CHAPTER 19

Validation in Computable General Equilibrium Modeling

Peter B. Dixon, Maureen T. Rimmer
Centre of Policy Studies, Monash University

Abstract

Validity is a key issue for consumers of computable general equilibrium (CGE) modeling services. What assurance can producers of CGE results give to consumers that a CGE analysis: (i) is computationally sound, (ii) uses accurate up-to-date data, (iii) adequately captures behavioral and institutional characteristics of the relevant part of the economy, (iv) is consistent with history, and (v) is based on a model that has forecasting credentials? This chapter gives some answers. With regard to (i), CGE modelers have an obligation to conduct exhaustive test simulations. The value of these procedures goes beyond computational checking. Test simulations are a practical way to become familiar with a model and often reveal modeling weaknesses. On (ii) and (iii), the most effective approach for displaying the relevant data and describing what is going on is via a back-of-the-envelope (BOTE) model. A well-designed BOTE model has two properties: it reveals the roles of major behavioral, institutional and data assumptions in causing a model to generate a particular result; and it is small enough to be managed with pencil and paper (on the back of an envelope) and to be presented in a limited timeframe to policy advisors. On (iv) and (v), the chapter describes various aspects of statistical validation, concentrating mainly on historical simulation, baseline forecasting and the testing of baselines against reality. This work demonstrates that CGE models can produce forecasts at a highly disaggregated level that comfortably beat non-model-based trend forecasts. It also demonstrates that there is considerable potential for improved CGE forecasts through conscientious data work and improved methods for projecting trends from historical simulations into forecasting simulations.

Keywords

Validating CGE models, CGE test simulations, CGE back of the envelope, CGE baseline, CGE forecasting performance

JEL classification codes

C68, C52, C53, D58

19.1 INTRODUCTION

Validation in CGE modeling has several meanings and purposes. At the simplest level validation refers to a demonstration that results have been computed correctly − that they follow from the model's theory and data. At a second level, validation refers to a demonstration that a modeler's explanation of results is a legitimate reflection of the

Handbook of CGE Modeling - Vol. 1 SET
ISSN 2211-6885, http://dx.doi.org/10.1016/B978-0-444-59568-3.00019-5

way the model works. This type of validation is vital in assessing: what has been taken into account in an analysis; whether the model's data on the parts of the economy with which the analysis is concerned are up-to-date and accurate; and whether the mechanisms built into the model are an adequate representation of how the relevant parts of the economy behave. A third meaning of validation refers to a demonstration of a model's consistency with history and a fourth meaning refers to a model's forecasting ability.

Corresponding to the multiple meanings and purposes are multiple validating procedures. These include:

(i) Test simulations for which the correct results are known *a priori*.
(ii) Inclusion in a model of variables that check national accounts identities.
(iii) Construction of a back-of-the-envelope (BOTE) model to explain results from a particular application of a full-scale model.
(iv) Application of regression analysis to test explanations of results.
(v) Estimation of parameters via econometric methods applied to time-series data.
(vi) Tests of a model's ability to reproduce outcomes for endogenous variables given the true values of exogenous variables.
(vii) Historical simulation in which a model is forced to track history.
(viii) Out-of-sample forecasting tests.

As will be seen in this chapter, there is not a simple correspondence between purposes and procedures. A procedure can serve more than one purpose and a purpose can be met by more than one procedure. For example, the procedure of conducting a test simulation can be used not only as a check of computational validity but also to reveal aspects of how a model works. Going the other way around, the purpose of understanding how a model works can be served not only by the construction of a BOTE model but also by a test simulation.

The aim of this chapter is to describe the state-of-the-art in CGE validation in a way that will be of interest to new CGE modelers, experienced CGE modelers and policy economists who, while not being CGE modelers themselves, have responsibilities for absorbing CGE results and managing CGE modelers. For new modelers, we hope that the chapter will act as an introduction to validation procedures and encourage them to be ambitious about checking and explaining their results. For experienced modelers, we hope that there are a few new ideas particularly in the areas of BOTE modeling, historical simulation and out-of-sample forecasting. For policy economists, not all the technical detail in the chapter will be of interest. However, we hope to give them a realistic idea of what they should demand of CGE modelers. In particular, we hope to give them the confidence to demand explanations of CGE results that can be understood in terms of familiar economic constructs such as demand and supply diagrams, producer and consumer surplus, and data from input-output tables and the national accounts. Only in this way can policy economists make an informed assessment of results in the context of the policy issue at hand and thereby exercise their responsibility to neither accept results that are flawed nor reject results that contain valuable insights.

The chapter is organized around the eight procedures listed above. Section 19.2 describes test simulations. While the primary purpose of these simulations is code checking, they can also have interesting interpretations. This idea is developed with an example concerning the controversial issue of the Chinese exchange rate.

Section 19.3 is about GDP identities. Starting from an input-output database for a CGE model, we establish the equality of the expenditure and income measures of nominal GDP. We then invoke the CGE assumption of cost minimizing behavior by producers, enabling us to derive a GDP identity in *real* terms. Checking that GDP measures stay locked together in both nominal and real terms throughout a CGE simulation is a powerful test of computational validity. However, deriving and understanding the GDP identities has a deeper value which makes struggling with the rather tedious algebra in Section 19.3 worthwhile. The measure of real GDP from the income side provides a strong link from CGE modeling to elementary microeconomic concepts.

Section 19.4 is about validation through BOTE analysis and plausibility checks. We start in Section 19.4.1 by discussing CGE storytelling. This is the art of developing a convincing *qualitative* explanation of seemingly counterintuitive CGE results. Then in Section 19.4.2 we show how the income measure of real GDP is the foundation for *quantitative* BOTE checks of CGE results for the effects of changes in taxes, technologies and primary factor inputs. We give two illustrations. The first concerns a CGE carbon tax result presented to the White House in the run up to the UN Climate Change Conference held in Copenhagen at the end of 2009. The second concerns the interpretation by Australia's Productivity Commission of a CGE result for the effects of a reduction in tariffs on imports of motor vehicles. Section 19.4.3 describes the development of BOTE models for checking and justifying CGE results. We work through two examples. In the first, we set out a very small BOTE model that was used to analyze results for the effects on aggregate employment of an increase in tariffs. The second involves a more elaborate BOTE model that was used to explain macro results from a CGE analysis presented to the Australian parliament on the effects of a proposed major reform in the tax system. The last part of Section 19.4 describes the technique of checking explanations of CGE results by using regression equations. We illustrate the technique on a set of CGE results generated for the US International Trade Commission on the employment effects, by state, in the US, of a unilateral removal of import restraints.

Section 19.5 describes three types of historical validation. The most obvious is estimation of a model's parameters by time-series econometrics. As discussed in Section 19.5.1, this has proved to be more difficult and less valuable than anticipated in the early days of CGE modeling. In Section 19.5.2 we describe a goodness-of-fit approach to historical validation in which observed values for endogenous variables are compared with values for these variables computed from the model on the basis of observed values for exogenous variables. This approach was adopted in Johansen (1960) seminal work on Norway and by other early CGE modelers. Section 19.5.3 describes historical simulations. These have been popular for the last 20 years with MONASH-style modelers. Historical simulations involve forcing a model to track history by exogenizing and shocking observable variables, and allowing the model to determine implied paths for naturally exogenous but unobservable shift variables such as industry technologies, household preferences, required rates of return on capital and positions of export-demand curves and import-supply curves. Validation is not assessed in terms of goodness-of-fit: the model fits perfectly by construction. Instead, validation consists of an assessment of the plausibility of the endogenously generated paths of the numerous shift variables.

Quantification in historical simulations of trends in shift variables representing technologies, preferences, etc., provides a key input to baseline forecasting. This is discussed in Section 19.6. Baseline forecasting opens up a possibility in CGE modeling for out-of-sample forecasting validation tests. We describe such a test that we have carried out with the USAGE model of the US. The model was used to generate a baseline forecast for 1998–2005 for the outputs of 500 commodities. This forecast is genuine in that it uses only information that was available up to 1998. To assess forecasting performance we compared the forecast growth rates for the 500 commodities with actual outcomes. As described in Section 19.6, we subsequently took the research in three directions. (i) We reran the 1998–2005 forecast with the true movements in macro, energy, technology and trade variables introduced successively. Analysis of the change in forecasting performance from stage to stage provides guidance on how research resources would best be used to improve CGE forecasting performance. (ii) We went back to the 1998 data sources to see if there was information that we overlooked which would improve forecasting performance. (iii) We reassessed the details of our forecasting technique to see if trends calculated with data available up to 1998 could have been used more effectively in the forecast.

Section 19.7 contains a brief summary and concluding remarks. We emphasize the importance of confronting modeling results with actual outcomes as a means of stimulating model improvements.

19.2 CHECKING THE CODE: HOMOGENEITY TESTS AND OTHER CHECKING SIMULATIONS

The most basic form of validation is checking for coding and data-handling errors. One effective way to do this is to run simulations for which the correct answers are known *a priori*.

A check favored by researchers who formulate and solve their models in levels of the variables is to run a simulation with the exogenous variables set at their base-period levels.[1] This computation should reproduce the base-period values for the endogenous variables. For Johansen-style modelers who formulate and solve their models in deviations of the variables away from the base-period solution, a base-period reproduction simulation requires zero shocks. Consequently, it produces zero effect and does not provide a coding check. However, there are several other checking simulations that can be applied by both types of modelers.

The most commonly used checking simulation is a nominal homogeneity test. If a CGE model is set up with no nominal rigidities,[2] then a 10% shock to all of the exogenous nominal variables should increase all endogenous nominal variables by 10% while leaving all real variables unchanged. Another check is a real homogeneity test. If a model is set up with constant returns to scale in all production activities, then a 10% shock to all real exogenous variables should increase all real endogenous variables by 10% while leaving all nominal variables unchanged. In applying these tests, it is sometimes a little difficult to decide what variables should be shocked, but working this out is instructive. For example, consider a nominal homogeneity test in a one-country model in which the exchange rate and foreign currency prices of imports are exogenous. In this case the exchange rate is a nominal variable that should be shocked: it is the reciprocal of the *domestic* dollar price of a foreign dollar. The foreign currency prices of imports are "real" variables which should not be shocked: they do not involve domestic dollars in their definition. Next, consider a real homogeneity test. While it is clear that shocks should be applied to exogenous quantities of factor inputs and exogenously specified real demands (e.g. demands by government), it might not be immediately obvious what to do about exports. If we are conducting a 10% real homogeneity test in a one-country model, we will need to shift export-demand curves 10% to the right to represent a 10% increase in the size of the world economy to match the 10% increase in the size of the domestic economy. Otherwise the increase in the size of the domestic economy will induce real effects via changes in the terms of trade.

Other checking simulations can be devised to suit particular models. For example, in a one-country model we can shift export-demand and import-supply curves vertically upwards by 10%, implying a 10% increase in foreign currency prices for any given quantities. If the exchange rate is exogenous and held constant but all other exogenous nominal variables are shocked by 10%, then in the absence of nominal rigidities, the result should be the same as for the nominal homogeneity test: 10% increases in

1 See Adelman and Robinson (1978, p. 65).

2 By nominal rigidities, we mean relationships that prevent some prices from adjusting fully over the time horizon of interest to an increase in the overall price level. An example is sluggish adjustment of nominal wages (discussed later in this section).

endogenous nominal variables and zero change in endogenous real variables. Further possible checking simulations are described in Dixon *et al.* (1992, pp. 246–251).

19.2.1 Using checking simulations to reveal properties of a model: The effects of appreciation in China

Complete documentation of detailed policy-relevant CGE models is time-consuming to prepare and time-consuming to read. Regrettably this means that complete documentation is often unavailable and even when it is available is often not very useful. So how can we become familiar with a model created by someone else?

One way is to apply a battery of checking simulations. These not only test a model's code but they can also be a way of discovering non-standard features of a model. In this section we work through the example of the 57-industry McHuge model[3] of China built by our colleague Yinhua Mai and applied in several papers, including Mai (2006) and Peng and Mai (2008).

McHuge is in the MONASH family of models (see Dixon *et al.* in Chapter 2 of this Handbook). It is recursive dynamic and annual. Simulation of the effects of policy and other shocks requires two runs of the model: a baseline run and a perturbation run. The two runs usually have different closures. If the exogenous variables in the perturbation run are set at the same values that they had either endogenously or exogenously in the baseline, then the perturbation run produces the same results as the baseline. However, in perturbation runs we shock some of the exogenous variables away from their baseline values. The effects of these shocks are revealed as differences between the perturbation and baseline values of variables that are endogenous in the perturbation run.

To get us started in understanding McHuge, we set up a perturbation closure in which the exchange rate is exogenous. Other exogenous variables included foreign currency prices for China's imports and the positions of foreign-demand curves for China's exports.[4] On shocking the exchange rate away from its baseline path we obtained the results in columns (1) and (2) of Table 19.1. These result show the effects of the shock as percentage deviations away from the baseline.

The shock we applied was an appreciation of 11.11% in 2006. This higher exchange rate was maintained in 2007. In McHuge the exchange rate is defined as the number of foreign dollars per domestic dollar (in this case RMB). Thus in the absence of nominal rigidities we would expect the effect to be a 10% reduction [$= 100 * (1/1.1111 - 1)$] in domestic prices with no changes to real variables. As can be seen from columns (1)

3 *M*onash *C*hina *H*unan *U*niversity *G*eneral *E*quilibrium model.

4 In a two-country world, China and rest of the world (ROW), an appreciation of the RMB could lead to an adjustment in China's price level or an adjustment in the ROW price level or both. These last two exogeneity assumptions imply that ROW is large relative to China so that appreciation of the RMB affects the Chinese price level rather than that of ROW.

Table 19.1 McHuge results for China on the effects of an 11.11% appreciation the RMB (% deviations from baseline)

		Sticky nominal wage rates		Sticky real wage rates		Sticky real wages and exogenous consumption	
		2006 (1)	2007 (2)	2006 (3)	2007 (4)	2006 (5)	2007 (6)
1	Exchange rate ($Foreign/RMB)	11.11[a]	11.11[a]	11.11[a]	11.11[a]	11.11[a]	11.11[a]
2	Price deflator for GDP	−8.79	−9.42	−9.95	−10.07	−10.01	−10.01
3	Price of private and public consumption	−9.18	−9.60	−9.94	−10.06	−10.00	−10.00
4	Real private and public consumption	−2.95	−1.58	0.14	−0.12	0.00[a]	0.00[a]
5	Real investment	−3.25	−1.07	0.03	−0.03	0.00	0.00
6	Export volumes	−3.90	−1.97	−0.14	0.12	0.00	0.00
7	Import volumes	−3.00	−1.33	0.03	−0.02	0.00	0.00
8	Real GDP	−3.53	−1.67	0.01	0.00	0.00	0.00
9	Aggregate employment	−5.84	−2.54	0.00	0.00	0.00	0.00
10	Aggregate capital	0.01	−0.46	0.00	0.01	0.00	0.00
11	Real wage	4.14	1.54	0.00	0.00	0.00	0.00
12	Average propensity to consume out of GNP	0.00[a]	0.00[a]	0.00[a]	0.00[a]	−0.14	0.12
13	Nominal GNP	−11.86	−11.03	−9.81	−10.17	−9.88	−10.11
14	Nominal GDP	−12.01	−10.93	−9.94	−10.06	−10.00	−10.00

[a]Exogenous in the perturbation run.

and (2), the price effects are less than a 10% reduction: reductions of 8.79 and 9.18% in the price deflators for GDP and consumption in 2006 (rows 2 and 3) and 9.42 and 9.60% in 2007. For real variables the effects are far from zero. The results for employment, GDP and real expenditure aggregates are strongly negative in 2006 and somewhat less negative in 2007.

The key to these negative results is the real wage effect: a deviation of 4.14% in 2006 and 1.54% in 2007 (row 11). Higher real wages reduce employment and GDP. This in turn reduces the expenditure aggregates. But why does our McHuge simulation show increases in real wages in response to appreciation? We found the answer in a nominal rigidity built into the specification of wage movements. In the perturbation closure adopted in our initial appreciation simulation, nominal wage adjustment is sticky. Appreciation reduces the prices of imports and exportables in China, thereby reducing the overall price level. Sticky adjustment of nominal wages leaves real wages elevated, at least temporarily.

In algebraic terms the sticky nominal wage *specification* in McHuge is:

$$\left\{\frac{W(t)}{W_{\mathrm{b}}(t)}-1\right\} = \left\{\frac{W(t-1)}{W_{\mathrm{b}}(t-1)}-1\right\} + \alpha * \left(\frac{L(t)}{L_{\mathrm{b}}(t)}-1\right). \tag{19.1}$$

In this equation the subscript b indicates a baseline value, i.e. a value in the run without the policy or other shock under consideration. $W_{\mathrm{b}}(t)$ and $L_{\mathrm{b}}(t)$ are the nominal wage rate and the level of employment in year t in the baseline. $W(t)$ and $L(t)$ are the nominal wage rate and the level of employment in year t in the perturbation run. α is a positive parameter.

Under (19.1), we assume in perturbation runs that the deviation in the wage rate from its baseline level in year t will be the same as that in year $t-1$ if employment in year t is at its baseline level. If employment in the perturbation run in year t is above (below) its baseline level then the deviation in the wage rate will increase (decrease), driving employment back towards it baseline level. This specification is consistent with an exogenous non-accelerating inflation rate of unemployment (NAIRU) and explains why the 2007 results in column (2) of Table 19.1 are closer to the expected effects in a homogeneity test than those for 2006 in column (1).

To test whether there are other nominal rigidities affecting McHuge results, we reran the appreciation experiment with wage specification (19.1) being implemented with real wages [nominal wages deflated by the consumer price index (CPI)] rather than nominal wages. The results are in columns (3) and (4) of Table 19.1. They are close to the homogeneity ideal: −10 for nominal variables and zero for reals. However, they are not close enough for us to stop the search for nominal rigidities that may be hiding errors.

Looking down column (3) of Table 19.1 we see that there is a noticeable gap between the deviations in nominal GNP (−9.81%, row 13) and nominal GDP (−9.94%, row 14).

We also see that the average propensity to consume (combined private and public) out of GNP is exogenous, unaffected by the shock to the exchange rate. This means that nominal consumption moves in line with nominal GNP. While the result for real GDP in column (3) is close to zero, the gap between GNP and GDP allows the deviation in consumption to move away from zero (0.14%, row 4). On this basis, we conclude that a potentially fruitful area in which to look for the source of non-homogeneity in column (3) is the McHuge treatment of interest and dividends on China's foreign assets and liabilities. These interest and dividend flows are the difference between GNP and GDP.

