of a draw being less than z declines as T_i increases.

EK derive equation (T3.1.12) by assuming that country i supplies commodity j to country n only if $P_{in}(j)$ is less than or equal to $P_{sn}(j)$ for all s ≠ i. Then X_{in} is determined by adding demands in n for i's products over all commodities j that meet this criterion.

Equation (T3.1.13) determines the cost of a unit of inputs in country n as a Cobb-Douglas function of wage rates and prices of intermediate inputs in country n. Prices of intermediate inputs are simply the price in n of the bundle of goods used in n.

Equation (T3.1.14) shows that the price of goods in country n depends on productivity and production costs in all countries (T_i and C_i for all i) and on trade costs applying to flows into country n (D_{in} for all i). What is not immediately apparent from (T3.1.14) is the weighting scheme applying to i-variables in the determination of prices in n. This becomes clearer in the percentage change version of (T3.1.14), see next subsection. There we show that the importance of i in determining prices in n depends on the share of n's expenditure that is devoted to goods from i, Π_{in}.

Equation (T3.1.15) specifies trade-cost markups as combinations of what we think of as transport costs, $DTRAN_{in}$, and tariffs, TAX_{in}. For example, if transporting goods from i to n adds 20 per cent to their costs and n imposes a 10 per cent tariff, then the trade cost markup on i-to-n flows is 1.32 (= 1.20*1.10). That is, D_{in} = 1.32.

3.2. Preparation for GEMPACK: initial solution, the EK model in deviation form (col. 2 of Table 3.1) and closures

To implement the EK model in GEMPACK, we need three things: (1) an initial solution; (2) a deviation system in which the variables are deviations from the initial solution; and (3) a closure, that is a division of the variables into endogenous and exogenous.

Initial solution

The last column of Table 3.2 shows how an initial solution can be developed from data on expenditures in each country n on goods sourced from each country i (X_{in}), values for the

Table 3.4. Expenditures in 2015 ($USb) in destination (column) from source (row)*

Dest. Source	USA	Canada	Mexico	Japan	SKorea	China	Germany	EU26	UK	RoW	Total
USA	27793	277	199	127	84	181	111	296	104	736	29909
Canada	329	2841	8	17	9	33	9	26	24	74	3371
Mexico	283	16	1748	4	3	16	7	14	4	52	2147
Japan	149	12	13	10789	76	230	34	70	22	384	11778
SKorea	75	7	11	46	2517	180	16	51	8	280	3190
China	563	43	39	210	122	22912	129	333	75	1058	25483
Germany	128	16	12	26	22	123	5309	743	101	439	6918
EU26	313	43	24	71	41	140	713	20282	326	1171	23123
UK	93	16	3	14	8	24	75	282	3892	205	4614
RoW	965	83	43	516	335	1217	386	1303	246	38459	43554
Total	30692	3354	2099	11820	3219	25055	6790	23401	4802	42857	154087

* These data are for 2015 and were derived from GTAP, see Dixon *et al.* (2019b).

power of the tariff markups (TAX_{in}) and the values of the parameters. Tables 3.4 and 3.3 show the 10-region expenditure data and parameter values for the illustrative application in the next subsection. We assume for illustrative purposes that there were no tariffs in the base period (TAX_{in} = 1 for all i,n).

There is no need to derive the initial solution. The last column of Table 3.2 establishes the existence of an initial solution. We will see shortly that only the initial values for the expenditure flows and tariff markups (X_{in} and TAX_{in}) together with values for the parameters β and θ affect the deviation results generated by the model.

Deviation system

The third column of Table 3.1 shows the EK equations in deviation form. The deviation equations can be derived by applying the rules in Table 2.1.

With two exceptions, the variables in the deviation equations refer to percentage changes away from the initial solution. These percentage change variables are denoted by lowercase versions of the uppercase variables. For example, x_{in} is the percentage deviation in X_{in} away from its initial value, tax_{in} is the percentage deviation in TAX_{in} away from its initial value, etc. The exceptions are d_bt_i and d_bt2gdp_n. These are the ordinary changes in BT_i and $BT2GDP_n$ away from their initial values. While we prefer to work with unit-free percentage changes, this is not possible for variables such as the balance of trade whose value can pass through zero.

By looking through the deviation equations, we see that the only variables for which the initial solution affects the deviation results are X_{in} and TAX_{in}. The values for the parameters θ and β are important. On the other hand, γ does not appear in the deviation equations and has no role in our deviation results. Our arbitrary decisions concerning the initial solutions for W_n and T_i, and the consequent values in the initial solution for L_n, P_n, C_i, D_{in}, are harmless. None of these initial values is in the system of deviation equations.

As mentioned in section 2, we often find deviation equations are more readily interpretable than the corresponding levels equations. An example from Table 3.1 can be obtained by combining the deviation versions of (T3.1.12) and (T3.1.14). This gives:

$$\pi_{in} = -\theta * \left(c_i - \frac{t_i}{\theta} + d_{in} - p_n \right) \tag{3.4}$$

The term t_i/θ can be interpreted as the percentage deviation in country i's productivity (output per bundle of inputs). It is the percentage deviation in i's productivity at every percentile on the cumulative distribution of productivities in response to an increase in T_i. This can be established from the Fréchet distribution by considering the productivity level at a given percentile. We calculate the change in this productivity level, dz, that takes us to the same percentile on the cumulative distribution of productivities after a change in T_i. To do this, we set the total differential of the RHS of (3.3) at zero:

$$\frac{\partial \exp(-T_i * z^{-\theta})}{\partial z} * dz + \frac{\partial \exp(-T_i * z^{-\theta})}{\partial T_i} * dT_i = 0 \tag{3.5}$$

leading to

$$100 * \frac{dz}{z} = 100 * \frac{dT_i}{T_i} * \frac{1}{\theta} = \frac{t_i}{\theta} \tag{3.6}$$

With this interpretation of t_i/θ, equation (3.4) shows that the percentage deviation in country i's share in n's market depends on the percentage movement in the price of i's products in n relative to the percentage movement in the general price level in n. The percentage movement in i's prices in n is given by the percentage movement in the cost of an input bundle in i deflated by productivity growth in i, plus the percentage movement in the power of trade costs on the i-to-n link.

Equation (3.4) demonstrates that at least in this simple model, EK's trade specification is equivalent to Armington. This becomes even more obvious if we rewrite (3.4) as

$$q_{in} = q_n - (1+\theta) * (p_{in} - p_n) \tag{3.7}$$

where

q_{in} is the percentage deviation in the volume of goods sent from i to n;
q_n is the percentage deviation in the volume of goods absorbed in n; and
q_{in}, q_n and p_{in} are given by

$$q_{in} = x_{in} - p_{in} \tag{3.8}$$

$$q_n = x_n - p_n \tag{3.9}$$

and

$$p_{in} = c_i - \frac{t_i}{\theta} + d_{in} \tag{3.10}$$

Equation (3.7) is the percentage deviation form for n's demand for products from i in an Armington model in which agents in n regard products from different source countries as CES substitutes with the Armington substitution elasticity set at 1 plus θ.

Closures

Our EK model contains $4N^2 + 10N + 1$ independent equations (Table 3.1) and $7N^2 + 14N + 1$ variables (Table 3.2) where N is the number of countries (10 in our illustrative simulations). Thus, $4N^2 + 10N + 1$ variables must be endogenous while the remaining $3N^2 + 4N$ must be exogenous. Table 3.2 shows our choice of endogenous and exogenous variables in what we refer to as the standard closure.

In this closure, the shift variables on the powers of the tariffs ($TAXIN_n$, $TAXOUT_{in}$ and $FTAX_{in}$), employment by country (L_n), the state of technology in each country (T_i) and the transport markup factors ($DTRAN_{in}$) are exogenous. The real trade balance (RBT_n) is exogenous for N-1 countries. Equation (T3.1.1) implies that $\sum_{i=1}^{N} BT_i = 0$, and consequently $\sum_{i=1}^{N} RBT_i = 0$. Thus, the trade balance for one country must be left endogenous. The world price level (PWORLD) is exogenous to provide the numeraire.

Other closures are possible. We could exogenize final demands (Y_n) for some countries and endogenize corresponding trade balances. Another possibility is to introduce wage stickiness in some countries by exogenizing W_n. The corresponding endogenization would be employment in these countries, L_n. In the next subsection, there are simulations in which $BT2GDP_n$ is exogenous for N-1 countries instead of RBT_n. We also report a simulation in which U.S. tariffs are endogenized to achieve an exogenous target for U.S. dependence on trade.

Table 3.5 shows the GEMPACK closure files for the illustrative simulations described in the next subsection. The standard closure is in the first column. In the swap statement we endogenize the deviation in the real trade balance for the USA and exogenize the deviation in the world price level.

3.3. Illustrative simulations

Table 3.6 gives a selection of results for six simulations. The first 3 simulations are validation checks. These are simulations for which we already know what the answer should be. If the simulation does not show the expected answer, then we know there must be a problem with either our coding or our thinking.

Simulation (1) is a price-homogeneity check. We applied a 1 per cent shock to PWORLD. In GEMPACK version of the model this is

Shock pworld = 1;

In the standard closure, which we used in simulation (1), PWORLD is the numeraire. There are no other exogenous price variables. As expected, the results for this simulation show:

1 per cent increases in prices and wage rates (P & W); 1 per cent increases in nominal variables such as costs (C), expenditures (X & Y) and GDP; 0 per cent changes in price-deflated variables such as real wages and real final demands (WREAL & YREAL); and zero change in BT2GDP which is the ratio of two nominals. Results for the balance of trade (BT) are given as ordinary changes, which can be shown to be 1 per cent of the initial values.

Simulation (2) is a quantity-homogeneity check. We applied a 1 per cent shock to labour input in each country (L_n). In GEMPACK this is:

Shock lab = uniform 1;

Table 3.5. Closures in the illustrative simulations

Sim (1), (4) and (5)	*Sim (2) and (3)*	*Sim (6)*
Exogenous ptaxin ptaxout fptax lab t dtran d_rbt ; **Rest Endogenous** ; **swap** d_rbt("USA") = pworld; **! End of standard closure**	**Exogenous** ptaxin ptaxout fptax lab t dtran d_rbt ; **Rest Endogenous** ; **swap** d_rbt("USA") = pworld; **! End of standard closure** **! Creates closure for sims (2) and (3)** **swap** pworld =d_rbt("USA") ; **swap** d_rbt = d_rat_bt2gdp; **swap** d_rat_bt2gdp("USA") = pworld;	**Exogenous** ptaxin ptaxout fptax lab t dtran d_rbt ; **Rest Endogenous** ; **swap** d_rbt("USA") = pworld; **! End of standard closure** **! Creates closure for sim (6)** **swap** ptaxin("USA") = pi("USA","USA");