China's foreign assets are predominantly bonds issued by foreign governments. In McHuge, the RMB value of interest on these bonds is calculated by applying an exogenous interest rate to the average RMB value of assets through the year. Exchange rate appreciation is modeled as taking place in the middle of the year. Consequently, according to McHuge an 11.11% appreciation reduces the average RMB value of China's foreign assets by approximately 5% (a 10% reduction in the end-of-year value but zero effect on the start-of-year value). This gives a 5% reduction in the RMB value of interest receipts. China's foreign liabilities are predominantly equity holdings by foreigners in businesses operating in China. With a 10% reduction in the Chinese price level, McHuge generates a 10% reduction in the RMB value of profits accruing to foreigners. This asymmetric treatment of interest on assets and dividends on liabilities in the first year of the appreciation generates an increase in China's GNP relative to its GDP, which in turn explains the increase in consumption that we see in row 4, column (3) of Table 19.1. However, the timing in McHuge of within-year interest and dividend payments is arbitrary. While our test simulation has not revealed an error (a result that does not follow from the theory and data) it has identified a shortcoming of the model. Until the model is refined, results that depend critically on within-year timing of interest and dividend payments on foreign assets and liabilities should be treated skeptically.

To continue the search for errors and modeling imperfections, we ran a further appreciation simulation in which the link between GNP and consumption is eliminated. This is done by endogenizing the average propensity to consume and exogenizing real consumption with zero shock. The results are in columns (5) and (6) of Table 19.1. Reflecting the within-year timing problem in the modeling of interest/dividend flows on foreign assets and liabilities, the deviation in GNP continues to show non-homogeneity [−9.88% in row 13 of column (5) rather than −10]. But this effect is now absorbed by a movement in the average propensity to consume (−0.14%, row 12). With no link between consumption and GNP, the simulation shows the ideal homogeneity result, indicating that we have now exhausted what can be found out about nominal rigidities and errors via the exchange rate homogeneity test.

Apart from being a checking device, the exchange rate homogeneity test has an important policy interpretation. Over the last few years, the US has been advising China to appreciate. The US authorities argue that this would help to alleviate what they see as a harmful imbalance in the trade between the two countries. The Chinese authorities have been reluctant to follow US advice. They are wary of the possibility [supported by columns (1) and (2) in Table 19.1] that appreciation could cause strongly negative real impacts.

From the Chinese point of view, the least damaging results occur when there are no significant nominal price rigidities. In these circumstances, the results in columns (3)−(6) of Table 19.1 imply that the main macro effects of appreciation would be an alteration in the Chinese rate of inflation. There would be little impact on real variables including imports and exports. Thus there would be little impact on the foreign currency value of Chinese surpluses on trade and current account. At the same time, there might be unintended consequences for the distribution of wealth in China with gains for creditors and losses for debtors.

However, the assumption that all prices in China are fully flexible is unrealistic. With sluggish adjustment in some RMB prices, appreciation causes real effects. For example, if wage rates in RMB terms are slow to adjust downwards then real appreciation can cause significant negative effects on output and employment. This is illustrated in Figure 19.1 that presents McHuge results from columns (1) and (2) of Table 19.1, and extends them to cover the period out to 2015.

To us, the main economic message from the exchange rate homogeneity test is that global trade imbalances are unlikely to be addressed successfully purely through exchange rate adjustments. What is required is changes in saving behavior. If we care about

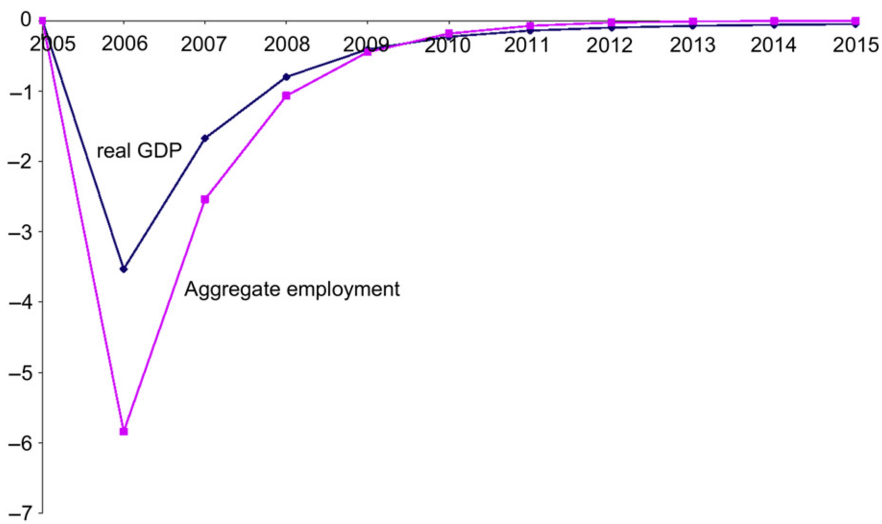

Figure 19.1 Effects on China's employment and GDP of an 11.11% appreciation under sticky adjustment of nominal wage rates (% deviations from baseline).

Chinese surpluses and US deficits, then we must look for policies that increase Chinese absorption while reducing US absorption.

19.3 VALIDATION THROUGH THE GDP IDENTITY

All economists know that GDP can be calculated as a sum of final demands or as a sum of incomes:

$$\mathrm{GDP} = P_{\mathrm{c}} * C + P_{\mathrm{i}} * I + P_{\mathrm{g}} * G + P_{\mathrm{x}} * X - P_{\mathrm{m}} * M \tag{19.2}$$

$$\mathrm{GDP} = \underset{\text{(factor incomes)}}{\sum_i W_i * F_i} + \underset{\text{(indirect taxes)}}{\sum_k T_k * P_k * X_k}, \tag{19.3}$$

where C, I, G, X and M are quantities of private consumption, investment, public consumption, exports and imports, and the Ps are corresponding price deflators; F_i and W_i are the employment and returns (wages or rentals) for factor i; X_k and P_k are the quantity and associated price for sale k; and T_k is the *ad valorem* tax rate applying to sale k. However, economists are sometimes less clear about what kind of prices and returns should be used in (19.2) and (19.3). Are the Ps purchasers, producer, basic, c.i.f. (cost, insurance and freight) or f.o.b. (free on board) prices? Are the returns pre or post tax?

Another set of questions relates to how to move from nominal to real GDP. The change in real GDP can be defined from (19.2) by:

$$100 * \Delta\mathrm{RealGDP} = P_{\mathrm{c}} * C * c + P_{\mathrm{i}} * I * i + P_{\mathrm{g}} * G * g + P_{\mathrm{x}} * X * x - P_{\mathrm{m}} * M * m, \tag{19.4}$$

where c, i, g, x and m are percentage changes in the quantities C, I, G, X and M. But how should we value $P_{\mathrm{c}} * C$, $P_{\mathrm{i}} * I$, etc.? If the percentage changes refer to movements between years t and $t+1$, should $P_{\mathrm{c}} * C$, $P_{\mathrm{i}} * I$, etc., refer to year t, year $t+1$ or somewhere in between? Finally, how can we calculate the change in real GDP from (19.3)? One intuitively appealing but false method is to write:

$$100 * \Delta\mathrm{RealGDP} = \sum_i W_i * F_i * f_i + \sum_k T_k * P_{\mathrm{k}} * X_k * x_k \quad \text{FALSE}, \tag{19.5}$$

where f_i and x_k are percentage changes in the quantities F_i and X_k. Formula (19.5) is clearly inadequate and will not in general produce results which match those obtained via (19.4). This is because (19.5) leaves out technical change. It falsely implies that if all tax rates are zero then real GDP can change only if there is a change in employment of factors.

In this section, we clarify these issues by deriving GDP identities in the context of CGE modeling. We start by looking at an input–output database for a detailed CGE model, one that identifies: production activities; capital creating activities; final demands by households, investors, government and foreigners (exports); demands for primary factors; and the taxes, technical changes and margins associated with every commodity and factor flow. Using this database we establish the GDP identity in nominal terms: we show that GDP defined as the purchasers value of private consumption plus investment plus public consumption plus exports less the c.i.f. value of imports equals GDP defined as pretax returns to primary factors plus sales taxes.[5] We then establish in (19.22) and (19.23) an identity in real change terms.

We have two reasons for devoting space to the GDP identity in this chapter. First, it provides a powerful device for checking the validity of CGE calculations. In any CGE solution, GDP calculated as nominal income [the right-hand side of (19.3)] should match GDP calculated as nominal expenditure [the right-hand side of (19.2)]. Similarly changes in real GDP calculated from real incomes [the right-hand side of (19.22) below] should match those calculated from real expenditures [the right-hand side of (19.23) below]. These checks are powerful because the two measures of GDP involve distinct sets of variables which are linked indirectly through a large number of equations in the CGE framework.

Our second reason for emphasizing GDP identities will become apparent in Section 19.4 where we will be concerned with explaining CGE results. Since CGE models include many direct and indirect connections between variables, it is often challenging to break into a set of results and explain them in a way that avoids circularity. We do not want to say that consumption increased because income increased and income increased because there was an increase in demand by consumers. What we want is a sequence in which the result for endogenous variable A is explained in terms of the theory and data of the model and the shocks to exogenous variables. Then we can explain endogenous variable B in terms of A and theory, data and exogenous variables, and so on. The GDP identity, particularly the real income measure, is a valuable starting point for organizing a non-circular explanation of CGE results. It also connects CGE results to standard microeconomic concepts such as consumer surplus, producer surplus, deadweight loss, input-saving technical change and marginal productivity of capital and labor.

19.3.1 Input-output database for a CGE model

Table 19.2 is a representation of an input-output database for a single country CGE model. The (1,1)-block shows sales of domestically produced commodities to activities. Activities produce commodities and units of capital.[6] For concreteness and simplicity, we

[5] In this exposition we treat all indirect taxes as sales taxes.

[6] In most CGE models the columns of inputs to current production and the columns of inputs to capital creation are presented in separate matrices. Here, it is algebraically convenient to present all these columns in a single matrix.

Table 19.2 Input-output database for a CGE model

		Activities: production of commodities and units of capital	**Final demands: purchases by households, government, investors and foreigners**	**Totals**
		1	**2**	
Domestic products: commodities and units of capital	1	$P(k) * X(k,j)$ $k \in A,\ j \in A$	$P(k) * X(k,j)$ $k \in A,\ j \in D$	Direct demands
Imported commodities	2	$P(k) * X(k,j)$ $k \in NM,\ j \in A$	$P(k) * X(k,j)$ $k \in NM,\ j \in D$	Imports by commodity
Primary factors	3	$P(k) * X(k,j)$ $k \in NF,\ j \in A$	$P(k) * X(k,j)$ $k \in NF,\ j \in D$	Returns to primary factors
Margins	4	$P(m) * XM(m,k,j)$ $m \in A,\ k \in A \cup N,\ j \in A$	$P(m) * XM(m,k,j)$ $m \in A,\ k \in A \cup N,\ j \in D$	Margin demands
Sales taxes	5	$T(k,j) * P(k) * X(k,j)$ $k \in A \cup N,\ j \in A$	$T(k,j) * P(k) * X(k,j)$ $k \in A \cup N,\ j \in D$	Sales taxes by taxed object
Total		Activity outputs	Totals for C, I, G and X	

assume that each commodity is produced by just one activity (or industry). Similarly we assume that each type of capital is produced by just one activity. By types of capital we have in mind industry specificity, with each industry using its own particular type of capital.[7] Thus, for example, the agricultural-capital-creating activity uses substantial inputs of agricultural machinery in creating capital for the activity that produces the agricultural commodity.

The row label for the (1,1)-block includes both domestic products and units of capital. In practice much of the (1,1)-block will be empty because there are no sales of units of capital to activities. Units of capital are sold to investors who rent them out to commodity-producing activities.

Each element in the (1,1)-block is a quantity [$X(j,k)$] multiplied by a basic price [$P(k)$]. Basic prices are those received by producers. They do not include sales taxes and margins (transport, retail trade, wholesale trade) that come between producers and users. As indicated in the (1,1)-block, the row and column subscripts, j and k, range over all the elements of the set A. This is both the set of all activities and the set of all commodities and capital types.

The (2,1)-block shows the basic value of sales of imported commodities to activities. We define the basic price of an imported commodity as the c.i.f. price.[8] The set of all imported commodities is denoted as NM, "N" for not-produced and "M" for imports.

Flows of primary factor services to activities are shown in block (3,1). This covers payments for labor, and rentals on capital and land. Basic prices for primary factors exclude sales taxes if any, but include income taxes. We avoid a user subscript on the basic price by imagining that there are many primary factors: a different type of capital for each industry and possibly many occupations and types of land. With many types of primary factors, it is legitimate to assume that any given primary factor is supplied to all users at a single basic price. Block (3,1) can contain many zeros: the commodity-i-producing activity will use capital of type i but not capital of type j; and capital-creating activities are often specified as using no primary factors. The set of all primary factors is denoted as NF, "N" for not-produced[9] and "F" for factor.

$XM(m,k,j)$ in block (4,1) is the use of margin service m to facilitate the flow of k to activity j. While we allow every $m \in A$ to act as a margin, we preclude the possibility of imported margins. This is reasonable because we are concerned only with facilitating flows within a country. This includes flows of exports to ports of exit and flows of

[7] If there are n industries, then the column dimension of the first column of matrices in Table 19.2 is $2n$.

[8] This is the price at the port of entry before customs. Basic prices for imported commodities are often defined as the c.i.f. price plus tariffs. For our present purposes it is convenient to treat tariffs as sales taxes applied to imports.

[9] In the case of labor and land, the idea of "non-produced" is clear. For capital, we assume that what is in use in production processes was not *currently* produced, it was produced in an earlier period.

imports after they have arrived at the port of entry.[10] In practice, block (4,1) contains many zeros. Only a few domestically-produced commodities such as transport, retail trade and wholesale trade fulfill the role of margins, and in most CGE models, it is assumed that primary factors are supplied without the assistance of margin services. However, it is notationally convenient to have m range over the whole of the set A and k range over $A \cup N$, where N is the set of non-produced inputs, i.e. $N = NM \cup NF$. A final noteworthy aspect of block (4,1) is that it does not allow for margin services in the delivery of margins. We assume that there are none.

In block (5,1), $T(k,j)$ is the sales tax rate applying to the basic value of the flow of k to activity j. This can refer to any (k,j) in blocks (1,1), (1,2) and (1,3). It is convenient to assume that there are no sales taxes on margins. A model incorporating this assumption can generate the effects of road transport sales taxes, for example, via taxes on inputs to the road-transport-producing activity. However, a limitation of this approach is that it rules out the possibility of differences in the sales tax rates applying to the use of road transport in facilitating the delivery of different kinds of goods.

Summing down the jth column in blocks (1,1), (2,1), (3,1), (4,1) and (5,1) gives the cost of all inputs to activity j. Under the zero-pure-profits condition, this is the basic value of the output of activity j.

Final users are identified in Table 19.2 as members of the set D. This includes households, government, investors and foreigners (demanders of exports). Basic values of flows of commodities, units of capital, imports and primary factors to final users are shown in blocks (1,2), (2,2) and (19.3). Associated margins and sales taxes are in blocks (4,2) and (5,2). In practical CGE modeling many of the flows in the final demand blocks are zero. Investors buy only units of capital and, in most models, there are no primary factor flows to final demands.

Summing down columns in blocks (1,2), (2,2), (3,2), (4,2) and (5,2) gives the national accounts aggregates for consumption (C), investment (I), government expenditure (G) and exports (X).

19.3.2 Input-output balance conditions and the GDP identity in nominal terms

Published data of the type illustrated in Table 19.2 usually satisfy a balance condition: the basic value of output from activity j equals the basic value of demands for the activity's product. If this condition is not initially satisfied then CGE modelers make adjustments so that it is satisfied. With the balance condition satisfied for all individual activities, it is satisfied in aggregate, that is the sum of the basic values of activity outputs equals the sum of the basic values of demands for these outputs. In terms of Table 19.2, the sum of the

[10] The country may provide transport and other margin services outside its borders. However these services are normally treated in CGE models as direct exports.

entries in blocks (1,1) to (5,1) equals the sum of the entries in blocks (1,1), (1,2), (4,1) and (4,2). Eliminating the common blocks (1,1) and (4,1) we can write:

$$\underbrace{\sum_{k\in N}\sum_{j\in A} P(k)*X(k,j)}_{\text{(non-produced inputs to activities, basic value)}} + \underbrace{\sum_{k\in A\cup N}\sum_{j\in A} T(k,j)*P(k)*X(k,j)}_{\text{(sales taxes on input to activities)}}$$
$$= \underbrace{\sum_{k\in A}\sum_{j\in D} P(k)*X(k,j)}_{\text{(final demands, domestic items, basic)}} + \underbrace{\sum_{m\in A}\sum_{k\in A\cup N}\sum_{j\in D} P(m)*XM(m,k,j)}_{\text{(margins on final demands)}}. \tag{19.6}$$

Now we work towards the GDP identity by making three adjustments to (19.6). Our strategy is to assemble the income-side measure of GDP on the left-hand side and in the process reveal the expenditure-side measure of GDP on the right-hand side. First we add the sales taxes on final demands to both sides. Second we take the imports out of the first term on the left-hand side and subtract them from the right-hand side. Third we add to the left-hand side the sales to final demand of primary factors. The balancing adjustment on the right-hand side is achieved by: (i) Adding to the first term the sales to final demand of all non-produced items (factors and imports) and (ii) subtracting import sales to final demand. After these adjustments (19.6) becomes:

$$\underbrace{\sum_{k\in NF}\sum_{j\in A\cup D} P(k)*X(k,j)}_{\text{(primary factor inputs to activities and final demand, basic)}} + \underbrace{\sum_{k\in A\cup N}\sum_{j\in A\cup D} T(k,j)*P(k)*X(k,j)}_{\text{(all sales taxes)}}$$
$$= \underbrace{\sum_{k\in A\cup N}\sum_{j\in D} P(k)*X(k,j)}_{\text{(all final demands, basic)}} + \underbrace{\sum_{m\in A}\sum_{k\in A\cup N}\sum_{j\in D} P(m)*XM(m,k,j)}_{\text{(margins on final demands)}}$$
$$+ \underbrace{\sum_{k\in A\cup N}\sum_{j\in D} T(k,j)*P(k)*X(k,j)}_{\text{(sales taxes on final demands)}} - \underbrace{\sum_{k\in NM}\sum_{j\in A\cup D} P(k)*X(k,j)}_{\text{(all imports c.i.f. value)}}. \tag{19.7}$$

Next we combine the first three terms on the right-hand side of (19.7) by introducing the concept of purchasers prices. This gives:

$$GDPINC \equiv \underbrace{\sum_{k\in NF}\sum_{j\in A\cup D} P(k)*X(k,j)}_{\text{(primary factor inputs to activities and final demand, basic)}} + \underbrace{\sum_{k\in A\cup N}\sum_{j\in A\cup D} T(k,j)*P(k)*X(k,j)}_{\text{(all sales taxes)}}$$
$$= \underbrace{\sum_{k\in A\cup N}\sum_{j\in D} P1(k,j)*X(k,j)}_{\text{(all final demands, purchasers prices)}} - \underbrace{\sum_{k\in NM}\sum_{j\in A\cup D} P(k)*X(k,j)}_{\text{(all imports c.i.f. value)}} \equiv GDPEXP. \tag{19.8}$$

where $P1(k,j)$ is the purchasers price (basic price plus sales taxes plus margins) of item k to user j, defined via:

$$\underset{\text{(purchasers' value)}}{P1(k,j) * X(k,j)} = \underset{\text{(basic value)}}{P(k,j) * X(k,j)} + \underset{\text{(sales taxes)}}{T(k,j) * P(k) * X(k,j)} + \underset{\text{(margins)}}{\sum_{m \in A} P(m) * XM(m,k,j)} . \tag{19.9}$$

Equation (19.8) is the familiar GDP identity. It says that payments to primary factors plus indirect taxes (GDP from the income side, *GDPINC*) equals the purchasers value of sales to households, investors, government and foreigners (exports) less the c.i.f. value of imports (GDP from the expenditure side, *GDPEXP*).