Table 3.6. GEMPACK results for 6 illustrative EK simulations (% deviations from initial solution)

	Price homogeneity	Quantity homogeneity	Tech homogeneity	U.S. tariff, no retaliation	U.S. tariff, with retaliation	U.S. tariff to reduce trade dependence by 50%
	Shock: pworld =1	Shock: ℓ_i =1 for all i	Shock: t_i =1 for all i	Shock: $taxin_{USA}$ = 30	Shock: $taxin_{USA}$ = 30 $taxout_{USA,i}$ = 30 for i ≠ USA	Shock: $\pi_{USA,USA}$ = 5.2144; $taxin_{USA}$ goes endo
Simulation	(1)	(2)	(3)	(4)	(5)	(6)
US tariffs on all imports	0	0	0	30	30	30.1704
Tariffs against all US exports	0	0	0	0	30	0
YREAL: Real final demand (welfare)						
USA	0	1	0.2406	0.3128	-0.5407	0.3094
Canada	0	1	0.2406	-1.5407	-1.7563	-1.5448
Mexico	0	1	0.2406	-2.1042	-2.3860	-2.1101
Japan	0	1	0.2406	-0.1518	-0.1852	-0.1521
SKorea	0	1	0.2406	-0.3510	-0.4271	-0.3517
China	0	1	0.2406	-0.3524	-0.3660	-0.3537
Germany	0	1	0.2406	-0.4256	-0.4299	-0.4272
EU26	0	1	0.2406	-0.1059	-0.1553	-0.1059
UK	0	1	0.2406	-0.0455	-0.1638	-0.0448
RoW	0	1	0.2406	-0.4431	-0.4615	-0.4446
RWAGE: Real wage rate						
USA	0	0	0.2406	-1.2171	-1.5819	-1.2203
Canada	0	0	0.2406	-1.4977	-1.8987	-1.5016
Mexico	0	0	0.2406	-1.8684	-2.3663	-1.8733
Japan	0	0	0.2406	-0.1779	-0.2262	-0.1784
SKorea	0	0	0.2406	-0.4183	-0.5310	-0.4194
China	0	0	0.2406	-0.2044	-0.2579	-0.2049
Germany	0	0	0.2406	-0.2816	-0.3534	-0.2823
EU26	0	0	0.2406	-0.1911	-0.2442	-0.1916
UK	0	0	0.2406	-0.3151	-0.4072	-0.3159
RoW	0	0	0.2406	-0.3144	-0.3955	-0.3152

Table 3.6 continues …

... Table 3.6 continued

	Price homogeneity	Quantity homogeneity	Tech homogeneity	U.S. tariff, no retaliation	U.S. tariff, with retaliation	U.S. tariff to reduce trade dependence by 50%
	Shock: pworld =1	Shock: ℓ_i =1 for all i	Shock: t_i =1 for all i	Shock: $taxin_{USA}$ = 30	Shock: $taxin_{USA}$ = 30 $taxout_{USA,i}$ = 30 for i ≠ USA	Shock: $\pi_{USA,USA}$ = 5.1886; $taxin_{USA}$ goes endo
Simulation	(1)	(2)	(3)	(4)	(5)	(6)
C: Costs						
USA	1	0	0.1202	14.5448	10.2636	14.6324
Canada	1	0	0.1202	-3.3458	-2.3501	-3.3635
Mexico	1	0	0.1202	-3.9952	-3.3234	-4.0154
Japan	1	0	0.1202	-3.5457	-2.6987	-3.5647
SKorea	1	0	0.1202	-3.4932	-2.6270	-3.5121
China	1	0	0.1202	-3.9942	-3.2711	-4.0145
Germany	1	0	0.1202	-3.4836	-2.6086	-3.5024
EU26	1	0	0.1202	-3.4226	-2.5413	-3.4412
UK	1	0	0.1202	-3.3078	-2.4047	-3.326
RoW	1	0	0.1202	-3.6018	-2.7590	-3.6211
BT2GDP: Ratio of balance of trade to GDP						
USA	0	0	0	0.0074	0.0052	0.0074
Canada	0	0	0	0.0004	0.0003	0.0004
Mexico	0	0	0	0.0023	0.0020	0.0023
Japan	0	0	0	-0.0003	-0.0002	-0.0003
SKorea	0	0	0	-0.0007	-0.0005	-0.0007
China	0	0	0	0.0014	0.0012	0.0014
Germany	0	0	0	0.0014	0.0010	0.0014
EU26	0	0	0	-0.0009	-0.0006	-0.0009
UK	0	0	0	-0.0029	-0.0021	-0.0029
RoW	0	0	0	0.0012	0.0010	0.0013

We expected that a uniform 1 per cent increase in the only scarce factor (labour) would have zero effect on price variables (P & W) and cause a 1 per cent increase in the size of all economies. With zero effect on prices and a 1 per cent increase in size, we expected 1 per cent increases in nominal and price-deflated nominal variables (C, X , Y, GDP and YREAL); and 0 per cent changes in price-deflated price variables (WREAL).