Invariably, CGE models impose the conditions of zero pure profits and market clearing. With these conditions, the GDP identity (19.8) should continue to hold throughout all simulations. This provides a good check on CGE computations. As a simulation takes us away from the initial database, no discrepancies should emerge between GDP calculated according to the left-hand side of (19.8) and GDP calculated according to the right-hand side. If discrepancies do emerge, then the model has not been properly implemented. So that we can check for discrepancies we include in our models two nominal GDP variables, *GDPINC* and *GDPEXP* defined separately by the left and right-hand sides of (19.8).[11]

19.3.3 GDP identity in real terms

In the previous section we derived the equality between the income and expenditure measures of nominal GDP. In doing this we used zero-pure-profit and market-clearing conditions. Deriving a useful GDP identity in real terms requires deeper use of behavioral assumptions built into CGE models, particularly cost-minimizing assumptions. In addition, the concept of real GDP is best thought of as an index with meaning only as an indicator of change. Thus, in this section we concentrate on defining expenditure and income measures of *changes* in real GDP. These can be changes through time or they can be policy-induced changes at a given point of time.

The derivation of the real GDP identity is quite intricate. However the final result, equation (19.21), is readily interpretable. Readers who do not want to follow the algebraic process can move straight to (19.21) and then look at our interpretation in Section 19.3.4.

[11] Equation (19.8) is not included in our models because, as demonstrated in this subsection, it is implied by the zero pure profits and market clearing conditions. Inclusion of (19.8) would make the model singular.

19.3.3.1 Production functions, technical change, cost minimization and zero pure profits

We consider a model in which activity j chooses inputs to minimize costs subject to a constant-returns-to-scale production function incorporating input-saving technical change, i.e. for all $j \in A$:

$$X(k,j), \quad k \in A \cup N,$$

minimizes:

$$\sum_{k \in A \cup N} P1(k,j) * X(k,j), \tag{19.10}$$

subject to:

$$Y(j) = F_j\left(\frac{X(k,j)}{A(k,j)}, \quad k \in A \cup N\right), \tag{19.11}$$

where $Y(j)$ is the quantity of output from activity j, F_j is j's constant-returns-to-scale production function and A(k,j) is a variable allowing for k-saving technical change in activity j. [A 1% reduction in A(k,j) introduces 1% k-saving technical change − it allows X(k,j) to be reduced by 1% with no change in any other input to j and no change in j's output.] We assume that there are zero pure profits in each activity, that is the value of output from activity j, $P(j) * Y(j)$ equals the value of inputs. Then, as shown in the Appendix, (19.10) and (19.11) lead to:

$$\begin{aligned} P(j) * Y(j) * y(j) = & \sum_{k \in A \cup N} P1(k,j) * X(k,j) * x(k,j) \\ & - \sum_{k \in A \cup N} P1(k,j) * X(k,j) * a(k,j), \quad j \in A. \end{aligned} \tag{19.12}$$

Where $y(j)$, $x(k,j)$ and $a(k,j)$ are percentage changes in the variables denoted by the corresponding upper case symbols. Equation (19.12) implies that the percentage change in output in activity j is a weighted average of the percentage changes in the technology-augmented inputs $[x(k,j) - a(k,j)]$ where the weights are cost shares, $[P1(k,j) * X(k,j)/(P(j) * Y(j))]$.

Similarly, we assume that the composition of margin services (e.g. rail versus road transport) chosen to facilitate any flow of n to m for $n \in A \cup N$ and $m \in A \cup D$: minimizes:

$$\sum_{k \in A} P(k) * XM(k,n,m), \tag{19.13}$$

subject to:

$$X(n,m) = FM_{n,m}\left(\frac{XM(k,n,m)}{AM(k,n,m)}, \quad k \in A\right), \tag{19.14}$$

where $FM_{n,m}$ is a constant-returns-to-scale production function describing the overall level of margin service generated to facilitate the flow of n to m; and $AM(k,n,m)$ is a variable allowing for k-saving technical change in the generation of margin services for the delivery of n to m. In (19.14) we assume that one unit of aggregate margin service is required to facilitate the flow of one unit of n to m. This is not a limiting assumption. It can be accommodated by scaling the FM function. Also note that basic prices are used in (19.13). This reflects our earlier assumption that there are no sales taxes or margins on margins.

From (19.13) and (19.14) we can obtain:

$$\begin{aligned}\sum_{k \in A} P(k) * XM(k,n,m) * x(n,m) &= \sum_{k \in A} P(k) * XM(k,n,m) * xm(k,n,m) \\ &\quad - \sum_{k \in A} P(k) * XM(k,n,m) * am(k,n,m), \quad n \in A \cup N \text{ and } m \in A \cup D.\end{aligned} \tag{19.15}$$

If we divide (19.15) through by $\sum_{k \in A} P(k) * XM(k,n,m)$, then it says that the percentage change in overall margin requirement [which is simply $x(n,m)$] to facilitate the flow of n to m is a weighted average of the percentage changes in the technology-augmented inputs [$xm(k,n,m) - am(k,n,m)$] to the creation of n-to-m facilitating services. The weights are shares for each type of margin in the total margin cost for the n-to-m flow.

19.3.3.2 Market clearing

Apart from cost-minimizing behavior and the assumption of zero pure profits in each activity, the only other economic assumption needed in the derivation of our real GDP identity is market clearing. We write this in a change form as:

$$\begin{aligned}Y(k) * y(k) &= \sum_{j \in A \cup D} X(k,j) * x(k,j) \\ &\quad + \sum_{n \in A \cup N} \sum_{m \in A \cup D} XM(k,n,m) * xm(k,n,m) \quad \text{for all } k \in A.\end{aligned} \tag{19.16}$$

The left-hand side of (19.16) is 100 times the change in the output of activity k. The terms on the right-hand side are 100 times the changes in demands for k. Multiplying through by $P(k)$ and aggregating over k gives:

$$\underset{\text{(basic value of activity outputs)}}{\sum_{k\in A} P(k) * Y(k) * y(k)} = \underset{\text{(basic value of direct demand for activity outputs)}}{\sum_{k\in A}\sum_{j\in A\cup D} P(k) * X(k,j) * x(k,j)}$$

$$+\underset{\text{(basic value of demands for activity outputs to be used as margins)}}{\sum_{k\in A}\sum_{n\in A\cup N}\sum_{m\in A\cup D} P(k) * XM(k,n,m) * xm(k,n,m)}. \quad (19.17)$$

All the terms in (19.17) are 100 times the change in a value associated with changes in quantities. However in the captions below the terms we adopt a shorthand: we just mention the relevant value.

19.3.3.3 Deriving the real GDP identity

Our first step in deriving the real GDP identity is to insert purchasers prices into the first term on the right-hand side of (19.17) with corresponding subtractions of taxes and margins. In doing this we use (19.9) and obtain:

$$\begin{aligned}
\underset{\text{(basic value of activity outputs)}}{\sum_{k\in A} P(k) * Y(k) * y(k)} &= \underset{\text{(purchasers value of direct demand for activity outputs)}}{\sum_{k\in A}\sum_{j\in A\cup D} P1(k,j) * X(k,j) * x(k,j)} \\
&- \underset{\text{(sales taxes on direct demand for activity outputs)}}{\sum_{k\in A}\sum_{j\in A\cup D} T(k,j) * P(k) * X(k,j) * x(k,j)} \\
&- \underset{\text{(basic value of margins on flows of activity outputs, assuming no technical change)}}{\sum_{k\in A}\sum_{j\in A\cup D}\left[\sum_{m\in A} P(m) * XM(m,k,j)\right] * x(k,j)} \\
&+ \underset{\text{(basic value of demands for activity outputs to be used as margins on all flows).}}{\sum_{k\in A}\sum_{n\in A\cup N}\sum_{m\in A\cup D} P(k) * XM(k,n,m) * xm(k,n,m)}
\end{aligned} \quad (19.18)$$

Next we extend the first term on the right-hand side of (19.18) to cover all inputs $k \in A \cup N$ and then split it into two terms on the j dimension: $j \in A$ and $j \in D$. Corresponding to the extension we must make a subtraction. In making the subtraction we include a new term [the last term on the right-hand side of (19.19)] to cover the basic value of the extension and we adjust the domains of the tax and margin terms to cover the rest of the extension. The result of these operations is:

$$\begin{aligned}
&\underbrace{\sum_{k\in A} P(k)*Y(k)*y(k)}_{\text{(basic value of activity outputs)}} = \underbrace{\sum_{k\in A\cup N}\sum_{j\in A} P1(k,j)*X(k,j)*x(k,j)}_{\text{(purchasers value of all inputs to activities)}} \\
&+ \underbrace{\sum_{k\in A\cup N}\sum_{j\in D} P1(k,j)*X(k,j)*x(k,j)}_{\text{(purchasers value of final demand)}} - \underbrace{\sum_{k\in A\cup N}\sum_{j\in A\cup D} T(k,j)*P(k)*X(k,j)*x(k,j)}_{\text{(all sales taxes)}} \\
&- \underbrace{\sum_{k\in A\cup N}\sum_{j\in A\cup D}\left[\sum_{m\in A} P(m)*XM(m,k,j)\right]*x(k,j)}_{\text{(basic value of margins on all flows, in the absence of technical change)}} \\
&+ \underbrace{\sum_{k\in A}\sum_{n\in A\cup N}\sum_{m\in A\cup D} P(k)*XM(k,n,m)*xm(k,n,m)}_{\text{(basic value of demands for activity outputs to be used as margins on all flows)}} \\
&- \underbrace{\sum_{k\in N}\sum_{j\in A\cup D} P(k)*X(k,j)*x(k,j)}_{\text{(basic value of demands for non–produced items)}} \quad .
\end{aligned} \tag{19.19}$$

Now we use (19.12) and (19.15) to eliminate activity inputs and outputs and margin flows:

$$\begin{aligned}
0 = &\underbrace{\sum_{k\in A\cup N}\sum_{j\in A} P1(k,j)*X(k,j)*a(k,j)}_{\text{(input–saving technical change in activites)}} + \underbrace{\sum_{k\in A\cup N}\sum_{j\in D} P1(k,j)*X(k,j)*x(k,j)}_{\text{(purchasers value of final demand)}} \\
&- \underbrace{\sum_{k\in A\cup N}\sum_{j\in A\cup D} T(k,j)*P(k)*X(k,j)*x(k,j)}_{\text{(all sales taxes)}} \\
&+ \underbrace{\sum_{k\in A}\sum_{n\in A\cup N}\sum_{m\in A\cup D} P(k)*XM(k,n,m)*am(k,n,m)}_{\text{(input–saving technical change in the provision of margins)}} \\
&- \underbrace{\sum_{k\in N}\sum_{j\in A\cup D} P(k)*X(k,j)*x(k,j)}_{\text{(basic value of demands for non–produced items)}} \quad .
\end{aligned} \tag{19.20}$$

The final step is to split the last term on the right-hand side of (19.20) into two terms on the k dimension: $k \in NM$ and $k \in NF$ and then rearrange as:

$$
\begin{aligned}
&100 * \Delta\mathrm{RealGDPINC} \equiv \\
&\underbrace{\sum_{k \in NF} \sum_{j \in A \cup D} P(k) * X(k,j) * x(k,j)}_{\text{(basic value of primary factors)}} + \underbrace{\sum_{k \in A \cup N} \sum_{j \in A \cup D} T(k,j) * P(k) * X(k,j) * x(k,j)}_{\text{(indirect taxes)}} \\
&- \underbrace{\sum_{k \in A \cup N} \sum_{j \in A} P1(k,j) * X(k,j) * a(k,j)}_{\text{(input-saving technical change in activities)}} - \underbrace{\sum_{k \in A} \sum_{n \in A \cup N} \sum_{m \in A \cup D} P(k) * XM(k,n,m) * am(k,n,m)}_{\text{(input-saving technical change in provision of margins)}} \\
&= \underbrace{\sum_{k \in A \cup N} \sum_{j \in D} P1(k,j) * X(k,j) * x(k,j)}_{\text{(purchasers value of final demand)}} - \underbrace{\sum_{k \in NM} \sum_{j \in A \cup D} P(k) * X(k,j) * x(k,j)}_{\text{(change in c.i.f. value of imports)}} \\
&\qquad \equiv 100 * \Delta\mathrm{Real}GDPEXP.
\end{aligned}
\tag{19.21}
$$

19.3.4 Interpreting the real GDP identity

For interpretation, it is helpful to divide (19.21) by the level of GDP and then split it into two equations. This gives:

$$
\begin{aligned}
\mathrm{realgdpinc} = &\underbrace{\sum_{k \in NF} \sum_{j \in A \cup D} S_{\mathrm{factor}}(k,j) * x(k,j)}_{\text{(basic value of primary factors)}} + \underbrace{\sum_{k \in A \cup N} \sum_{j \in A \cup D} S_{\mathrm{tax}}(k,j) * x(k,j)}_{\text{(indirect taxes)}} \\
&- \underbrace{\sum_{k \in A \cup N} \sum_{j \in A} S_{\mathrm{input}}(k,j) * a(k,j)}_{\text{(input-saving technical change in activities)}} \\
&- \underbrace{\sum_{k \in A} \sum_{n \in A \cup N} \sum_{m \in A \cup D} S_{\mathrm{margin}}(k,n,m) * am(k,n,m)}_{\text{(input-saving technical change in privision of margins)}}
\end{aligned}
\tag{19.22}
$$

$$
\mathrm{realgdpexp} = \underbrace{\sum_{k \in A \cup N} \sum_{j \in D} S_{\mathrm{final}}(k,j) * x(k,j)}_{\text{(purchasers value of final demand)}} - \underbrace{\sum_{k \in NM} \sum_{j \in A \cup D} S_{\mathrm{imp}}(k,j) * x(k,j)}_{\text{(change in c.i.f. value of imports)}},
\tag{19.23}
$$

where realgdpinc and realgdpexp are percentage changes in the income and expenditure measures of real GDP, and the Ss are ratios of values to GDP: $S_{\mathrm{factor}}(k,j)$ refers to payments to primary factor k in activity or final demand j, $S_{\mathrm{tax}}(k,j)$ refers to tax collections on flow (k,j), $S_{\text{input}}(k,j)$ refers to the purchasers value of flow of k to activity j, $S_{\text{margin}}(k,n,m)$ refers to the basic value of k used as a margin to facilitate flow of n to m, $S_{\text{final}}(k,j)$ refers to the purchasers value of flow k to final demand category j and $S_{\text{imp}}(k,j)$ refers to the c.i.f. value of imports of k used in activity or final demand j.

Real GDP is a measure of an economy's output: it is the quantity of goods and services that the economy supplies to final demanders net of all imports (including those used in activities as well as in final demands). Consistent with this, in specifying the percentage change in real GDP, (19.23) includes percentage changes in flows to final demands as positive terms and percentage changes in import flows as negative terms. These percentage changes are weighted according to the ratios of the underlying flows to GDP.

Equation (19.22) implies that the economy's output as measured by real GDP increases when there are: increases in primary factor inputs, increases in tax-carrying flows, input-saving technical change and margin-saving technical change. The weights attached to percentage changes in each of these variables reflect the ratio of the associated flows to GDP.

To understand (19.22) we start by ignoring for a moment tax-carrying flows: simply assume $S_{\text{tax}}(k,j)$ is zero for all (k,j). Then the remaining terms will be easily recognized as a multicommodity generalization of one-sector growth equation:

$$y = S_l * l + S_k * k + a, \tag{19.24}$$

where y, l and k are percentage changes in real GDP and inputs of labor and capital; a is total-factor-productivity growth[12]; and S_l and S_k are labor and capital shares in GDP.

The primary factor term in (19.22) generalizes the labor and capital terms in (19.24). The general form in (19.22) tells us that a 1% increase in the input of primary factor k to activity j, holding all other primary factor inputs and technology constant, generates an increase in real GDP of $S_{\text{factor}}(k,j)$%. Thus, for example, if employment in every activity increases by 1% and payments to labor are 60% of GDP, then GDP increases by 0.6%.

The technology terms in (19.22) generalize a in (19.24). The first technology term tells us that input-saving technical change of 1% operating on the flow of k to j $[a(k,j) = -1]$ will, holding other things constant, generate an increase in real GDP of $S_{\text{input}}(k,j)$%. Continuing our previous example, a 1% labor-saving technical change operating on the use of labor in every activity increases GDP by 0.6%. The second technology term in (19.22) implies, for example, that if the use of rail transport to facilitate the flow of coal to ports of exit costs 0.5% of GDP, then a technological improvement of 1% operating on this margin cost will increase GDP by 0.005%.

12 Following the usual notational convention in macroeconomics a positive value for a in (19.24) implies an improvement in technology. By contrast, positive values for the a's in (19.22) imply deteriorations in technology.

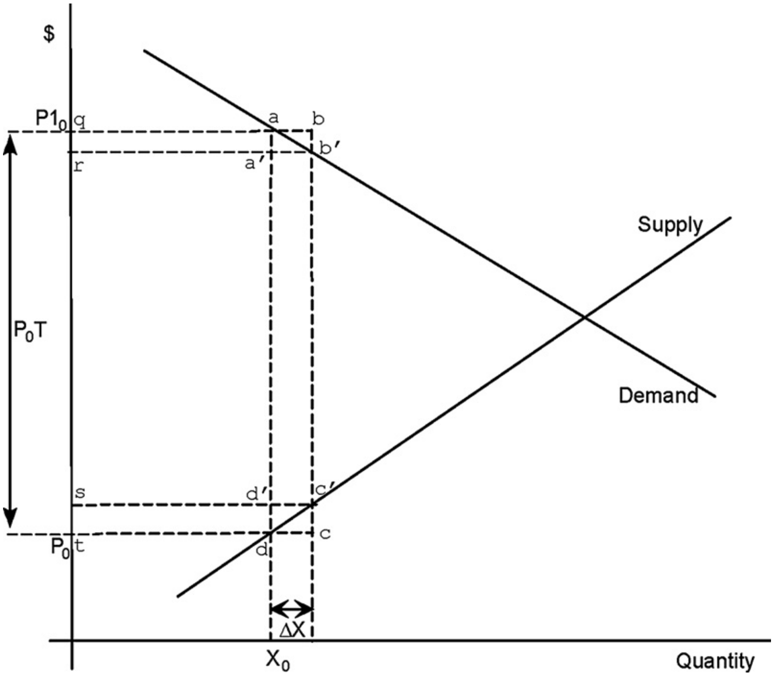

Figure 19.2 Effects of a small change in a sales tax.