However, with the standard closure this is not what the simulation told us. It showed price effects that deviated from zero and effects on nominal variables that deviated from one. The explanation is the exogeneity of real trade balances (RBT) on zero change. Recall that in the standard closure, RBT_n is exogenous for N-1 countries and endogenously determined as a residual for the remaining country. With zero shocks to RBT_n, deficit countries can't expand their real final demands (YREAL) by the expected full 1 per cent because the part of their consumption financed by their trade deficit does not expand. Surplus countries expand their real final demands by more than 1 per cent. They don't increase the amount of potential real consumption that they lose in creating their trade surpluses. To check this explanation, we reran the simulation with $BT2GDP_n$ replacing RBP_n on the exogenous list for N-1 countries. The replacement was made in GEMPACK by the three swap commands at the bottom of the middle column in Table 3.5. With the trade balances now allowed to expand with GDP, the simulation showed the expected results.

Simulation (3) is a technology homogeneity check. We applied a 1 per cent shock to the technology variable T_i for all i. In GEMPACK this is:

Shock t = uniform 1;

With homogeneous preferences and constant returns to scale in production, a uniform technology improvement does not change in the commodity composition of trade on any link. Consequently, after the uniform increase in the T_is, we would expect the productivity values in country i on its sales to n (includes i) to occupy the same positions in i's cumulative distribution of productivities as they did in the original situation. That is, if in the original situation, i sent commodity j to n, and commodity j was at the 85th percentile of i's cumulative distribution of productivities, then we would expect in the new situation that i would continue to send j to n, and that commodity j would continue to be at the 85th percentile of i's cumulative distribution of productivities.

These expectations suggest that in simulation (3), the percentage increase in productivity throughout the world will be revealed if we can evaluate by how much z has to change so that F(T,z) is unchanged when T increases by 1 per cent, where F(T,z) is the Fréchet function defined in (3.3). We make this evaluation by working with the equation

$$F(T(\text{initial}), z(\text{initial})) = F(T(\text{final}), z(\text{final})) \tag{3.11}$$

Under (3.3), (3.11) gives

$$-T(\text{initial}) * z(\text{initial})^{-\theta} = -T(\text{final}) * z(\text{final})^{-\theta} \tag{3.12}$$

For a 1 per cent increase in T we have

$$\frac{z(\text{final})}{z(\text{initial})} = 1.01^{1/\theta} \tag{3.13}$$

With θ = 8.28, see Table 3.3, (3.13) implies that a 1 per cent increase in T will cause productivity to increase by 0.1202 per cent.

The increase in productivity refers to the increase on output per unit of input of Cobb-Douglas bundles of labour and intermediate inputs. With labour representing only half the value of the input bundles (β = 0.5), the productivity increase applies to inputs worth twice the value of GDP. Because labour is the only scarce factor, the percentage increase in real wages must be approximately twice the percentage increase in productivity. In simulation (3) we see real wage increases of 0.2406 per cent (approximately twice 0.1202 per cent). This is also the percentage increase in real final demands. There are no changes in prices because the world price level is fixed, and the uniformity of the productivity changes rules out changes in relative prices between countries. Consequently, the percentage changes in nominal wages are the same as the percentage changes in real wages. With nominal wages rising by 0.2406 per cent and the prices of intermediate inputs fixed, costs per unit of output must increase by about half of 0.2406. This is borne out by the results in Table 3.6 for simulation (3) which show cost increases of 0.1202 per cent.

Simulations (4) and (5) are projections of the effects of tariff increases. Both simulations were conducted with the standard closure. In simulation (4), the U.S. imposes a 30 per cent tariff on all imports from all sources. In GEMPACK this is:

Shock ptaxin("USA") = 30;

In simulation (5), other countries respond to the U.S. tariffs by imposing 30 per cent tariffs on imports from the U.S. The shocks in this simulation are:

Shock ptaxin("USA") = 30;

Shock ptaxout("USA", REG) = uniform 30;

where REG is the set of all countries.

The results for real final demand (YREAL) are indicators of percentage effects on welfare. In simulation (4) the U.S. gains a welfare benefit of 0.3128 per cent. The U.S. tariffs increase costs in the U.S. Part of these costs are passed on to U.S. trade partners as higher export prices. At the same time, exporters to the U.S. lower their prices. This is required to maintain their competitiveness in the U.S. market at a sufficient level to fix their real trade balances (fixed in the standard closure). The net result of these price changes is a welfare-enhancing terms-of-trade improvement for the U.S. It turns out this terms-of-trade benefit is sufficient to outweigh the welfare reducing effects for the U.S. of a tariff-induced misallocation of resources.

In simulation (4), all other countries lose from the U.S. tariffs. The biggest losses are for Canada and Mexico (1.5407 and 2.1042 per cent).

Even without retaliation from other countries, the tariffs are not entirely benign for the U.S. They cause a loss in real wages of 1.2171 per cent. Potentially, this wage loss could be compensated by reductions in other taxes, or transfers of the tariff revenue to households. Investigating these possibilities is beyond the scope of the current model.

In simulation (5), retaliation turns the U.S. welfare gain into a welfare loss of 0.5407 per cent. Retaliation hurts the U.S. but doesn't benefit the retaliating countries. Their welfare losses are slightly greater in simulation (5) than in simulation (4).

Giesecke and Waschik (2025) used a detailed dynamic version of GTAP to simulate the effects of Trump tariffs without and with retaliation. In broad terms, their results are similar to those for simulations (4) and (5). Giesecke and Waschik show that without retaliation, terms-of-trade effects will leave the U.S. with a small welfare gain. Most other countries will lose with the biggest losses being for Mexico and Canada. In the Giesecke and Waschik study, retaliation wipes out the U.S. gains and exacerbates the losses for most other countries.