The tax term in (19.22) does not have an analogue in (19.24). This term arises because indirect taxes drive a wedge between the costs to a society of producing taxed goods represented by the basic price and margin costs and the benefits represented by the purchasers price. The nature of the tax term can be understood via Figure 19.2. This shows demand and supply curves for a commodity for which the purchasers price is separated from the basic price by a tax at rate T.[13] In the initial situation, the purchasers price that motivates demanders is $P1_0$, the basic price that motivates suppliers is P_0 and the quantity sold is X_0. Now we introduce a small cut in the tax rate, reducing the gap between the purchasers price and the basic price in Figure 19.2 from **qt** to **rs**. This increases sales by ΔX. We compute the effect on real GDP as:

$$\Delta\text{RealGDP} = P1_0 * \Delta X - P_0 * \Delta X. \tag{19.25}$$

The first term on the right-hand side of (19.25) recognizes that extra units of the commodity are valued in GDP at the price paid by users. In the second term we assume that the basic price represents the opportunity cost of the resources required to supply extra units. The change in real GDP calculated on the right-hand side of (19.25) is

13 For simplicity we assume that there are no margins. Alternatively we could think of the margins as being part of the price that motivates supply.

represented in Figure 19.2 by the rectangular strip **abcd**. More precisely, the change in real GDP should be calculated as **ab′c′d**, taking account of the reduction in the purchasers price and the increase in the basic price as X increases.[14] However, provided the change in the tax rate is small, we can ignore the two triangles **abb′** and **c′cd**. Then from (19.25) we obtain:

$$\Delta \text{RealGDP} = P_0 * T * X_0 * \left(\frac{\Delta X}{X_0}\right). \tag{19.26}$$

leading to a partial equilibrium one-commodity analogue of the indirect tax term in (19.22):

$$\text{realgdp} = S_{\text{tax}} * x. \tag{19.27}$$

where realgdp and x are the percentage changes in RealGDP and X, and S_{tax} is the ratio of tax collection to GDP.

What happens if we are dealing with a large tax change? Then in terms of a demand and supply diagram, the effect on GDP cannot be adequately approximated as a rectangular strip. We must take triangles into account.

As explained in Dixon *et al.* in Chapter 2 of this Handbook, in models formulated in Johansen's linear percentage-change style, accurate solutions for the effects of large changes can be obtained by a multistep procedure. If we want to know the effects of a 100% reduction in a tax, for example, we can start by working out the effects of a 10% tax reduction using linear percentage change equations such as (19.22) and (19.23) with the shares at their initial values. Then we update the shares to values reflecting the situation after the 10% tax reduction. With these new shares in place, we compute the effects of a further 10% reduction. We proceed in this way until we have computed the effect of the full 100% tax reduction. As illustrated in Figure 19.3, this procedure captures the GDP triangle **amd** as an addition of rectangular strips.

Throughout the steps of a multistep calculation, the equality in the movements of real GDP calculated from (19.22) and (19.23) should be maintained. If even a small discrepancy (a difference of more than about 0.0001 percentage points) emerges, then we should look for an error.

However, given the interpretation of (19.22), its inclusion in a model alongside (19.23), provides more than just a device for checking computations. As discussed in Section 19.4, (19.22) is a valuable tool for setting up back-of-the-envelope expectations about what a CGE model should imply about the effects of: changes in employment of labor and capital, changes in technologies, and changes in taxes. In applying versions of (19.22) in this way, the question arises as to what values should be used for the shares: should these reflect

[14] Area **ab′c′d** is also the welfare gain calculated as: the increase in consumer surplus (**qab′r**) plus the increase in producer surplus (**sc′dt**) minus the loss in tax revenue (**qaa′r** plus **sd′dt** minus **a′b′c′d′**).

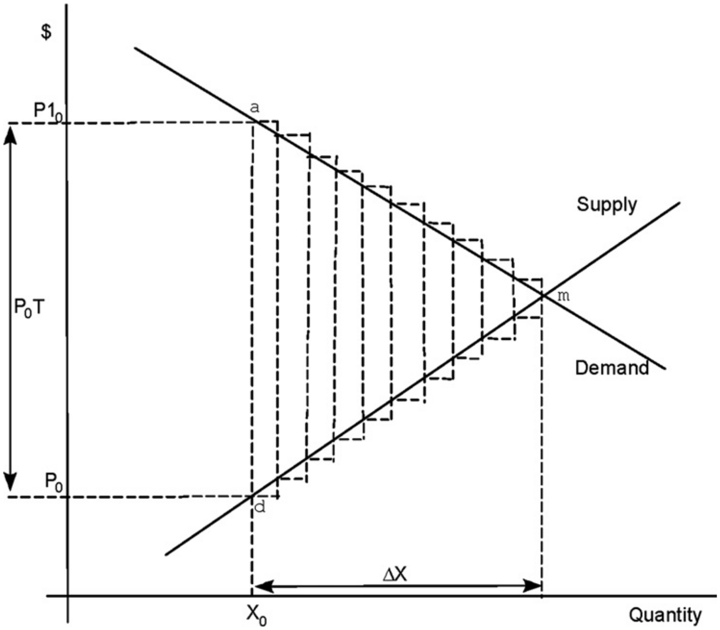

Figure 19.3 Effects of a removal of a sales tax. The increase in real GDP from the removal of the tax is given by the area of the triangle **amd**, which is $0.5 * P_0T * \Delta X$. In a multistep computation area **amd** is approximated by the sum of the dotted rectangular strips.

the initial situation or the final situation, or somewhere between? In the multistep procedure used to generate accurate solutions in Johansen-style CGE models, the transition in share values from their initial to their final values is normally quite smooth. Thus, in using equations such as (19.22) for back-of-the-envelope calculations, adequate explanations of CGE results can usually be obtained with the shares set at half way between their initial and final values. For example, in working out what to expect as the result for a simulation of the effect on GDP of the removal of a tax under the assumptions of no changes in primary factor inputs or technology, we can use a half-way version of (19.27):

$$\text{realgdp} = 0.5 * \left(S_{\text{tax}}^{\text{initial}} + S_{\text{tax}}^{\text{final}}\right) * x = 0.5 * S_{\text{tax}}^{\text{initial}} * x. \tag{19.28}$$

The percentage change in real GDP calculated in (19.28) will be recognized as the welfare effect of removing a tax in a partial equilibrium demand and supply framework (triangle **amd** in Figure 19.3) expressed as a percentage of GDP. This reminds us that the apparatus of diagrammatic demand and supply analysis is often a powerful starting point for explaining and understanding CGE results. However, as can be seen from the tax term on the right-hand side of (19.22), the general equilibrium approach offers much more than can be gleaned from demand and supply diagrams. For example, assume that every flow is taxed at the same rate. Then, the tax term in (19.22) tells us that if a small tax cut on flow (k,j) simply shifts resources between activities, then there will be no welfare effect. This is a result that is not easily explained in a partial equilibrium framework: it depends on the idea that in an economy with many taxed activities, basic prices depicted on the partial equilibrium supply curve understate the opportunity cost of resources required to generate an increase in output.

19.4 VALIDATION THROUGH BACK-OF-THE-ENVELOPE (BOTE) ANALYSIS AND OTHER PLAUSIBILITY CHECKS

In this section we consider two related questions: how can we decide that a set of CGE results is valid in the sense of being based on theory and data that are sufficiently comprehensive and realistic to capture the essence of the issue being analyzed; and how can we explain CGE results in a way that isolates and highlights the key mechanisms and data items on which they rest?

We give four answers: (i) Tell a qualitative story, (ii) tell a quantitative story based on the GDP identity, (iii) build a BOTE model that elucidates the relevant aspects of the full-scale model and (iv) use regression analysis to test ideas on what is going on in the CGE model. In any situation, it may be appropriate to pursue several of these approaches.

19.4.1 Qualitative validation: Telling a story

A form of validation for any set of CGE results is a story that explains the results in broad qualitative terms via mechanisms that are present in the model. Without such a story it is difficult to know whether a set of results rests on realistic assumptions about how the relevant part of the economy behaves.

Australian CGE modelers have been adept at finding CGE stories since the 1970s when they started presenting results in policy forums. Often in these situations they were faced with skeptical policy advisers who, while trained in economics, had little interest in or toleration of economic modeling. The CGE modelers used stories to persuade the policy makers that there might be some truth in what may have seemed counterintuitive results. For example, Johnson (1985) conducted a CGE simulation to assist government officials in their negotiations with Australian forestry workers who were campaigning against proposed restrictions on clear felling in old-growth native forests. The workers feared that the restrictions would destroy forestry jobs. Johnson's simulation demonstrated that the proposed regulations would in fact increase forestry employment. His non-technical but effective explanation of this result had three elements. (i) Restrictions on clear felling would increase labor required per log taken from existing forests. (ii) Restrictions would increase investment in plantation forestry: an activity that provided jobs for forestry workers. (iii) There would be a reduction in output of forestry products, reflecting higher prices, but the reduction would be small because the elasticity of demand for forestry products is low. Johnson found that the first two positive effects on forestry employment would easily outweigh the third negative effect.

Another example of a counterintuitive result that was explained convincingly by a qualitative story was that of Adams and Parmenter (1993, 1995). They were commissioned by the Bureau of Tourism Research to simulate the effects on the state economies in Australia of a general increase in inbound tourism. Their CGE simulation suggested that Queensland (Australia's main destination for foreign tourists) would be a small loser. In explaining this result, they pointed at data showing that international tourists spend a lot of time in Queensland but not a commensurate amount of money. Tourists do their shopping and pay for their within-Australia travel in Sydney (in New South Wales). So the upside for Queensland from a general stimulation of inbound tourism is not as strong as might be imagined. At the same time, Adams and Parmenter identified a significant downside. Queensland is a major exporting state for mining and agriculture. These industries would be hurt by tourism-induced real appreciation.

19.4.2 Quantitative validation: BOTE calculations based on the GDP identity

Being able to tell a convincing *qualitative* story is only a necessary condition for confidence about the computational and theoretical validity of a set of CGE results. We need a *quantitative* story to understand why a model produces particular numerical results. A simple quantitative story is often provided by a BOTE calculation based on the GDP identity, particularly Equation (19.22), and related demand and supply diagrams. Via this approach we can work out, for example, why a CGE model of the US implied that a unilaterally imposed \$20 per ton tax on CO_2 emissions would reduce US GDP by \$10 billion, and why a CGE model for Australia implied that a cut in motor vehicle tariffs from 10 to 5% would increase Australia's GDP by 0.06%. Why not \$100 billion for the carbon tax and 0.006% for the tariff cut?

19.4.2.1 Explaining the CO_2 result

In the case of the tax on CO_2 emissions, the \$10 billion result was obtained by our colleague Ashley Winston in a CGE simulation undertaken for the White House at the end of 2009 in the run up to the UN Climate Change Conference held in Copenhagen. Consistent with the sorts of figures being discussed in Washington at the time, Winston found that a \$20 per ton tax on CO_2 emissions would, in the long run, reduce total US emissions from 6 billion to 5 billion tons. In terms of (19.22), we can think of the right to emit CO_2 as being an input to production activities with a basic price of zero and a purchasers price accounted for entirely by the tax of \$20 per ton. Then a simple diagram (Figure 19.4) showing a demand curve for the right to emit CO_2 against the price of such emissions explains Winston's result.

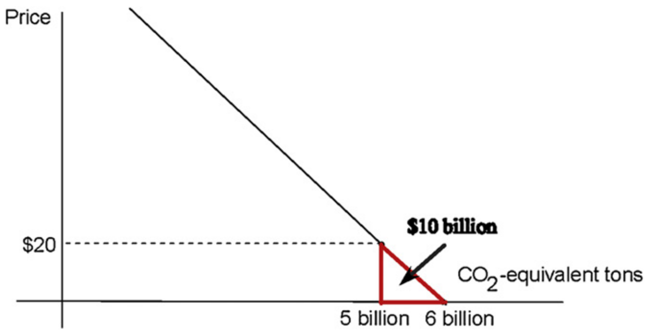

Figure 19.4 Demand for the right to emit CO_2.

19.4.2.2 Explaining the tariff-cut result

In their automotive industry report the Australian government's Productivity Commission (2008) announced as a headline result that a cut in tariffs on motor vehicle imports from 10 to 5% would increase Australia's annual GDP in the long run by 0.06% with a similar percentage increase in economic welfare. They based these claims on a CGE simulation.

In assessing the real GDP result, we start by noting that Australia's imports of motor vehicles in 2008 were about 2% of GDP, and the elasticity of demand for these imports assumed by the Productivity Commission was about −0.6. In the absence of changes in technology or in aggregate inputs of labor and capital, (19.22) suggests that the percentage effect on GDP of the tariff cut should be approximately:

$$\begin{aligned} \text{realgdp} &= 0.5 * \left[S_{\text{tax}}^{\text{initial}} \begin{pmatrix} \text{imported,} & \text{all} \\ \text{vehicles} & \text{users} \end{pmatrix} + S_{\text{tax}}^{\text{final}} \begin{pmatrix} \text{imported,} & \text{all} \\ \text{vehicles} & \text{users} \end{pmatrix} \right] * \eta * p1 \begin{pmatrix} \text{imported} \\ \text{vehicles} \end{pmatrix} \\ &= 0.5 * [0.002 + 0.001] * (-0.6) * (-4.5) = 0.004. \end{aligned} \tag{19.29}$$

In (19.29), $S_{\text{tax}}^{\text{initial}} \begin{pmatrix} \text{imported,} & \text{all} \\ \text{vehicles} & \text{users} \end{pmatrix}$ and $S_{\text{tax}}^{\text{final}} \begin{pmatrix} \text{imported,} & \text{all} \\ \text{vehicles} & \text{users} \end{pmatrix}$ are tariff collections on motor vehicles as a share of GDP in the initial and final situations. With the tariff moving from 10 to 5% and the initial level of imports being 2% of GDP, we would expect these shares to be about 0.002 and 0.001. η is the elasticity of demand for imported vehicles (about −0.6). $p1\begin{pmatrix} \text{imported} \\ \text{vehicles} \end{pmatrix}$ is the average percentage effect of the tariff cut on purchasers prices of imported vehicles. An upper bound on the expected reduction in purchasers prices caused by the cut in tariffs is 4.5% [= 100 ∗ (1.05 − 1.1)/1.1]. This would be the result if there were no margins and other motor vehicle taxes. Using this upper bound in (19.29) generates a BOTE real GDP effect of 0.004%.

This is only 1/15th of the Productivity Commission's estimate (0.004 compared with 0.06). Consequently, we must look for some other source of real GDP gain in the Commission's simulation. We were able to confirm that the Commission assumed no changes in technology: all the *a*'s in (19.22) were set at zero. The Commission also assumed that the cut in tariffs would have no long-run effect on employment. On the other hand, they assumed that capital stock in each industry would adjust so that there would be no long-run effect on industrial rates of return. Under this assumption, shocks that are favorable to the use of capital generate an increase in the economy's aggregate capital/labor ratio. With labor fixed, aggregate capital stock is increased, generating an increase in real GDP via the first term on the right-hand side of (19.22).

The Commission did not report the result for the economy-wide capital stock but after some detective work we found that their simulation implied an increase of about 0.14%. With the capital share in GDP being about 40%, the increase in capital stock accounts for an increase in real GDP of about 0.056% ($= 0.4 * 0.14$). Together, 0.056% from the primary factor term in (19.22) and 0.004% from the tax term explain the Commission's result of 0.06% for the increase in real GDP.

There are two reasons that locating the source of the Commission's GDP gain is important. First, we now know that in assessing their GDP result, we should think about whether it is reasonable to suppose that a cut in motor vehicle tariffs from 10% to 5% would cause a long-run increase in the economy's capital stock of 0.14%. Dixon (2009) considers this issue and agrees qualitatively with the Commission that an increase in K/L is realistic, mainly because a cut in motor vehicle tariffs would reduce the price deflator for investment (a vehicle intensive component of expenditure) relative to the price deflator for GDP.[15] He queries the size of the increase arguing that it is too large because the Commission assumed too much flexibility in industry capital/labor ratios by adopting values for K/L substitution elasticities that were too large. However, this is a relatively minor point.

The second reason that the source of the GDP gain is important relates to welfare. The Commission assumed that most of the extra GDP counts as an increase in the economic welfare of Australia. However, the increase in GDP is generated almost entirely by an increase in capital. An increase in capital must be financed either by foreign or domestic savings. In either case the extra income generated by the extra capital should not be considered a welfare gain: if finance is via foreign savings then the extra income belongs predominantly to foreigners; if it is by domestic savings then consumption must have been foregone to provide these savings. Dixon (2009) estimated that even if we

15 On the assumption that the rental rate on capital is the value of its marginal product, and that rates of return are approximated by the ratio of rental value to asset value, we can write $ROR = (P_{gdp}/P_i) * MPL(K/L)$. If the rate of return (ROR) is held constant in the long run, and the price deflator for GDP (P_{gdp}) increases relative to the price deflator for investment (P_i), then the marginal product of capital (MPK) must fall implying an increase in K/L.

accept the Commission's estimate of a long-run GDP gain of 0.06%, then the welfare gain would be no more than 0.01%.

19.4.3 Validation through BOTE models

A BOTE model is a small model designed to explain a particular application of a full-scale CGE model. The ideal BOTE model is simple enough to be managed with pen and paper but detailed enough to reproduce key aspects of the results from the full model.

BOTE modeling is as old as CGE modeling. In his seminal work, Johansen (1960) used a one-sector BOTE model to guide his discussion of the huge number of results generated by his CGE model of Norway. The BOTE model told him what to look for and what to expect in his full-scale model for the effects of exogenous changes in aggregate capital, aggregate employment, the number of households, technologies, exogenous demands and the price of non-competitive imports.[16] A similar approach was adopted by Taylor *et al.* (1980) to help explain the results from their model of Brazil.

BOTE modeling has a long history in Australia starting with efforts to explain results from the ORANI model. This was a 100-industry model designed for applied work, particularly on trade policy. In the initial ORANI publication, BOTE models were described as having three roles:

> *First, there is a purely practical point. With a model as large as ORANI, the onus is on the model builders to provide convincing evidence that the computations have been performed correctly, i.e., that the results do in fact follow from the theoretical structure and database. Second [BOTE calculations are] the only way: to "understand" the model; to isolate those assumptions which "cause" particular results; and to assess the plausibility of particular results by seeing which real-world phenomena have been considered and which have been ignored. Third, … by modifying and extending [BOTE] calculations … the reader will be able to obtain a reasonably accurate idea of how some of the projections would respond to various changes in the underlying assumptions and data.*
>
> ***(Dixon* et al.*, 1977, pp. 194–195)***

The BOTE models constructed to explain ORANI results often had just one domestically produced good which was consumed domestically and exported, one imported good, one type of labor and one type of capital. However, the exact nature of the BOTE models varied from application to application. For an ORANI simulation of the effects of an increase in oil prices, the corresponding BOTE model included the price of oil and the share of oil in the economy's production costs (Dixon *et al.*, 1984). For an ORANI simulation of the effects of a tariff increase, the corresponding BOTE model included a tariff rate (Dixon *et al.*, 1977, pp. 214–222). As demonstrated in many papers,

[16] Johansen's use of BOTE modeling is discussed in Dixon and Rimmer (2010a, section 19.2.2).

BOTE calculations reproduced ORANI results and those from subsequent models with considerable precision.[17] This led some economists to wonder why we needed full-scale models. The answer is that in most cases it would not have been possible to think of the BOTE model without first having the results from the main model. In any case, critical ingredients of the BOTE models were supplied from the database of the main model.

An important idea arising from BOTE modeling is that CGE results are usually best explained in a top-down fashion: macro to micro rather than micro to macro. Consider, for example, the short-run effects on employment of an increase in tariffs under the assumptions of fixed real wages, slack labor markets and a fixed nominal exchange rate. Given the task of explaining why a CGE model such as ORANI shows a decrease in aggregate employment, beginner students in CGE modeling typically adopt a micro-to-macro strategy, starting with industry results. These show reductions in employment in export-related industries. Students argue correctly that higher tariffs raise costs by increasing intermediate input prices and nominal wage rates: prices of imported inputs rise directly, domestic input prices rise in response to reduced competitive pressure from imports and nominal wage rates rise to maintain real wage rates by offsetting tariff-induced increases in consumer prices. Students recognize that export industries face high demand elasticities and are hurt because they are poorly placed to pass cost increases into higher prices. However, what about employment in tariff-protected import-competing industries? For these industries higher tariffs generate increased employment by inducing favorable demand switches towards their products and away from imported substitutes. Thus the micro-to-macro approach leaves students with the correct but *inadequate* explanation that aggregate employment decreases because the losses in export-oriented industries outweigh the gains in import-competing industries.[18] Via a BOTE model we can develop an *adequate* explanation.

19.4.3.1 BOTE model for explaining the short-run aggregate employment effect in a CGE model of an increase in tariffs

To explain the ORANI result for the aggregate employment effect of an increase in tariffs, we build a BOTE model in which the economy produces and exports one good (grain) and imports another good (vehicles). Grain production is via a constant returns to scale production function of capital and labor inputs. The cost of a unit of consumption is a Cobb–Douglas function of the consumer prices of grain and vehicles. Finally, we assume that the cost per hour of employing labor equals the value to the employer of

17 Recent application of BOTE methodology can be found in Adams (2005) and Dixon *et al.* (2011).

18 In their applications of the ORANI model for Australia's Industry Commission, Dixon *et al.* (1977, p. 208) and Dixon *et al.* (1982, p. 291) rely on this inadequate explanation.

labor's marginal product. Under these assumptions, which are representative of those underlying the full ORANI model, we have:

$$W = P_g * \frac{1}{A} * F_l\left(\frac{K}{L}\right) \tag{19.30}$$

$$P_c = P_g^{\alpha_g} * (P_v * T_v)^{\alpha_v}, \tag{19.31}$$

leading to:

$$\frac{W}{P_c} = \left(\frac{P_g}{P_v}\right)^{\alpha_v} * \left(\frac{1}{T_v}\right)^{\alpha_v} * \frac{1}{A} * F_l\left(\frac{K}{L}\right). \tag{19.32}$$

where W is the nominal wage rate; P_g and P_v are the basic prices of grain and vehicles; T_v is the power of the tariff $(1 + \text{rate})$ on vehicles; P_c is the price of a unit of consumption; α_g and α_v are positive parameters summing to one; K and L are capital and labor inputs; A is the technology variable in the production function relating grain output to factor inputs, i.e. grain output $= (1/A) * F(K,L)$; and F_l is the derivative of F with respect to L and, on the assumption of constant returns to scale, is an increasing function of K/L.

For the tariff simulation, (19.32) immediately puts us on the right track for explaining the reduction in aggregate employment. Under fixed real wages, the tariff increase has no effect on the left-hand side of (19.32). On the right-hand side, there is a reduction in $1/T_v$. We can also anticipate that our full-scale CGE model will behave as if there is an increase in A, i.e. a technological deterioration: tariff increases cause inefficiencies in the allocation of capital and labor between import-competing and export industries that cannot be captured directly in a one-sector BOTE model.[19] Against these two effects, we can expect an increase in P_g/P_v. This is because tariff increases generate an improvement in the terms of trade by restricting the supply of exports (thereby increasing their foreign currency price) and restricting the demand for imports (thereby reducing their foreign currency price). If the reductions in $1/T_v$ and A outweigh the increase in P_g/P_v, then (19.32) implies that F_l must increase. In the short run, K is fixed and so L must fall. Provided that terms-of-trade effects are weak, reflecting the adoption of high export-demand and import-supply elasticities, we can understand why the full-scale model implies that an increase in tariffs reduces aggregate employment in the short run.

Our macro-based BOTE explanation can be tested in various ways. If the BOTE explanation is really capturing what is going on, then a repeat of the simulation in the

[19] The appropriate increase in A could be worked out in a BOTE calculation along the lines of (19.29). It is likely to be small because with realistic values for the initial tariff rates and policy-relevant values for the tariff changes, efficiency effects are small.

full-scale model with lower export-demand and import-supply elasticities would show a smaller reduction in employment.

As we will see shortly, BOTE models such as (19.30)−(19.32) can be extended to explain other results including those for investment, consumption and the balance of trade. BOTE models can also be made numeric by giving F_l, for example, a specific form, and plugging in stylized values from the full database for parameters and cost shares.

Once the behavior of macro variables is understood via a BOTE model, then it is usually comparatively straightforward to explain results for micro variables. For example, if we can understand why a simulation shows a reduction in aggregate investment, then we can understand why it shows a reduction in output and employment in the construction industry.

By giving us readily understood explanations of results, BOTE models help to pinpoint limitations and weaknesses of particular applications of CGE models. For example, in the ORANI tariff-increase simulation an underlying assumption is that there are no compensating changes in other tax rates. It may have been more realistic to assume that extra tariff revenue is used to reduce other taxes. If income taxes were reduced then it might be reasonable to assume that workers factor this into their wage demands and allow the pretax real wage rate to fall. As is apparent from (19.32), this might eliminate the simulated reduction in employment.

With BOTE models highlighting the potential importance of revenue-use assumptions, this issue has been handled more explicitly in subsequent simulations of tariff and tax changes in Australia. A prime example is CGE work carried out for the Australian parliament in 1999 on the effects of introducing of a goods and services tax (GST).

19.4.3.2 BOTE model for explaining the CGE results for the effects of a GST in Australia

In 1998 the Australian government proposed an overhaul of the taxation system involving three parts:

- **(i)** The introduction of a GST. This is a value-added tax and is equivalent to a broad-based consumption tax.
- **(ii)** The abolition of the wholesale sales tax (WST). This was a sales tax without any particular logic that was charged at high rates mainly on durables such as vehicles, refrigerators, washing machines and television sets.
- **(iii)** A switch away from direct taxation towards indirect taxation. The net effect of the GST/WST substitution was projected to be an annual increase in indirect taxes of A$6 billion, allowing a balanced budget reduction in direct taxes of A$6 billion. To make the tax reform more attractive to voters the government offered an additional unbalanced A$6 billion reduction in direct taxes, bring the total reduction in direct taxes to A$12 billion.

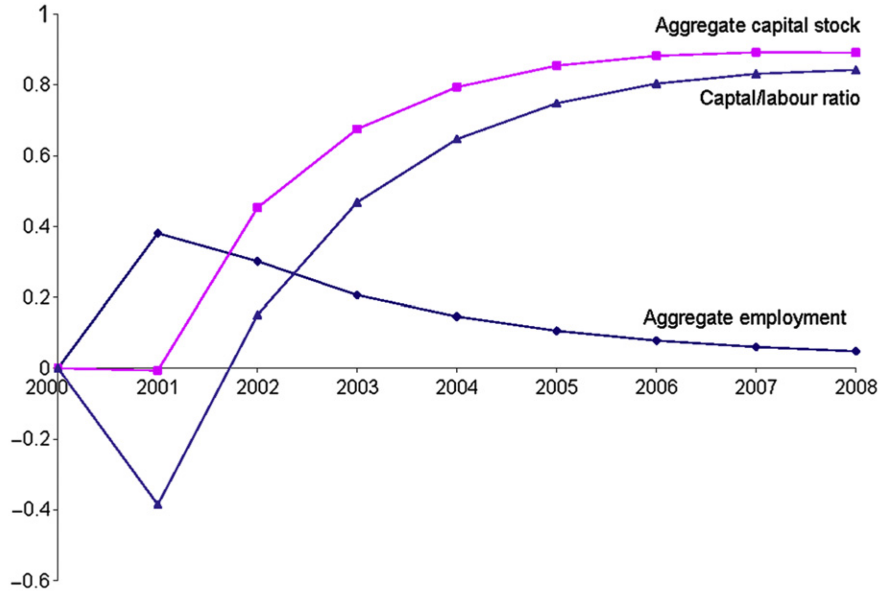

Figure 19.5 MONASH results: effects of introducing a GST under after-tax wage bargaining (% deviation from baseline).

Opposition parties in the Australian parliament forced the government to commission the Centre of Policy Studies (CoPS),[20] Econtech[21] and several other economic consultants to provide CGE analysis of the proposals. The politicians demanded results at a highly detailed level covering effects on industries, regions, occupations and income classes. However, in the end the debate was mainly about macroeconomic effects, particularly on employment. With some modifications, the proposals were passed by the parliament after a lengthy consideration of the modeling results.

Figures 19.5 and 19.6 show results for two simulations produced by the CoPS team using the MONASH model. The simulations differ with respect to assumptions about the reaction of the labor market. The results in Figure 19.5 were generated assuming workers accept cuts in income tax as compensation for GST-related increases in taxes on consumption. Under this assumption, workers focus on the real wage *after* income tax in the wage bargaining process. The results in Figure 19.6 were generated under the assumption that workers do not accept cuts in income tax as compensation for the GST. They continue their traditional behavior of bargaining over the real wage *before* income tax. As can be seen from the two charts, the results for the all-important employment

20 See Dixon and Rimmer (1999a,b,c).
21 See Murphy (1999).

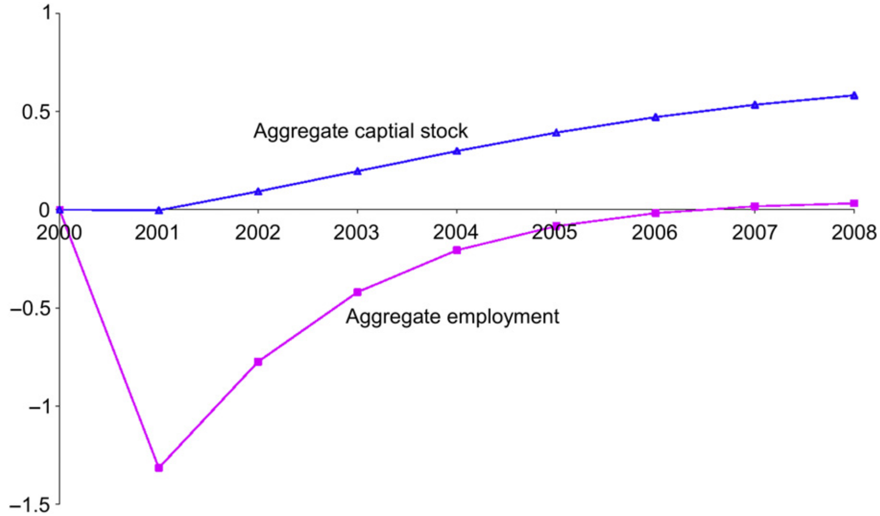

Figure 19.6 MONASH results: effects of introducing a GST under before-tax wage bargaining (% deviation from baseline).

variable are radically different: favorable in the after-tax case (Figure 19.5) and unfavorable in the before-tax case (Figure 19.6). The CoPS team understood these results via a BOTE model and then explained them at various levels of technicality to the politicians and their advisors.

The BOTE model that CoPS used was a version of the grain/vehicles model described in (19.30)–(19.32) extended to include five types of taxes and investment undertaken as a Cobb–Douglas combination of grain and vehicle purchases. The BOTE equations were as follows:

$$P_c = P_g^{\alpha_{gc}} * P_v^{\alpha_{vc}} * T_c \tag{19.33}$$

$$P_i = P_g^{\alpha_{gi}} * P_v^{\alpha_{vi}} * T_i \tag{19.34}$$

$$W * T_w = \left(P_g/T_g\right) * M_l\left(K/L\right) \tag{19.35}$$

$$Q * T_q = \left(P_g/T_g\right) * M_k\left(K/L\right) \tag{19.36}$$

$$W_{realA} = W/P_c \tag{19.37}$$

$$W_{realB} = W * T_w/P_c \tag{19.38}$$

$$\rho = Q/P_i, \tag{19.39}$$

where P_g and P_v are the basic price of grain and the c.i.f. price of vehicles; P_c and P_i are the purchasers prices of a unit of consumption and a unit of investment; T_c, T_i and T_g are the powers (one plus rates) of the taxes applying to consumption, investment and production; Q and W are after-tax rental and wage rates; T_w and T_q are the powers (one plus rates) of the income taxes applying to labor and capital income; M_l and M_k are the marginal products of labor and capital; W_{realA} and W_{realB} are the *after*-tax real wage rate and the *before*-tax real wage rate; ρ is the rate of return on capital calculated as the after-tax rental (i.e. after-tax user price of capital) divided by the cost (i.e. the asset price) of a unit of capital; and the αs are positive parameters reflecting the shares of grain and vehicles in consumption and investment, such that $\alpha_{gc} + \alpha_{vc} = 1$ and $\alpha_{gi} + \alpha_{vi} = 1$. From these equations we find that:

$$M_l\left(\frac{K}{L}\right) = W_{realA} * T_w * T_g * T_c * \left(\frac{P_v}{P_g}\right)^{\alpha_{vc}}, \tag{19.40}$$

or equivalently:

$$M_l\left(\frac{K}{L}\right) = W_{realB} * T_g * T_c * \left(\frac{P_v}{P_g}\right)^{\alpha_{vc}}, \tag{19.41}$$

and that:

$$M_k\left(\frac{K}{L}\right) = \rho * T_q * T_g * T_i * \left(\frac{P_v}{P_g}\right)^{\alpha_{vi}}. \tag{19.42}$$

In (19.40)–(19.42), we emphasize that marginal products are functions of K/L. M_l is an increasing function of K/L and M_k is a decreasing function of K/L.

While the government's proposal involved changes in hundreds of sales tax and income-tax rates that were taken into account in the MONASH simulations, in terms of the BOTE model these could be condensed into:

- A cut in T_g from 1.036 to 1.022 reflecting elimination of the WST on inputs to production.
- A cut in T_i from 1.037 to 1.013 reflecting elimination of the WST on inputs to capital creation.
- A cut in T_w from 1.250 to 1.215 reflecting reduction in income tax on labor.
- A cut in T_q from 1.250 to 1.233 reflecting reduction in income tax on capital.
- An increase in T_c from 1.070 to 1.105 reflecting the net effect on consumption taxes of the introduction of the GST and the elimination of the WST.

With the change in the tax mix, $T_w * T_g * T_c$ declines by 1.0% (from 1.250 * 1.036 * 1.070 to 1.215 * 1.022 * 1.105). Under the after-tax bargaining assumption, W_{realA} is sticky in the short run. In the absence of a change in the terms of trade (a movement in P_v/P_g),[22] the BOTE model, see (19.40), which indicates that the change in the tax mix will cause a short-run reduction in the marginal product of labor. Consequently K/L will decrease. As K moves slowly, there must be a short-run increase in L. This is confirmed in Figure 19.5 where we see that employment moves above the baseline.

Looking now at (19.42), we ask what is the short-run impact of the change in the tax mix on the rate of return (ρ)? With a decrease in K/L, M_k rises. $T_q * T_g * T_i$ falls by 4.9% (from 1.250 * 1.036 * 1.037 to 1.233 * 1.022 * 1.013). Again, ignoring changes in P_v/P_g, we see that ρ must rise. Thus the MONASH simulation showed a short-run increase in investment followed by an upward movement in capital which can be seen in Figure 19.5.

The labor market assumption underlying this simulation is that when the labor market is tight the bargaining position of workers improves, generating an increase in real after-tax wages relative to the baseline. Tightness in the labor market is indicated by employment being above baseline. Thus, the short-run increase in employment shown in Figure 19.5 led in the MONASH simulation to increased after-tax real wages (increases in W_{realA}). Consequently, after initially moving below its forecast path, M_l moved up towards this path [see (19.40)]. This meant that after its initial fall, K/L rose (see Figure 19.5). This forced M_k to fall back towards baseline generating a reduction in ρ, thereby slowing the rise in K. With growth in K being choked off, further rises in W_{realA} (necessitating increases in K/L) were achieved by reductions in L. As can be seen in Figure 19.5, L fell back towards the baseline.

In the very long run, the deviation in ρ is zero. In the absence of terms-of-trade effects, the reduction in $T_q * T_g * T_i$ leaves the K/L ratio permanently increased (Figure 19.5). Thus M_l is permanently raised. This together with the reduction in $T_w * T_g * T_c$ leaves W_{realA} permanently increased with L approximately on its baseline path.

Supporters of the government's proposal were happy with the results in Figure 19.5: a temporary increase in employment and a permanent increase in real after-tax wages. However, the CoPS team made two enthusiasm-damping points. First, they reminded the politicians that the proposal involved a net reduction in tax revenue of A$6 billion. This accounts for about half of the planned reduction in income taxes. Without the reduction in tax revenue, T_w would decline from 1.250 to 1.233 rather than to 1.215 and $T_w * T_g * T_c$ would increase by about 0.5%. This would generate a short-term decline in L. Thus the short-term employment gain associated with the package is entirely the result of fiscal stimulation that could be undertaken without disrupting the economy with a major change in the tax system.

22 Dixon and Rimmer (1999c) discusses the terms-of-trade effects of the GST in detail. However, these effects are not important for the main arguments.

The second point made by the CoPS team was that the favorable results illustrated in Figure 19.5 depended on the after-tax wage bargaining assumption. What seemed at the time to be an equally likely assumption was that wage bargaining would be conducted in before-tax terms. In terms of the BOTE model, the short-run employment outcome would then depend on the movement in $T_g * T_c$ [see (19.41)]. Implementation of the government's proposal would raise $T_g * T_c$ by about 2% (from 1.036 * 1.070 to 1.022 * 1.105). Thus, with sticky before-tax real wages (W_{realB}), the CoPS team warned that M_l would rise, necessitating a short-run fall in L as shown by the MONASH results in Figure 19.6. The long-run macro effects under either wage assumption are quite similar. This is because under both assumptions, employment in the long run is driven close to its baseline level.