The similarity between the Giesecke and Waschik results and those from our EK model is interesting from a methodological point of view. In the version of GTAP used by Giesecke and Waschik, trade is modelled on conventional Armington lines. As in Bekkers *et al.* (2023), it appears that switching between Armington and EK specifications of trade need not have significant implications for the results of tariff simulations.

Simulation (6) is a demonstration of closure flexibility. Rather than asking what are the effects of tariff changes, in simulation (6) we ask what tariff increases would be required to reduce U.S. dependence on trade by 50 per cent.

We measure trade dependence as the import share in total U.S. expenditure, that is

$$\text{TradeDependence}_{\text{USA}} = 1 - \Pi_{\text{USA,USA}} \qquad (3.14)$$

From Table 3.4 we see that the initial value for U.S. trade Dependence is 0.0956 (= 1-27793/30692). In simulation (6), we require a reduction in trade dependence from 0.0956 to 0.0478. To achieve this, we need to move $\Pi_{\text{USA,USA}}$ from 0.9044 to 0.9522, a percentage increase of 5.2144. So that we can apply a shock to $\Pi_{\text{USA,USA}}$, and find the corresponding tariff increase, we need to exogenize trade dependence and shock it with the target movement while endogenizing the tariff movement. The necessary closure swap is indicated in the simulation (6) column of Table 3.5.

The endogenously calculated U.S. tariff increase in simulation (6) is 30.1704 per cent. This is close to the tariff increase (30 per cent) that was applied exogenously in simulation 4. With no retaliation in simulations (4) and (6), the results for the two simulations are almost the same.

4. Concluding remarks

Since the 1980s, GEMPACK has been a major facilitating factor in the application and development of CGE models. This is also true of GAMS, the main alternative to GEMPACK.

GEMPACK is superior to GAMS in solution speed, and for large models this advantage is overwhelming. To the best of our knowledge, GEMPACK can solve all operational CGE models that are solvable in GAMS, but not the other way around. GEMPACK copes with implicit functions that arise from production functions such as CRESH and utility functions such as AIDADS. These are a problem for GAMS.

With a team of 4 software specialists at the Centre of Policy Studies, GEMPACK continues to evolve and improve. The work of the GEMPACK team includes not only computational algorithms, but also auxiliary programs for assisting in building models and in analysing and presenting results.

Originally, GEMPACK was used on CGE models with an Armington specification of trade. But there are now GEMPACK-based CGE models incorporating Melitz (2003).[10] Motivated by Markusen (2023), we have been adjusting Melitz from large-group monopolistic competition to small-group monopolistic competition in a GEMPACK-based version of the GTAP model, see Dixon and Rimmer (2024). Also in a GEMPACK-based version of GTAP,

[10] See for example, Akgul *et al.* (2016), Balistreri & Rutherford (2013), Bekkers & Francois (2018) and Dixon *et al.* (2018 & 2019a).

Bekkers *et al.* (2023) include EK features. Nevertheless, Armington remains the dominant trade specification in CGE models.

The paper by Bekkers *et al.* and the illustrative application of the EK model in section 3, produce results that look like those that could be obtained from an Armington model. This is not surprising in view of the Armington/EK equivalence demonstrated in section 3 by the equation (3.7), but it led us to think about EK's contribution relative to Armington.

In the early models [e.g. Johansen (1960) and Evans (1972)], good j produced in country n is a perfect substitute for good j produced anywhere else in the world. This led to a crippling problem: extreme specialization. Import shares for good j in country n tended to flip between 0 and 1 in response to tiny shocks. Similarly, the simulated composition of a country's exports tended to be far too specialized: if it was a good idea for country n to export a small quantity of j, then the models tended to show that country n should export a lot of j and little or nothing of most other commodities. To achieve realistic results, the CGE pioneers resorted to *ad hoc* fixes such as exogenous maximum and minimum shares for imports of each commodity and exports from each industry.

Armington solved the specialization problem. His specification is based on differences in the characteristics of commodities in the same classification but produced in different countries. Such differences are easy to visualize when the classifications are those normally available in data published by statistical agencies. Australian coal is different from Chinese coal; a vacation in Fiji is different from a vacation in Iceland; and French wine is different from U.S. wine, etc. Any broadly defined commodity j produced in one country can be modelled in the Armington style as an imperfect substitute for commodity j produced in another country. With the adoption of the Armington specification, a CGE model can produce realistic (muted) responses of trade shares to policy and other shocks.

Eaton and Kortum (2002) imagine going to very fine commodity classifications. At that level, there are no differences in the characteristics of commodities in the same classification but produced in different countries. Yet, Eaton and Kortum produce a trade specification in which trade shares behave in an empirically reasonable manner. If costs rise in country n in a broadly defined industry j, then in the background to an EK model, country n entirely losses its exports in some of the narrowly defined subcategories of j. This is not explicit in an EK-style CGE model. It shows up as a loss by n of a fraction of its exports of broadly defined commodity j. Thus, the EK results look like Armington results.

What EK have contributed is an attractive micro-economic story that justifies the Armington specification applied for commodities defined at statistically relevant levels of disaggregation. The productivity dispersion parameter θ in EK plays a similar role to the import-domestic substitution elasticity in Armington. Like a high import-domestic substitution elasticity in Armington, a high value for θ in EK implies high sensitivity of trade flows to changes in relative costs.

References

Akgul, Z., N.B. Villoria and T.W. Hertel (2016), "GTAP-HET: Introducing Firm Heterogeneity into the GTAP Model", *Journal of Global Economic Analysis*, 1(1), pp. 111-180.