19.4.4 Validation through regression analysis applied to results

BOTE models are an effective way of checking and explaining macroeconomic results from CGE models. As mentioned in Section 19.4.3, macro results are often the dominant determinants of results at the industry, regional and occupational levels. However, top-down explanations of these micro results can often be deepened by regression analysis. Under this approach, modelers check and strengthen their understanding of micro results by regressing them against what they think are the determining factors in the model.[23]

Consider, for example, the results in Figure 19.7 obtained by Dixon *et al.* (2007) from the USAGE model[24] for the effects on employment by state in the US of removal of major import restraints (tariffs and quotas).[25] The worst affected states are Idaho and North Carolina, which lose 0.498% and 0.477% of their jobs, while the most favored state, Washington, obtains a 0.214% increase in jobs. Idaho suffers from over-representation in its employment of highly protected Sugar crops, Sugar products and Dairy products, while North Carolina suffers from over-representation of highly protected Textiles. Washington benefits from over-representation of export-oriented commodities such as Aircraft and Aircraft equipment. Do state employment shares and percentage changes in commodity outputs at the national level explain all of the regional employment results? To answer this question Dixon *et al.* regressed state employment results from USAGE against a national index worked out for region r as:

$$NatIndex\left(r\right) = \sum_j \mathrm{Sh}(j,r) \times Nat\_emp(j), \tag{19.43}$$

[23] Perhaps the first application of this method is in Dixon *et al.* (1977, pp. 223–259). See also Dixon *et al.* (1982, pp. 313–319).

[24] USAGE is a large-scale MONASH-style model of the US, see Dixon *et al.* in Chapter 2 of this Handbook.

[25] An earlier version of these results was used by the US International Trade Commission (2004, pp. 125–126).

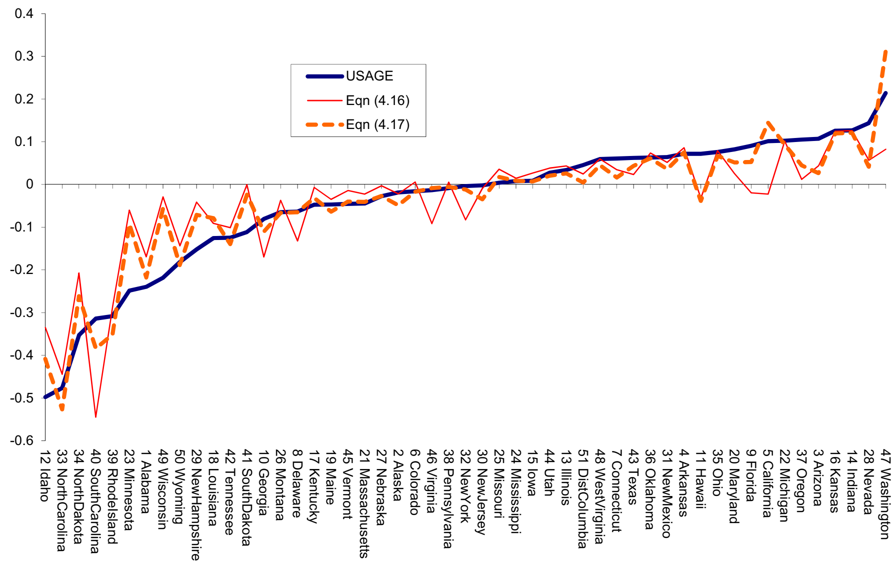

Figure 19.7 State employment effects (%) of removing major US tariffs and quotas explained by regression equations.

where Sh(*j*,*r*) is the share of employment in region *r* accounted for by production of good *j* and *Nat_emp*(*j*) is the percentage change in employment at the national level in the production of commodity *j*.

The outcome of the regression was:

$$Reg\_emp(r) = -0.023 + 2.755 * NatIndex(r), \quad R^2 = 0.73, \tag{19.44}$$

where *Reg_emp*(*r*) is the USAGE percentage change result for employment in state *r*.

On studying Figure 19.7, Dixon *et al.* (2007) saw that regression Equation (19.44) strongly underpredicts the USAGE employment results for Washington, California and South Carolina. A factor that these three states have in common is major ports. In the USAGE simulation, a state benefits from having a major port via the trade-expanding effects of the removal of import restraints. On this basis Dixon *et al.* decided to add a port index to their regression explanation of the USAGE results. The index chosen was a ratio of two shares: the state's share of US trade going through its ports and the state's share of national employment. With the port index included, the regression equation became:

$$Reg\_emp(r) = -0.050 + 3.164 * NatIndex(r) + 0.056 * PortIndex(r), \quad R^2 = 0.88. \tag{19.45}$$

The port index enters the regression with the expected sign and raises R^2 to 0.88. Nevertheless, Figure 19.7 still shows some large gaps between the USAGE results and those explained by the regression.

By investigating these gaps, Dixon *et al.* uncovered further mechanisms in USAGE that were relevant for their simulation. They noted that regression equation (19.45) strongly under predicts the USAGE employment results for Hawaii, Nevada and Arizona. By adding a state tourism-related variable to the regression equation they demonstrated that underprediction for these states is related to tourism effects specified in USAGE but not captured by either NatIndex or PortIndex.

19.5 CONSISTENCY WITH HISTORY

For researchers with a background in statistics a natural form of validation for a model is a demonstration that it fits history. In CGE modeling, efforts to achieve historical consistency can take various forms including: time-series estimation of parameters, tests of goodness-of-fit and historical simulation.

19.5.1 Time-series estimation of parameters

For Jorgenson's models (see particularly Jorgenson, 1984) consistency with the past means that parameters (e.g. substitution elasticities) are estimated by econometric techniques applied mainly with time-series data. For the early MONASH models, particularly ORANI, there were also considerable efforts to use time-series data to estimate parameters including: import—domestic substitution elasticities (Armington elasticities), product—product transformation elasticities in agriculture, capital—labor substitution elasticities, export-demand elasticities, and consumer income and price elasticities. This work is summarized in Dixon *et al.* (1982, pp. 181—189). Implicitly it was thought that with statistically justified parameter values we would have models that would perform well in forecasting and policy analysis.

Valuable time-series econometric research continues in the CGE field.[26] However, it is fair to conclude that this type of research has not delivered as much to CGE modeling as was initially hoped and anticipated. Parameters estimated by time-series econometrics have often proved unrealistic in a simulation context. For example, econometricians have estimated that export-demand elasticities are quite low, even less than one.[27] But in a simulation context, low values can lead to implausible results suggesting that cost increases in export-oriented industries can improve welfare by generating increases in export revenue despite reductions in export quantities. Another example is econometric estimates of substitution elasticities between capital and labor. These are often close to one. However, in simulation, values as high as this lead to unrealistic sensitivity of employment with respect to changes in real wages.[28] Implications such as these have led

26 See, e.g., Hertel *et al.* (2007), and Hillberrry and Hummels in Chapter 18 of this Handbook.
27 Export-demand elasticities are discussed in Dixon and Rimmer (2010b, section VI).
28 See Dixon (2009).

some CGE modelers to abandon econometric estimates of parameters, even when such estimates are available. In the majority of influential CGE analyses, settings of key parameters reflect judgments sometimes supplemented by sensitivity analyses.

19.5.2 Tests of goodness-of-fit

The approach used by early CGE modelers to fit history was to perform goodness-of-fit tests, i.e. tests of the ability of a model to project endogenous variables given the correct values for the exogenous variables. For example, Johansen (1960) compared observed average annual growth rates around 1950 in endogenous variables such as industry outputs, employment and capital inputs with predictions from his model computed on the basis of observed or estimated values for average annual growth rates in six sets of exogenous variables: aggregate employment; aggregate capital stock; the number of households; Hicks-neutral technical change by industry; the price of non-competing imports; and exogenous demand, including public consumption, aggregate investment and net exports (exports minus imports) by commodity. He used this comparison to pinpoint weaknesses in his model and to organize a discussion of real-world developments. For agriculture he found that the computed growth rate in output closely matched reality, but that the computed growth rate in employment was too high while that in capital was too low. This led to a discussion of reasons, not accounted for in the model, for exodus of rural workers to the cities. For forestry, the computed growth rates for output and primary-factor inputs were too high. He thought that the income elasticity of demand for forestry products may have been set too high and also that there may have been a taste change, not included in his model, against the use of forestry products as fuel. By going through his results in this way, Johansen developed an agenda for model improvement.

Other early tests of goodness-of-fit for CGE models were conducted by Taylor *et al.* (1980) on a 25-sector Brazilian model for the period 1959–1971 and Dixon *et al.* (1978) on a 109-sector model of Australia for the period 1963–1972. Perhaps the most detailed of these early tests was carried out by Cook (1980) who measured the accuracy of year-on-year projections from a 22-sector ORANI model of Norway for 1949–1961.[29]

Since the 1980s relatively little work has been undertaken on checking the validity of CGE projections against statistically observed outcomes. Perhaps this reflects the focus of most modern CGE studies on the effects of a particular shock (e.g. the effects of a change in a tax or tariff). At any time the economy is affected by a myriad of exogenous shocks.

[29] Adelman and Robinson (1978, pp. 62–76) fed base-period values for the exogenous variables into their Korean model and showed that it then closely reproduced base-period values for the endogenous variables. They describe this as "a strong test of both model structure and model behavior." However, it seems to us that the test they performed is merely computational (see Section 19.2). Passing it is evidence that they successfully calibrated their model but says nothing about the predictive or explanatory power of the model or its consistency with history apart from the initial data year.

Isolating the actual impact of any particular shock independently of a model is problematic. Thus, it is usually difficult to check a CGE projection of the effects of a particular shock against an independently observed outcome.[30]

19.5.3 Historical simulation

Our own approach to fitting history is the technique of historical simulation.[31] Using this technique we force a model to track observed movements in outputs, inputs and final demands. In this process the model reveals changes in consumer preferences, technologies, world trading conditions and other naturally exogenous but unobservable variables. For example, given data on consumption of tobacco products (the variable on the left-hand side of a demand equation) and data on prices, incomes and number of households (variables on the right-hand side of a demand equation), an historical simulation will fit the demand equation for tobacco products by introducing an endogenous shift variable representing household preferences (part of the household utility function). Fitting history in this way is not a goodness-of-fit test: we fit the data perfectly. However, assessing the plausibility of generated changes in consumer preferences, technologies, etc., against other information becomes a form of validation.

An example of this type of validation is our historical simulation with USAGE for the period 1992–1998 (Dixon and Rimmer, 2004). Plausibly, this quantified increasing interest by US households in health and lifestyle issues: with preference shifts towards Boats, Luggage, Travel trailers, Sporting clubs and Cable TV, and against Cigarettes, Malt beverages, Wine and spirits, and Distilled liquors. Fashion changes were also evident: with preference shifts away from Bowling centers and Newspapers. Among the technology changes quantified by the historical simulation were shifts by industries in favor of the use of Computer equipment, Computer services, Personnel supply services, Job training and Management services. In capital creation, there were shifts against Glass, Sawmill products, and Brick and clay tiles, and towards Ready-mix concrete. It appears that between 1992 and 1998, US building methods became more basic and less artistic.

Less plausibly, our initial historical simulations showed preference shifts against nearly all food products. This alerted us to the possibility that the expenditure elasticities of demand for food products (deduced from an econometric study) were too high, so that income growth was generating too much growth in modeled food consumption, thus requiring negative preference changes to allow USAGE to reproduce observed growth. When lower elasticities were adopted the problem of preference shifts against food

30 Attempts by Kehoe (2005) and other to validate CGE projections of the effects of particular shocks are discussed in Section 19.7.

31 This is described in Dixon *et al.* in Chapter 2, Section 2.5, of this Handbook. The initial application of historical simulation was Dixon and McDonald (1993).

disappeared. In this way, results from historical simulations are used to refine parameter estimates.

19.6 FORECASTING PERFORMANCE

In recent years, CGE modelers have become increasingly interested in baseline forecasting.[32] There are two obvious reasons for this. First, it is a response to demand by private and public sector consumers of CGE analyses who want projections not only of the effects of particular shocks, but also of the evolution of the economy without the shock. Second, answers to policy questions can be improved by generating them as deviations around a realistic forecast. For example, in Australia's GST debate (see Section 19.4.3.2) CGE models were in agreement that the government's proposed tax reforms would harm industries with a heavy reliance on inbound tourism.[33] However, the models disagreed on the implications of this for the rest of the economy. The MONASH model, which incorporated a baseline forecast implying a strongly increasing share for inbound tourism in Australia's GDP, gave a more pessimistic picture than the Murphy model which produced comparative static results on a mid-1990s database (see Dixon and Rimmer, 1999c).

For this chapter we are interested in baseline forecasting for a third reason: it opens up a possibility for model validation. Baseline forecasts that account for the myriad of shocks to which the economy is subjected at any time, can be checked against outcomes. Reasons for discrepancies can be investigated and avenues for model improvements can be found.[34] The rest of this section is a description of a baseline validation test that we carried out for USAGE – a MONASH-style model of the US.[35]

19.6.1 Checking USAGE 1998–2005 forecasts for 500 industry outputs against reality

Creation of a baseline for a MONASH-style model uses two simulations: historical and forecast. The relationship between these simulations is illustrated in Figure 19.8. The current year is denoted by t_0, the historical period is $t_0 - \tau_1$ to t_0 and the forecast period is t_0 to $t_0 + \tau_2$. As can be seen from Figure 19.8, baseline forecasts developed according to the MONASH methodology incorporate available expert forecasts and build in trends in

32 See Dixon *et al.* in Chapter 2, Section 2.5, of this Handbook.

33 Unlike other exports, under the government's proposal tourism exports were not GST exempt.

34 A fourth reason for being interested in baseline forecasts is that they are necessary in studies of adjustment costs. For example, in considering the adjustment costs following a tariff cut, it makes all the difference whether the reductions in employment in adversely affected industries are accommodated by an increase in the rate of firing or a reduction in the rate of hiring, but to know which applies we cannot avoid forecasting, see Dixon and Rimmer (2002, pp. 289–299) and Productivity Commission (2000).

35 Section 19.6.1 up to the end of Section 19.6.1.4 is an abridged version of Dixon and Rimmer (2010c).

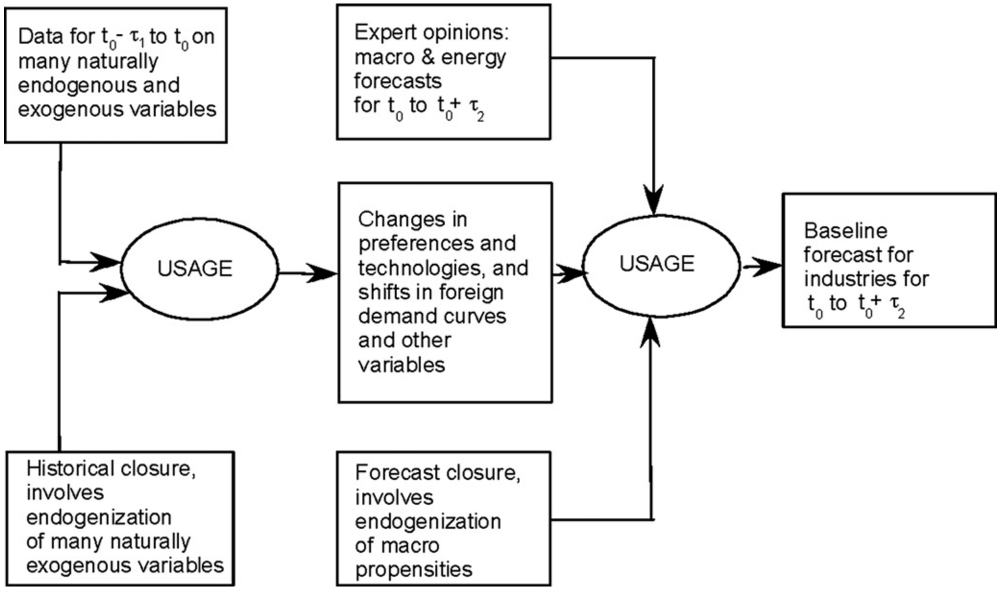

Figure 19.8 Relationship between historical and forecast simulations.

preferences, technologies and world trading conditions derived from the historical simulation.

In the case of USAGE, expert opinions used in baselines include forecasts prepared by the Congressional Budget Office, the US Department of Agriculture, the Bureau of Labor Statistics and the Energy Information Administration. These cover the usual array of macro variables (C, I, G, X, M, GDP and employment) and energy variables, such as the outputs and prices of electricity, crude oil and natural gas. The most convenient source for both macro and energy forecasts is the Energy Information Administration's annual publication entitled *Annual Energy Outlook*. With regard to trends, USAGE historical simulations provide these at the 500-sector level: for consumer preferences measured by coefficients in the underlying household utility function; For several aspects of industry technologies including intermediate-input-saving, primary-factor-saving, labor-capital bias and import-domestic bias; and for trading conditions measured by the positions of foreign demand curves for US exports and foreign supply curves for US imports.

For assessing the validity of our baseline forecasting method, we generated USAGE forecasts for 1998–2005 for 500 commodity outputs using an historical simulation for 1992 to 1998 and expert opinions available in 1998. In terms of Figure 19.8, $t_0 = 1998$, $\tau_1 = 6$ and $\tau_2 = 7$. We then compared these genuine output forecasts for 1998–2005 with actual outcomes. We also conducted an historical simulation for 1998–2005. This revealed the "true" movements in preferences, technologies and world trading conditions for the period. Having made the genuine output forecast for 1998–2005, we then performed a series of forecast simulations in which we successively introduced the "truth" for the movements in different groups of exogenous variables. The aim was to assess the importance of different exogenous factors in determining the accuracy of forecasts for outputs by commodity.

19.6.1.1 Measuring forecast performance

The first measure of USAGE's forecast performance that came to mind is average error (AE), defined as:

$$AE = \left(\frac{1}{N}\right) * \sum_c \left|f_c - a_c\right| / \left(1 + \frac{a_c}{100}\right), \tag{19.46}$$

where f_c is the forecast of the percentage change in the output of commodity c between 1998 and 2005, a_c is the actual percentage change in the output of commodity c, and N is the number of commodities (503 in the present application of USAGE). AE is an average across the 503 USAGE commodities in percentage gaps between forecast levels of commodity outputs and actual levels in 2005.[36]

If AE had turned out to be close to zero, then there would have been no difficulty in declaring the forecast to be a success. However, as we will see in Section 19.6.1.2, the AE value that we obtained seems high at 19%. Rather than being disappointed, we decided to look at what can be done without a model. The most obvious non-model approach is historical trends. On this logic, a performance measure for the USAGE forecasts that builds in a fair comparison is:

$$M = \frac{\sum_c \left|f_c - a_c\right| / \left(1 + \frac{a_c}{100}\right)}{\sum_c \left|h_c - a_c\right| / \left(1 + \frac{a_c}{100}\right)}, \tag{19.47}$$

where h_c is the percentage change in the output of commodity c across the historical period, 1992–1998, extrapolated to make it apply for a seven-year period rather than a six-year period. M is the ratio of the average error in the USAGE forecast to the average error in a forecast based on extrapolation. If $M = 1$, then the USAGE-based forecast is no more or less accurate than a non-model-based forecast generated by trends. If, on the other hand, $M = 0.7$, we can say that by using USAGE we have eliminated 30% of the error involved in simply relying on historical trends.

[36] In (19.46) the forecasting error for each commodity receives equal weight. This is appropriate if we do not wish to emphasize forecasting performance for any particular group of commodities. Dixon and Rimmer (2010c) use weighted versions of (19.46) with weights reflecting commodity shares in US imports, exports and total trade. These versions are appropriate when there is a special interest in the performance of USAGE in forecasting trade-exposed industries.

19.6.1.2 Overall performance of the genuine forecast

Table 19.3 presents results from applying (19.46) and (19.47). In this section we consider the results in the first row, those for the genuine forecasts. Section 19.6.1.3 considers the results in the other rows.

What conclusions should we draw for the genuine forecasts? At first glance the *AE* values seem large. GDP growth, an indicator of average output growth across commodities, was 21.6% between 1998 and 2005. Yet the forecast error for the output of a typical commodity is 19%. This would be a disastrously large average error if all commodities had actual growth rates in a tight band around 21.6. However, they did not. The actual growth rates were spread over the range −66 (for Slippers) to 218 (for Computer peripheral equipment). Only 151 out of the 503 USAGE commodities exhibited output growth within 10 percentage points of the average.

The *M* coefficient gives a more optimistic view of the USAGE forecasts than obtained from *AE*. It indicates that USAGE reduces the forecast error by 42% ($M = 0.58$) relative to a simple non-modeling extrapolation approach. The difference between the *AE* and *M* perspectives for measuring forecast performance is illustrated in Figure 19.9, which shows for each of the 503 USAGE commodities the trend forecast error, $|h_c - a_c|/(1 + a_c/100)$, on the horizontal axis and the USAGE forecast error, $|f_c - a_c|/(1 + a_c/100)$, on the vertical axis. With the *AE* perspective we consider only the values on the vertical axis: *AE* is a simple average of these values. With the *M* perspective, we consider the values on both axes. As the bulk of the points in Figure 19.9 are below the 45° line, the *M* perspective tells us that the USAGE forecasts outperform simple trend forecasts.

19.6.1.3 Improving performance by getting the exogenous movements right

Rows (2)−(5) of Table 19.3 show how the errors in the forecasts are reduced as we successively introduce the true 1998−2005 movements in the exogenous variables. In row (2), the true values for the 1998−2005 movements in macro and energy variables replace the forecast values. In row (3), the true values for the trade variables (movements in import/domestic preferences, import prices, export-demand curves and tariff equivalents of import restraints) replace the forecast values, and in row (4) we introduce

Table 19.3 *M* and *AE* coefficients for USAGE forecasts of growth in commodity outputs between 1998 and 2005

	Simulation	*AE*	*M*
(1)	Genuine forecast	19	0.58
(2)	(1) + Macro and energy	18	0.54
(3)	(2) + Trade variables	12	0.36
(4)	(3) + Technology and preferences	5	0.14
(5)	Truth for all exogenous	0	0.0

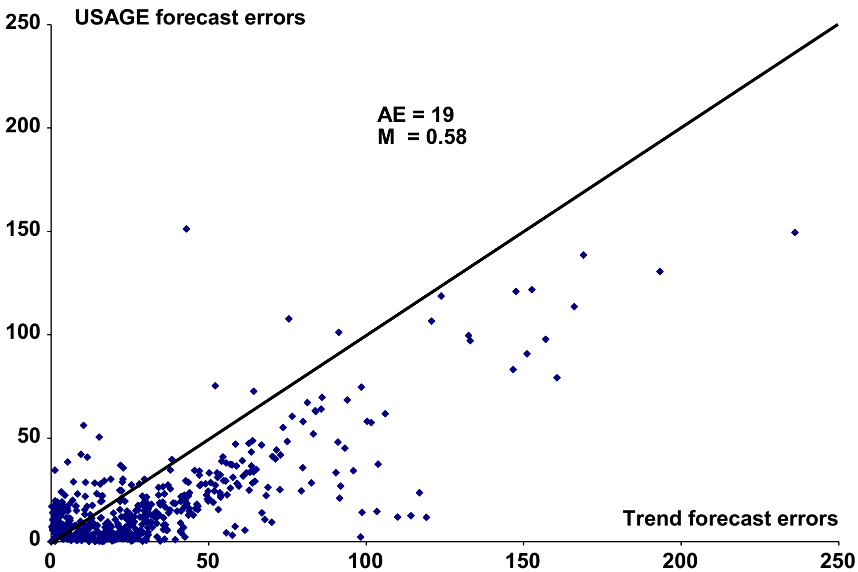

Figure 19.9 Percentage forecast errors for commodity outputs, 1998–2005: Extrapolated 1992–1998 trend versus genuine USAGE forecast.

the true values for the movements in technology and preference variables. Row (5) introduces the true values for movements in all remaining exogenous variables (e.g. required rates of return on capital and the commodity composition of public expenditure).[37]

Comparison of rows (1) and (2) in Table 19.3 shows that the availability in 1998 of accurate macro and energy forecasts for the period 1998–2005 would have generated a useful improvement in our ability to forecast commodity outputs. The *M* coefficient is reduced from 0.58 to 0.54 while the *AE* coefficient falls 19 to 18. As discussed in Dixon and Rimmer (2010c), the introduction of accurate macro and energy forecasts had their greatest impact on the USAGE forecasting performance for export-oriented commodities. USAGE output forecasts for export commodities were particularly poorly informed by the macro forecasts available in 1998. These showed aggregate exports increasing by 57.3% between 1998 and 2005, whereas the truth was an increase of only 23.8%. Telling USAGE about the 23.8% significantly improved the model's ability to forecast output growth for heavily exported commodities.

A major improvement in the USAGE forecasting performance comes when the model is given the truth about exogenous trade variables. In Table 19.3, *M* and *AE* fall

[37] The true values for industry/commodity variables such as technologies and positions of export-demand curves that are introduced as we move down Table 19.3 are relative movements rather than absolute movement. A degree of freedom is needed to hit each macro target.

from 0.54 to 0.36 and from 18 to 12. The importance of trade variables is perhaps surprising because for most commodities import shares in the US market and export shares in output are quite small: 64% of USAGE commodities in 1998 had import shares of less than 15 and 73% had export shares of less than 15%.

The reason that accurate projections of trade variables would make a major contribution to the accuracy of the commodity output forecasts is that our present projections for these variables are poor. Export growth between 1998 and 2005 at the 500-commodity level regressed against the corresponding growth between 1992 and 1998 gives an R^2 of 0.03. A similar regression for imports gives an R^2 of 0.07. Consequently, a forecasting method based on the assumption that trends from the first period continue into the second period is bound to fail.

Getting the movements right in the technology and preference variables makes a similar contribution to forecasting performance as the introduction of accurate forecasts for exogenous trade variables. The introduction of the truth for technology and preference variables reduces M from 0.36 to 0.14.

In thinking about the three sets of exogenous variables (macro/energy, trade and technology/preferences) and on looking at Table 19.3, our impression is that the greatest payoffs in terms of forecasting accuracy are likely to come from work in the trade area.[38] Enormous resources have already been devoted by US government agencies and other agencies around the world to macro and energy forecasting. Consequently, we should not expect significant improvements in those areas. On technology and preference variables, it is not clear that we can do much better than our historical extrapolations, although it may be helpful to make a detailed review of the technology assumptions built into the forecasts of demand for labor by occupation prepared by the Bureau of Labor Statistics (see, e.g. Bureau of Labor Statistics, 2008; Dixon and Rimmer, 2006). Little work has been done on projecting movements in import/domestic preferences and export-demand curves. Given the apparent importance of these variables in determining the commodity structure of US output, and the underdeveloped nature of research in the area, this seems to be the obvious direction to go.

19.6.1.4 How can USAGE forecasts beat or get beaten by trend forecasts?

USAGE forecasts incorporate trend assumptions for nearly all technology, preference and trade-shift variables. Movements in these variables are major determinants of changes in the commodity composition of US output. A reasonable question therefore is how does

[38] True values for exogenous variables could be introduced in a different order from that adopted in Table 19.3. We found that the contributions of accurate trade and technology forecasts to the USAGE performance does not depend strongly on the order in which these forecasts are introduced. In Table 19.3, M falls from 0.54 to 0.36 to 0.14 with the introduction of the correct information on trade followed by that on technology. When technology is introduced before trade, M goes from 0.54 to 0.35 to 0.14. In the first sequence the trade and technology contributions to the reduction in M are 0.18 and 0.22. In the second sequence they are 0.21 and 0.19.

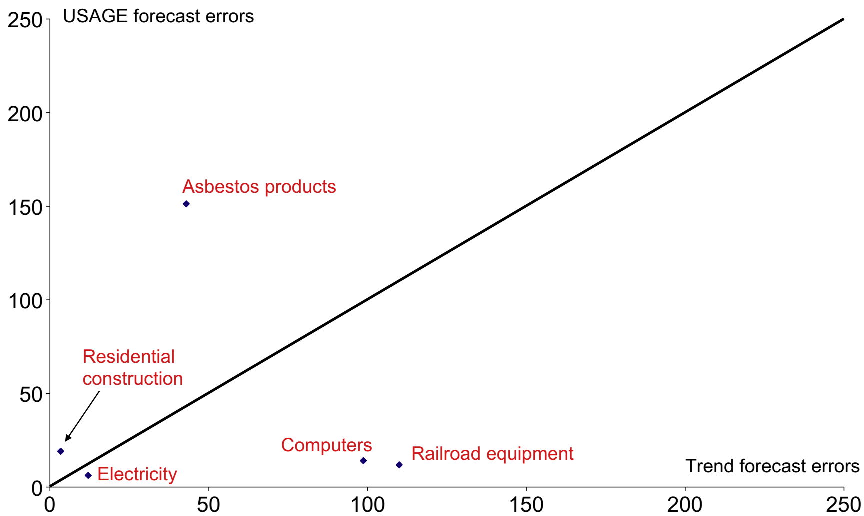

Figure 19.10 Percentage forecast errors for selected commodity outputs 1998–2005: Extrapolated 1992–1998 trend versus genuine USAGE forecast.

USAGE generate forecasts for commodity outputs that are distinctly different from trend forecasts.

There are three factors that contribute to the answer. (i) USAGE forecasts for commodity outputs are driven by macro and energy forecasts that deviate from trends. (ii) The starting point for the USAGE forecasts is different from the starting point for the historical trends. Sales structures for each commodity and rates of return and cost structures for each industry in 1998 differ from those in 1992. These differences mean that trends imposed with 1998 being the starting point can have different effects from similar trends imposed in 1992. For example, a given trend in foreign demand for commodity *i* will have a different effect on output depending on the share of exports in sales of commodity *i*. (iii) USAGE recognizes detailed demand linkages. Thus, for example, if the USAGE forecast for output or investment in industry *i* differs from the trend forecast and output of *j* depends heavily on sales to either output or investment in *i*, then the USAGE forecast for output of *j* is likely to differ from the trend forecast. To illustrate the operation of these factors we have taken results from Figure 19.9 for five commodities and reproduced them in Figure 19.10.

In the case of Computers, the superior USAGE forecast relative to the trend forecast is explained mainly by the macro input to the USAGE forecast simulation. The macro forecast for 1998–2005 correctly specified that the rate of growth in non-residential investment would be considerably lower than occurred in the period 1992–1998. This caused USAGE to correctly forecast a reduced rate of growth in the output of Computers which rely heavily on demand from investment. While the macro forecast for non-residential investment was relatively accurate, the macro forecast that we used for 1998–2005 for residential investment turned out to be far too low (a forecast growth of 18.2% compared with actual growth of 45.3%), causing USAGE to produce an inaccurate forecast for Residential construction. As residential investment grew strongly between 1992 and 1998, and continued to grow strongly between 1998 and 2005, the USAGE forecast for Residential construction for 1998–2005 was easily beaten by the trend forecast.

For Electricity, the superior performance of USAGE relative to the trend is explained by a good forecast, built into USAGE, from the Energy Information Administration. They correctly predicted a slowdown in the rate of growth of electricity consumption from 1998 to 2005 relative to the growth that occurred between 1992 and 1998.[39]

The superior performance of USAGE in forecasting the output of Railroad equipment illustrates the importance of both linkages and the starting point. Output of Railroad equipment is linked to investment in the railroad industry. Between 1992 and 1998 investment in the railroad industry grew rapidly. Thus, there was strong growth in the output of Railroad equipment. Under the trend forecasting method, Railroad equipment was given good prospects for 1998–2005. In fact, as correctly predicted by USAGE, output of Railroad equipment between 1998 and 2005 exhibited weak growth. The USAGE simulation encapsulated the fact that strong railroad investment between 1992 and 1998 left the railroad industry with abundant capital in 1998. With only moderate growth predicted for 1998–2005 in the use of railroad services, USAGE projected very little need for further investment purchases of Railroad equipment in the period 1998–2005.

The USAGE forecast for Asbestos products generated a point in Figures 19.9 and 19.10 further above the 45° line than that for any other commodity. The USAGE error is 151% whereas the trend error is 43%. The actual growth in the output of Asbestos products between 1998 and 2005 was −40%, following −12% between 1992 and 1998. The USAGE and trend forecasts are 51 and −14%. The trend forecast is not good (−14 versus −40%) but why was the USAGE forecast so awful (+51 versus −40%)? The answer can be found in the 1998 starting point.

Between 1992 and 1998, there was an increase in the use of Asbestos products per unit of output in using industries such as aircraft and motor vehicle parts. This was captured in the historical simulation as Asbestos-products-using technical change. Despite this, output of Asbestos products shrank due to a collapse in exports, portrayed in the 1992–1998 historical simulation as a sharp inward movement of the foreign demand curve. Both the favorable technical change and the unfavorable movement in the foreign demand curve were projected forward in the USAGE forecast for 1998–2005. With

39 For the genuine forecast simulation for 1998–2005 we used forecasts from Energy Information Administration (1996). In the historical simulation for 1998–2005 we used data from Energy Information Administration (2006).

virtually no exports of Asbestos products at the 1998 starting point for the forecast, the inward movement in the foreign demand curve did no harm to the commodity's output. However, with favorable technical change continuing to have a stimulatory effect, USAGE incorrectly forecast strong growth in the commodity's output.

19.6.1.5 Next steps: should we have done better?

As described in Section 19.5.2, Johansen (1960) analyzed gaps between projections from his 22-commodity model of Norway and outcomes with a view to pinpointing model weaknesses and making improvements. In this spirit, our colleague Peter Mavromatis has done about 18 months work on an analysis of the gaps between USAGE projections and outcomes.

He started this process by looking at commodities with the largest percentage errors (contributions to *AE*) in our 1998–2005 forecast. For each of these commodities he has been asking two questions:

(i) Was there information available in 1998 that should have been used to improve the forecast?

(ii) Was there a better way of using the information from the historical simulation for 1992–1998?

So far, for six of the large-error commodities he has given an affirmative answer to Question (i). These commodities are Asbestos products, Commercial fishing, Dolls, Theatres (mainly production and distribution of motion pictures), Electron tubes and Recordmedia (mainly blank tapes and videos). For Asbestos products, information not used in our forecast but available in 1998 suggested poor growth prospects on health and safety grounds: we should not have extrapolated Asbestos-product-using technical change from the 1992–1998 period into the forecast for 1998–2005. For Commercial fishing, Mavromatis argued that USAGE overestimated growth because information available in 1998 on environment constraints was overlooked. For Dolls, industry information available in 1997 and 1998 suggested, on the basis of a preference change towards collectable dolls, much stronger prospects than were forecast by USAGE. By the late 1990s, Theatres faced a well-recognized emerging threat from piracy that could have been taken into account in the USAGE forecasts. In 1998 it was clear that Electron tubes were being phased out in favor of plasma and LCD flat panels in televisions and computer monitors. Taking this into account would have produced a much less optimistic 1998–2005 forecast than was generated by USAGE. Mavromatis found credible information available in 1998 suggesting that Recordmedia was about to face serious competition from Chinese varieties and from digital technologies. Again, taking this into account would have damped our 1998–2005 forecast.

In relation to Question (ii), Mavromatis has concentrated on the Textile, clothing and footwear (TCF) sector which contains 31 USAGE commodities. For most of these, our 1998–2005 forecast overstates output growth. Mavromatis found that this was caused mainly by overestimation of the movement between 1998 and 2005 in the ratio of landed-duty-paid import prices (P_M) to basic prices of domestic products (P_D). In other words, our forecasts implied that domestic TCF commodities were more competitive than turned out to be the case. This caused Mavromatis to think about how we had projected P_M and P_D in our forecasts.

For each commodity, the movement in P_M is determined by movements in the nominal exchange rate, the foreign currency price and the US tariff. Between 1992 and 1998 the exchange rate appreciated by 13% and tariffs were reduced. However, the reduction in the P_Ms was muted by growth in the foreign currency prices. In our forecasts for 1998–2005 we projected forward the increases in foreign currency prices and assumed correctly that there would be little change in the exchange rate and that US tariffs would continue to fall at approximately the same rate as in the earlier period. Thus, via the exchange rate effect, our 1998–2005 forecasts showed stronger growth in the P_Ms in the TCF sector than was the case in 1992 to 1998. For the P_Ds, an important determinant is primary factor productivity growth. In our 1998–2005 forecast we projected forward rapid TCF primary factor productivity growth that occurred between 1992 and 1998.

Mavromatis thinks that our projection methods for both the P_Ms and P_Ds in the TCF sector can be improved. He notes that TCF imports to the US are heavily regulated by tariffs, quotas, voluntary export restraints and antidumping actions. He argues that TCF exporters to the US recognize that if they allow their product prices to fall too rapidly on the US market then they are likely to face restrictions. Thus they adjust their foreign currency prices in response to movements in the US exchange rate. In these circumstances, Mavromatis argues that it is better to project forward movements in the real prices of TCF imports (the ratio of P_M to the CPI) than movements in their foreign currency prices. He sees this as an assumption of constancy in the way that US policy makers handle the balance between the interests of the US TCF sector (that are harmed by falling P_Ms) and those of US households (that are enhanced by falling P_Ms). With regard to the P_Ds, Mavromatis argues that productivity growth in the TCF sector was enhanced in the 1992–1998 period by merger activity and closure of low-productivity plants. He thinks that it was clear by 1998 that this process had run its course and that the rate of productivity growth in the US would slow. He suggests that we should have built a productivity pause into our 1998–2005 forecast.

Mavromatis has rerun our 1998–2005 forecast. He has incorporated extra information for the six large-error commodities, projected real import prices for TCF commodities as a trend from 1992 to 1998 instead of projecting the trend in foreign currency prices and assumed no further primary factor productivity growth in the TCF sector instead of projecting forward the rapid productivity growth from 1992 to 1998. His results (labeled 'improved') together with our original results are presented in Table 19.4.