Armington, P.S. (1969) "A theory of demand for products distinguished by place of production", *Staff Papers-International Monetary Fund*, 16(1): 159–178.

Arrow, K.J., H.B. Chenery, B.S. Minhas and R.M. Solow (1961), "Capital-labor substitution and economic efficiency", *Review of Economics and Statistics*, 43, 225-250.

Balistreri, E. and T. Rutherford (2013), "Computing general equilibrium theories of monopolistic competition and heterogeneous firms", chapter 23 in P.B. Dixon and D.W. Jorgenson (editors*), Handbook of Computable General Equilibrium Modeling*, Elsevier, pp. 1513-1570.

Bekkers, E. and J. Francois (2018), "A parsimonious approach to incorporate firm heterogeneity in CGE models", *Journal of Global Economic Analysis*, 3(2), pp. 1-68.

Bekkers E., E. Corong, J. Francois and H. Rojas-Romagosa (2023), "A Ricardian trade structure in CGE: modelling Eaton-Kortum based trade with GTAP", *Journal of Global Economic Analysis*, vol. 8(2), pp. 1-59.

Bergman L. (1978), "Energy policy in a small open economy the case of Sweden", RR-78-16, *International Institute of Applied Systems Analysis*, Laxenberg, available at https://pure.iiasa.ac.at/id/JGEAeprint/827/1/RR-78-016.pdf .

Dahlquist, G., A. Bjorck and N Anderson (translator) (1974), *Numerical Methods*, Prentice-Hall, New Jersey.

Dixon, P.B., B.R. Parmenter, G.J. Ryland and J. Sutton (1977), *ORANI, A General Equilibrium Model of the Australian Economy: Current Specification and Illustrations of Use for Policy Analysis,* Vol. 2 of the First Progress Report of the IMPACT Project, Australian Government Publishing Service, Canberra, pp. xii + 297.

Dixon, P.B., A.A. Powell and B.R. Parmenter (1979), *Structural Adaptation in an Ailing Macroeconomy*, Melbourne University Press, pp. xvii + 88.

Dixon, P.B., B.R. Parmenter, John Sutton and D.P. Vincent (1982), *ORANI: A Multisectoral Model of the Australian Economy*, Contributions to Economic Analysis 142, North-Holland Publishing Company, pp. xviii + 372.

Dixon, P.B. and M.T. Rimmer (2015), "Obituary: Kenneth Robert Pearson", *Gazette, The Australian Mathematical Society*, vol. 42(4), pp. 228-231.

Dixon, P.B., M. Jerie and M.T. Rimmer (2018), *Trade theory in computable general equilibrium models: Armington, Krugman and Melitz*, Advances in Applied General Equilibrium Modelling, Springer Nature, Singapore, pp. xi + 189.

Dixon, P.B., M. Jerie and M.T. Rimmer (2019a), "Melitz in GTAP made easy: the A2M conversion method and result interpretation", *Journal of Global Economic Analysis*, 4(1), pp. 97-127.

Dixon, P.B., M.T. Rimmer and N. Tran (2019b), "GTAP-MVH, A Model for Analysing the Worldwide Effects of Trade Policies in the Motor Vehicle Sector: Theory and Data", *CoPS Working Paper G-290*, available at https://www.copsmodels.com/ftp/workpapr/g-290.pdf .

Dixon, P.B., M. Jerie, D. Mustakinov, M.T. Rimmer, N.Sheard, F. Schiffmann and G.Wittwer (2024), "Constructing a Destructive Events Tool using Small Rectangular Areas, Computable General Equilibrium Modelling and Neural Networks", *CoPS/IMPACT Working Paper Number G-349*, available at https://www.copsmodels.com/elecpapr/g-349.htm

Eaton, J. and S. Kortum (2002), "Technology, Geography and Trade", *Econometrica*, vol. 70(5), pp. 1741-79, September.

Evans, H.D. (1972), *A General Equilibrium Analysis of Protection: the effects of protection in Australia*, Contributions to Economic Analysis 76, North-Holland Publishing Company, Amsterdam.

Geary, R. (1950-1), "A note on 'A constant utility index of the cost of living' ", *Review of Economic Studies*, vol 18, pp. 65-66.

Giesecke, J. and R. Waschik (2025), "Economic Analysis of U.S. Tariffs Introduced Over March-April 2025", *CoPS/IMPACT Working Paper Number G-352*, available at https://www.copsmodels.com/ftp/workpapr/g-352.pdf .

Hanoch, G. (1971), "CRESH production functions", *Econometrica*, 38, September, pp. 695-712.

Hertel, T.W., J.M. Horridge and K.R. Pearson (1992), "Mending the family tree: a reconciliation of the linearized and levels schools of AGE modelling", *Economic Modelling,* 9, pp. 385-407.

Hertel, T.W. (editor), (1997), *Global trade analysis: modeling and applications*, Cambridge University Press, Cambridge, UK, pp. xvii + 403.

Horridge, J.M., A Meeraus, K. Pearson and T. Rutherford (2013), "Solution software for computable general equilibrium modeling", Chapter 20, pp. 1331-1381 in P.B. Dixon and D.W. Jorgenson (editors) *Handbook of Computable General Equilibrium Modeling*, Elsevier.

Johansen, L. (1960), *A Multisectoral Study of Economic Growth,* Contributions to Economic Analysis 21, North-Holand, Amsterdam.

Keller, W.J. (1980), *Tax incidence: a general equilibrium approach,* Contributions to Economic Analysis 134, North-Holland, Amsterdam.

Klein, L. and H. Rubin (1948-9), "A constant utility index of the cost of living", *Review of Economic Studies*, vol 15, pp. 84-87.

Leontief, W.W. (1936), "Quantitative input-output relations in the economic system of the United States", *Review of Economics and Statistics*, 18, 105-125.

Markusen, J.R. (2023), "Incorporating theory-consistent endogenous markups into applied general equilibrium models", *Journal of Global Economic Analysis*, 8(2), pp. 60-99.

Melitz, M.J. (2003), "The impact of trade on intra-industry reallocations and aggregate industry productivity", *Econometrica*, 71(6), pp. 1695–1725.

Rimmer, M.T. and A.A. Powell (1996), "An implicitly additive demand system", *Applied Economics,* 28*,* pp. 1613-1622.

Staelin, C. (1976), "A general equilibrium model of tariffs in a non-competitive economy", *Journal of International Economics*, 6, pp.39-63.

Stone, R. (1954) "Linear Expenditure System and Demand Analysis: An Application to the British Pattern of Demand", *Economic Journal*, Vol. 64, No. 255 (September 1954), pp. 511-32.

Taylor, L. and S.L. Black (1974), "Practical general equilibrium estimation of resource pulls under trade liberalization", *Journal of International Economics*, 4, April, pp.35-58.