Segments (2) and (3) of Table 19.4 show that the Mavromatis changes dramatically improve the USAGE forecasting performance for the six large-error commodities and

Table 19.4 *M* and *AE* coefficients for USAGE forecasts of growth in commodity outputs between 1998 and 2005

Commodity coverage	Forecast type	AE	M
(1) Whole economy: 503 commodities	Original	18.9	0.58
	Improved	16.1	0.50
(2) Improvable "large-error" commodities excluding TCF sector: six commodities	Original	90.5	1.37
	Improved	34.9	0.53
(3) TCF sector: 31 commodities	Original	66.2	0.74
	Improved	25.2	0.28

for the 31 TCF commodities. *AE* coefficients for these two groups fall from 90.5 and 66.2 to 34.9 and 25.2. Rather than being outperformed by trend forecasts, with the extra information incorporated by Mavromatis the USAGE forecast for the six large-error commodities easily beats the trend forecast (M falls from 1.37 to 0.53). For the TCF commodities, the original forecast beat the trend forecast ($M = 0.74$), but with the Mavromatis approach to P_M/P_D, the superiority of the USAGE forecast is strongly enhanced ($M = 0.28$).

Segment (1) of Table 19.4 shows that for the whole economy the Mavromatis improvements reduce *AE* from 18.9 to 16.1 and M from 0.58 to 0.50. To a pessimist these reductions are disappointingly small considering the enormous effort that Mavromatis has put into achieving them. A pessimist could also emphasize that Mavromatis has concentrated on the low hanging fruit implying that there is likely to be sharply diminishing returns to further effort. To an optimist, Mavromatis has delved into the results and, in the spirit of Johansen, has found areas for improvement while at the same time he has demonstrated the robustness of our original method.

19.7 CONCLUSION

Validity is a key issue for consumers of CGE modeling services. What assurance can producers of CGE results give to consumers that a CGE analysis: (i) Is computationally sound, (ii) uses accurate up-to-date data, (iii) adequately captures behavioral and institutional characteristics of the relevant part of the economy, (iv) is consistent with history and (v) is based on a model that has forecasting credentials?

In this chapter we have given some answers. With regard to (i), CGE modelers have an obligation to conduct exhaustive test simulations and to check national accounts identities. The value of these procedures goes beyond computational checking. Test simulations are a practical way to become familiar with a model and often reveal modeling weaknesses. As illustrated with our Chinese exchange rate example, computational checks can sometimes have policy-relevant interpretations.

On (ii) and (iii), the most effective approach for displaying the relevant data and describing what is going on is via a BOTE model. A well-designed BOTE model has two properties: it reveals the roles of major behavioral, institutional and data assumptions in causing a model to generate a particular result, and it is small enough to be managed with pencil and paper (on the back of an envelope) and to be presented in a limited timeframe to policy advisors. A good BOTE model can be used to predict how the results from the main model will change in response to changes in data items, parameter values and behavioral assumptions. Tests of the BOTE model as a result-explaining device can be carried out by comparing these predictions with results from the main model computed with reset values for data and parameters and with alternative behavioral assumptions. Another device for testing the veracity of an explanation of CGE results is regression analysis applied to the results from the main model.

Given any CGE result, can we always design a BOTE model that explains it? A message from general equilibrium theory is that every part of the economy is connected with every other part. Consequently it is not clear on purely theoretical grounds that every CGE result can be explained by a BOTE model that is very much simpler than the original model. However, our experience is that an informative BOTE model *can* always be found. It is not easy to teach the BOTE technique – every analysis requires a different BOTE model to pick out and highlight the cogent features of the main model. What is required is perseverance and practice.

BOTE analysis is a critical component of validation and consumers of CGE analyses can often be won over by it. However, it is not the approach that first comes to their minds. Almost invariably they think in terms of statistical validation, (iv) and (v) above. In this chapter, we have described various aspects of statistical validation, concentrating mainly on historical simulation, baseline forecasting and the testing of baselines against reality. This work demonstrates that CGE models can produce forecasts at a highly disaggregated level that comfortably beat non-model-based trend forecasts. It also demonstrates that there is considerable potential for improved CGE forecasts through conscientious data work and improved methods for projecting trends from historical simulations into forecasting simulations. What it does not do is cast light on the statistical validity of CGE results for the effects of particular policies.

The problem is that during any period in which an economy is adjusting to a policy change, other factors will also be operating. This point was not adequately addressed in the often-cited validation exercise by Kehoe (2005).[40] In that exercise, Kehoe assesses the performance of various models in predicting the effects of NAFTA. He notes that the

40 The issue is better handled by Kehoe *et al.* (1995) who allow for other factors in a validation exercise concerning the short-run effects on consumer prices in Spain of a major change in indirect taxes. They recognized that during the adjustment period consumer prices were also affected by a drought-induced change in the price of agricultural products and a large change in the world price of oil.

model of Brown, Deardorff and Stern (BDS) predicted that NAFTA would increase Mexican exports by 50.8%. Over the period 1988–1999, Mexican exports went up by 140.6%. Kehoe invites us to draw the conclusion that the BDS model strongly underestimated the effects of NAFTA. However, what about all the other factors that affected Mexican trade volumes over these 10 years?

A possible methodology for tackling statistical validation of policy effects is the MONASH decomposition technique (see Dixon *et al.*, 2000; Harrison *et al.*, 2000 and Dixon *et al.* in Chapter 2 of this Handbook). This technique allows separation of the effects of policy changes including trade reforms from those of other factors, such as changes in technologies, changes in import–domestic preferences, changes in consumer preferences, changes in world commodity prices, changes in population, changes in required rates of return on capital and changes in transport costs. However, a major assumption in existing applications of decomposition simulations is that changes in policies do not affect these other factors. In a validation exercise concerned with a trade policy for example, this assumption would need to be tested. Links between technologies and trade policy have been suggested by a long stream of authors including Leibenstein (1966, X-efficiency), Krueger (1974, rent seeking), Harris (1984, scale economies and imperfect competition) and Melitz (2003, reallocation of resources between firms within an industry). Links between import/domestic preferences and trade policy have been hypothesized by several authors including Feenstra (1994, variety and the "price" motivating import demands) and Dixon and Rimmer (2002, p. 60 & 71, variety and the nature of import restraints).

Confrontation between reality and CGE results, whether for baseline forecasts or for the effects of particular policies, is vital in the continuing development of CGE modeling as a valuable tool in policy formulation.[41] To foster this confrontation, CGE modelers should welcome opportunities to explain results to specialists on different aspects of the economy and to absorb feedback. This was recognized by Leif Johansen in connection with projections made with his model in 1969 by the Norwegian Ministry of Finance for the period 1963–1990:

> *The projections described above aroused great interest in wide circles. … Several institutions, enterprises and persons approached the Ministry of Finance to get more details or suggest alternative assumptions. There was a clear need to continue the explorations and the Ministry invited interested persons to take part in an informal working group. Some 30 persons representing organizations, research institutes, private firms and other ministries participated. This*

[41] A good example of productive confrontation is given by Valenzuela *et al.* (2007). Relative to real-world outcomes, their initial modeling of the effects on wheat prices of variations in climatic conditions showed too much purchaser-price volatility in import destinations and too little producer-price volatility in export sources. They rectified this modeling weakness by introducing real-world mechanisms, such as variable tariffs, that damp price effects in importing countries. These same mechanisms accentuate price volatility in exporting countries by preventing price signals from exercising their full potential effect on demand in importing countries.

group met regularly with members of staff of the Ministry in 1970. Several aspects and possible uses of the projections were discussed and many alternatives were tried. ... Among alternatives tried were some with variations in the expansion of ocean transport (a very important element on the Norwegian balance of payment), some with variations in the rate of growth of total labor force and alternatives with lower overall growth generated by lower investment proportions and slower technological progress.

(Johansen, 1974, pp. 225−226)

Following Johansen, we envisage a process in our USAGE work in which preliminary forecasts are made and documented, and then tested in discussions with industry experts from the US International Trade Commission and the Departments of Commerce, Agriculture and Energy. These four agencies use the model and are well placed to provide feedback on institutional and technological factors that should be taken into account in analyses at a detailed level.

APPENDIX: RELATIONSHIP BETWEEN THE PERCENTAGE CHANGES IN OUTPUT, INPUTS AND TECHNOLOGY VARIABLES: DERIVATION OF (19.12) AND (19.15)

Assume that input quantities $X_1, \ldots, X_n$ are chosen to minimize:

$$\sum_i P_{\mathrm{i}} * X_i. \tag{19A.1}$$

subject to:

$$Y = F\left(\frac{X_1}{A_1}, \ldots, \frac{X_n}{A_n}\right). \tag{19A.2}$$

where P_{i} is the price of input i, Y is output, F is a constant-returns-to-scale production function and A_i is a technological coefficient (input-i-saving variable).

The first-order conditions are (19A.2) and:

$$P_{\mathrm{i}} = \Lambda * F_i * \frac{1}{A_i}, \quad i = 1, \ldots, n, \tag{19A.3}$$

where Λ is the Lagrangian multiplier and F_i is the partial derivative of F with respect to its ith argument.

Total differentiation of (19A.2) gives:

$$\mathrm{d}Y = \sum_i F_i * \mathrm{d}(X_i/A_i). \tag{19A.4}$$

Now substitute from (19A.3) into (19A.4):

$$\mathrm{d}Y = \sum_i \left(\frac{P_{\mathrm{i}} * A_i}{\Lambda}\right) * \mathrm{d}(X_i/A_i). \tag{19A.5}$$

From here we obtain:

$$y = \sum_i \left(\frac{P_i * X_i}{\Lambda * Y} \right) * (x_i - a_i). \tag{19A.6}$$

where y, x_i and a_i are percentage changes in Y, X_i and A_i.

Under constant returns to scale, Euler's theorem applied to (19A.2) gives:

$$Y = \sum_i F_i * \left(\frac{X_i}{A_i} \right). \tag{19A.7}$$

Via (19A.3), (19A.7) becomes:

$$Y = \sum_i \left(\frac{P_i * A_i}{\Lambda} \right) * \left(\frac{X_i}{A_i} \right). \tag{19A.8}$$

leading to:

$$\Lambda * Y = \sum_i P_i * X_i. \tag{19A.9}$$

Substituting from (19A.9) into (19A.6) gives:

$$y = \sum_i S_i * (x_i - a_i), \tag{19A.10}$$

where S_i is the share of input i in total costs.

Equations (19.12) and (19.15) are applications of (19A.10).

ACKNOWLEDGMENTS

We thank Janine Dixon for valuable comments.

REFERENCES

Adams, P.D., 2005. Interpretation of results from CGE models such as GTAP. J. Pol. Model. 27, 941–959.

Adams, P.D., Parmenter, B.R., 1993. The Medium-Term Significance of International Tourism for the State Economies. Australian Government Publishing Service, Canberra.

Adams, P.D., Parmenter, B.R., 1995. An applied general equilibrium analysis of the economic effects of tourism in a quite small, quite open economy. Appl. Econ. 27, 985–994.

Adelman, I., Robinson, S., 1978. Income Distribution Policy in Developing Countries: A Case Study of Korea. Oxford University Press, New York.

Bureau of Labor Statistics, 2008. National employment matrices, available at http://www.bls.gov/emp/empocc2.htm.

Cook, L.H., 1980. Validation of a Johansen-type multi-sectoral model: Norway, 1949–1961. PhD Thesis. Department of Economics, Monash University, Clayton.

Dixon, P.B., 2009. Comments on the Productivity Commission's modelling of the economy-wide effects of future automotive assistance. Econ. Paper. 28, 11–18.

Dixon, P.B., McDonald, D., 1993. An explanation of structural changes in the Australian economy: 1986–87 to 1990–91. Economic Planning Advisory Council Background Paper 29. Australian Government Publishing Service, Canberra.

Dixon, P.B., Rimmer, M.T., 1999a. The Government's Tax Package: Further Analysis Based on the MONASH Model. Appendix VI in Senate Select Committee on a New Tax System, First Report. Parliament of the Commonwealth of Australia, Canberra, pp. 173–247.

Dixon, P.B., Rimmer, M.T., 1999b. The Government's Tax Package: Short Run Implications for Employment by Industry, Region, Occupation, Age and Gender. Appendix 4 in Report of the Inquiry into The GST and a New Tax System. Senate Employment, Workplace Relations, Small Business and Education References Committee. Parliament House, Canberra.

Dixon, P.B., Rimmer, M.T., 1999c. Changes in indirect taxes in Australia: A dynamic general equilibrium analysis. Aust. Econ. Rev. 32, 327–348.

Dixon, P.B., Rimmer, M.T., 2002. Dynamic General Equilibrium Modelling for Forecasting and Policy: A Practical Guide and Documentation of MONASH. North-Holland, Amsterdam.

Dixon, P.B., Rimmer, M.T., 2004. The US economy from 1992 to 1998: Results from a detailed CGE model. Econ. Rec. 80 (Special Issue), S13–S23.

Dixon, P.B., Rimmer M.T., 2006. Employment by occupation and industry, 2004 and 2014. Technical Documentation, available at http://www.monash.edu.au/policy/ftp/techusage1.pdf.

Dixon, P.B., Rimmer, M.T., 2010a. Johansen's contribution to CGE modelling: Originator and guiding light for 50 years. CoPS/IMPACT Working Paper G-203, available at http://www.monash.edu.au/policy/ftp/workpapr/g-203.pdf.

Dixon, P.B., Rimmer, M.T., 2010b. Optimal tariffs: Should Australia cut automotive tariffs unilaterally? Econ. Rec. 86 (273), 143–161.

Dixon, P.B., Rimmer, M.T., 2010c. Validating a detailed, dynamic CGE model of the US. Econ. Rec. 86 (Special Issue), S22–S34.

Dixon, P.B., Parmenter, B.R., Ryland, G.J., Sutton, J., 1977. ORANI, a General Equilibrium Model of the Australian Economy: Current Specification and Illustrations of Use for Policy Analysis, Volume 2 of the First Progress Report of the IMPACT Project. Australian Government Publishing Service, Canberra.

Dixon, P.B., Harrower, J.D., Vincent, D.P., 1978. Validation of the SNAPSHOT model. Preliminary Working Paper SP-12, IMPACT Project, Melbourne.

Dixon, P.B., Parmenter, B.R., Sutton, J., Vincent, D.P., 1982. ORANI: A Multi-Sectoral Model of the Australian Economy. Contributions to Economic Analysis vol. 142. North-Holland, Amsterdam.

Dixon, P.B., Parmenter, B.R., Powell, A.A., 1984. The role of miniatures in computable general equilibrium modelling: Experience from ORANI. Econ. Model. 1, 421–428.

Dixon, P.B., Parmenter, B.R., Powell, A.A., Wilcoxen, P.J., 1992. Notes and Problems in Applied General Equilibrium Economics. North-Holland, Amsterdam.

Dixon, P.B., Menon, J., Rimmer, M.T., 2000. Changes in technology and preferences: A general equilibrium explanation of rapid growth in trade. Aust. Econ. Paper. 39, 33–55.

Dixon, P.B., Rimmer, M.T., Tsigas, M.E., 2007. Regionalising results from a detailed CGE model: Macro, industry and state effects in the US of removing major tariffs and quotas. Paper. Reg. Sci. 86, 31–55.

Dixon, P.B., Rimmer, M.T., Wittwer, G., 2011. Saving the Southern Murray–Darling Basin: The economic effects of a buyback of irrigation water. Econ. Rec. 87 (276), 153–168.

Energy Information Administration, 1996. Annual Energy Outlook 1996, with Projections to 2015. Office of Integrated Analysis and Forecasting, US Department of Energy, Washington, DC.

Energy Information Administration, 2006. Annual Energy Outlook 2006, with Projections to 2030. Office of Integrated Analysis and Forecasting, US Department of Energy, Washington, DC.

Feenstra, R.C., 1994. New product varieties and the measurement of international prices. Am. Econ. Rev. 84, 157–177.

Harris, R.G., 1984. Applied general equilibrium analysis of small open economies with scale economies and imperfect competition. Am. Econ. Rev. 74, 1016–1032.

Harrison, W.J., Horridge, J.M., Pearson, K.R., 2000. Decomposing simulation results with respect to exogenous shocks. Comput. Econ. 15, 227–249.

Hertel, T.W., Hummels, D., Ivanic, M., Keeney, R., 2007. How confident can we be of CGE-based assessments of Free Trade Agreements? Econ. Model. 24, 611–635.

Johansen, L., 1960. A Multisectoral Study of Economic Growth. Contributions to Economic Analysis, 21. North-Holland, Amsterdam.

Johansen, L., 1974. A Multisectoral Study of Economic Growth, second enlarged ed. Contributions to Economic Analysis 21' North-Holland, Amsterdam.

Johnson, D., 1985. The short-term economic effects of environmental constraints on forest industries. Rev. Market. Agr. Econ. 53, 149–156.

Jorgenson, D.W., 1984. Econometric methods for applied general equilibrium analysis. In: Scarf, H.E., Shoven, J.B. (Eds), Applied General Equilibrium Analysis. Cambridge University Press, New York, pp. 139–203.

Kehoe, T.J., 2005. An evaluation of the performance of applied general equilibrium models of the impact of NAFTA. In: Kehoe, T.J., Srinivasan, T.N., Whalley, J. (Eds), Frontiers in Applied General Equilibrium Modeling: Essays in Honor of Herbert Scarf. Cambridge University Press, Cambridge, pp. 341–377.

Kehoe, T.J., Polo, C., Sancho, F., 1995. An evaluation of the performance of an applied general equilibrium model of the Spanish economy. Econ. Theor. 6, 115–141.

Krueger, A.O., 1974. The political economy of the rent-seeking society. Am. Econ. Rev. 64, 291–303.

Leibenstein, H., 1966. Allocative efficiency versus X-efficiency. Am. Econ. Rev. 56, 392–415.

Mai, Y.-H., 2006. The Chinese economy from 1997 to 2015: Developing a baseline for the McHuge model, Cops/IMPACT Working paper G-161, available at http://www.monash.edu.au/policy/elecpapr/g-161.htm.

Melitz, M.J., 2003. The impact of trade in intra-industry reallocations and aggregate industry productivity. Econometrica 71, 1695–1725.

Murphy, C., 1999. Modelling a New Tax System (ANTS)—Comparing MONASH and MM303. Appendix VII in Senate Select Committee on a New Tax System, First Report. Parliament of the Commonwealth of Australia, Canberra, pp. 248–304.

Peng, X., Mai, Y.-H., 2008. Population Aging, Labor Market Reform and Economic Growth in China: A Computable General Equilibrium Analysis. Cops/IMPACT Working paper G-174, available at: http://www.monash.edu.au/policy/ftp/workpapr/g-174.pdf.

Productivity Commission, 2000. Review of Australia's General Tariff Arrangements: Inquiry Report. Australian Government, Canberra, available at: http://www.pc.gov.au/projects/inquiry/tariff/docs/finalreport.

Productivity Commission, 2008. Modelling Economy-wide Effects of Future Automotive Assistance. Productivity Commission Report, Melbourne.

Taylor, L., Bacha, E.L., Cardoso, E.A., Lysy, F.J., 1980. Models of Growth and Distribution for Brazil. Oxford University Press for the World Bank, New York.

US International Trade Commission, 2004. The economic effects of significant US import restraints: Fourth update, Investigation 332-325: Publications 3701. US International Trade Commission, Washington, DC.

Valenzuela, E., Hertel, T.W., Keeney, R., Reimer, J.J., 2007. Assessing global computable general equilibrium model validity using agricultural price volatility. Am. J. Agr. Econ. 89, 383–397.