Appendix: TABLO code for implementation of EK model in GEMPACK

```
!***************************************************************************!
!***************************************************************************!
!                              EK model                                     !
! prepared by Peter Dixon and Maureen Rimmer                                !
! August 2025                                                               !
!***************************************************************************!
!***************************************************************************!

File
DATA # file containing data for the EK model #;

Set
REG  # Regions # read elements from file DATA header "H1";

Coefficient
    (Parameter)
Beta  # Labour share in costs in all industries in all regions #;
    (all,i,REG)
C_BT(i)  #  Balance of trade for country i #;
    (Parameter) (All,i,REG)(All,n,REG)
KD(i,n)  #  Kronecker's delta: 1 on the diagonal, else 0 #;
     (all,n,REG)
C_GDP(n)  #  GDP in country n #;
C_PWORLD #  World price level #;
    (all,i,REG)(all,n,REG)
C_PI(i,n) # Share of expenditure in n accounted for by goods sourced form i #;
     (all,i,REG)(all,n,REG)
C_PTAX(i,n) # Power of the ad valorem tax (1 + rate) on i-to-n flows, tariff imposed by n #;
    (Parameter)
Theta # Dispersion parameter in the Fréchet distribution of productivities #;
    (all,n,REG)
C_X(n) # Total expenditure on goods in country n #;
    (Parameter) (all,n,REG) C_Xb(n) # Expenditure in n, base value #;
    (all,i,REG)(all,n,REG)
C_X_in(i,n) # Expenditure in n on goods sourced from i #;
    (all,n,REG)
C_WAGEBILL(n) # Payments to labor in n  #;
    (all,n,REG)
C_Y(n)  #  Expenditure in country n by final users  #;

Read
Beta from file DATA Header "Beta";
C_PTAX from file data Header "PTAX";
C_PWORLD from file data Header "PWLD";
C_X_in from file data Header "CXin";
Theta from file DATA Header "Thta";
```

```
Formula
(all,n,REG) C_X(n) = Sum(i,REG, C_X_in(i,n));
(all,i,REG)(all,n,REG) C_PI(i,n) =C_X_in(i,n) /C_X(n);
(all,n,REG) C_Y(n) = C_X(n) - (1-BETA)*Sum(i,REG, C_X_in(n,i)/C_PTAX(n,i));

(All,i,REG) C_BT(i)= Sum(n,REG, {C_X_in(i,n)/C_PTAX(i,n)}  - {C_X_in(n,i)/C_PTAX(n,i)}   );
(all,n,REG) C_WAGEBILL(n) =BETA*Sum(i,REG, C_X_in(n,i)/C_PTAX(n,i));
(all,n,REG) C_GDP(n)   = C_WAGEBILL(n) +Sum(i,REG, [C_X_in(i,n)/C_PTAX(i,n)]*(C_PTAX(i,n)-1));
(Initial) (all,n,REG) C_Xb(n)= C_X(n);

Formula (initial) (All,i,REG)(All,n,REG) KD(i,n) =0;
Formula (initial) (All,i,REG) KD(i,i) =1;

Variable
    (All,i,REG)
c(i) #  Cost of input bundle used in production in i #;
    (all,i,REG)(All,n,REG)
d(i,n) # Trade-cost markup factor on i-to-n flows determined as a combination of exogenous transport & tariff
markups #;
    (change) (All,i,REG)
d_bt(i)  #  Balance of trade for country i #;
    (Change) (All,n,REG)
d_rat_bt2gdp(n)  # Ratio of balance of trade to gdp in n # ;
    (Change) (All,n,REG)
d_rbt(n)  # Real balance of trade for country n: balance of trade deflated by the world price level #;
    (all,i,REG)(All,n,REG)
dtran(i,n) # Transport markup on flows from i to n    #;
    (All,i,REG)(All,n,REG)
fptax(i,n)  #  Shift variable, included to allow all options in setting tariffs  # ;
    (all,n,REG)
gdp(n)  #  GDP in country n #;
    (All,n,REG)
lab(n) # Employment in country n #;
    (All,n,REG)
p(n) #  Price level in country n #;
    (all,i,REG)(All,n,REG)
pi(i,n) # Share of expenditure in n accounted for by goods sourced form i #;
    (all,i,REG)(All,n,REG)
ptax(i,n) # Power of the ad valorem tax (1 + rate) on i-to-n flows, tariff imposed by n #;
    (All,n,REG)
ptaxin(n)  # Uniform shift in power of n's tariff on n's imports from all sources     #;
    (All,i,REG)(All,n,REG)
ptaxout(i,n)  # Shift in power of n's  tariff in imports from i , useful in tariff retalitation simulations #;
pworld #  World price level #;
    (All,i,REG)
t(i) #  State of technology in country i #;
    (All,n,REG)
w(n) # wage rate in n #;
    (All,n,REG)
wreal(n)   # Real wage rate #;
    (All,n,REG)
x(n) # Expenditure in n #;
    (all,i,REG)(All,n,REG)
x_in(i,n) # Expenditure in n on goods sourced from i #;
    (All,n,REG)
y(n)  #  Expenditure in country n by final users  #;
```

```
    (All,n,REG)
yreal(n)   # Real final demand  in n#;

Update
(all,i,REG)(all,n,REG) C_X_in(i,n) = x_in(i,n);
C_PWORLD = pworld;
 (All,i,REG) (all,n,REG)  C_PTAX(i,n)   =ptax(i,n);

Equation
!(T3.1.1)!  E_d_bt      #  Balance of trade for country i #
(All,i,REG) 100*d_bt(i)= Sum(n,REG, [{C_X_in(i,n)/C_PTAX(i,n)}*(x_in(i,n)-ptax(i,n))  -
                              {C_X_in(n,i)/C_PTAX(n,i)}*(x_in(n,i)-ptax(n,i)) ] );

!(T3.1.2)!   E_y  # Value of final use in n #
(All,n,REG) C_Y(n)*y(n) = C_X(n)*x(n) - (1-BETA)*Sum(i,REG, [C_X_in(n,i)/C_PTAX(n,i)]*(x_in(n,i) -
ptax(n,i)));

!(T3.1.3))! E_w  # Wage rate in i #
(All,n,REG) w(n)+lab(n)=  Sum(i,REG,
[C_X_in(n,i)/C_PTAX(n,i)]/[Sum(s,REG,C_X_in(n,s)/C_PTAX(n,s))]*(x_in(n,i) -ptax(n,i))  );

!(T3.1,4)!  E_x_in  # Value flow from i to n #
(all,i,REG)(All,n,REG)x_in(i,n) - x(n) = pi(i,n) ;

!(T3.1.5)!  E_pworld   #  World price level #
pworld = [1/Sum(ii,REG, C_Xb(ii))]*Sum(i,REG, C_Xb(i)*p(i));

!(T3.1.6)!  E_d_rbt      # Real balance of trade for country n #
(All,n,REG) 100*d_rbt(n) = 100*[1/C_PWORLD]*d_bt(n)  -[ C_BT(n)/C_PWORLD]*pworld;

!(T3.1.7)!  E_gdp  #  GDP in country n #
(All,n,REG) C_GDP(n)*gdp(n)=C_WAGEBILL(n)*(w(n)+lab(n))
 +  Sum(i,REG, C_X_in(i,n)*x_in(i,n)) -  Sum(i,REG, [C_X_in(i,n)/C_PTAX(i,n)]*(x_in(i,n) -ptax(i,n)));

!(T3.1.8)!  E_d_rat_bt2gdp      # Ratio of balance of trade to gdp in n #
(All,n,REG) 100*d_rat_bt2gdp(n) = 100*[1/C_GDP(n)]*d_bt(n)  -[ C_BT(n)/C_GDP(n)]*gdp(n);

!(T3.1.9)!  E_yreal # Real final demand #
(All,n,REG) yreal(n) = y(n) - p(n);

!(T3.1.10)! E_wreal # Real wage rate #
(All,n,REG) wreal(n)  = w(n) - p(n);

!(T3.1.11)! E_ptax    # Power of the ad valorem tax (1 + rate) on i-to-n flows
(All,i,REG)(All,n,REG) ptax(i,n) =[1-KD(i,n)]*[ptaxin(n)+ptaxout(i,n)] + fptax(i,n);

!(T3.1.12)! E_pi  # Value flow from i to n #
(all,i,REG)(All,n,REG) pi(i,n)
     =  t(i)  -Theta*c(i)  -Theta*d(i,n) -  Sum(s,REG, C_PI(s,n)*(t(s)  -Theta*c(s)  -Theta*d(s,n)) );

!(T3.1.13)! E_c  #  Cost of a unit of input in cnt i #
(All,n,REG) c(n) = Beta*w(n) +(1-Beta)*p(n);
```

```
!(T3.1.14)! E_p # Trade barriers on the i to n link #
(All,n,REG) p(n) = -(1/Theta)*Sum(i,REG, C_PI(i,n)*[ t(i)  -Theta*c(i)  -Theta*d(i,n)]) ;

!(T3.1.15)! E_d # Trade-cost markup factor on i-to-n flows #
(All,i,REG)(All,n,REG) d(i,n) = dtran(i,n)+ptax(i,n);


! Using GEMPACK program AnalyzeGE, modellers can evaluate pre-simulation values of coefficients.
With the code below, modellers can also evaluate post-simulation values of coefficients !
Postsim (begin);
File (new) Results # Some sim results #;
Write
  x  to file Results header "x";
Postsim (end);
```