This paper contributes to the ongoing debate on the usefulness of economic models for policy analysis. It argues that there is a need for models which incorporate both the modern inter-temporal approach to macroeconomics and short-run ad-hoc behaviour. This need cannot be met by the simple models that permeate the macroeconomics literature, but requires large-scale simulation models such as the MSG2 and G-Cubed multi-country models. It is shown that these models give insights into the adjustment process in a number of historical episodes which are not well explained by simple macroeconomic frameworks: Reaganomics in the 1980s; German reunification in the early 1990s; fiscal consolidation in Europe in the mid-1990s; the formation of NAFTA; the Asian Crisis; and the current productivity boom in the USA. The paper also argues that using a well-defined theoretical specification but introducing real-world rigidities can also help to explain the ‘six major puzzles in international macroeconomics’ highlighted in a recent paper by Obstfeld and Rogoff (2000).

I. INTRODUCTION

There is a thriving debate on the usefulness of different kinds of models for the analysis of economic policy. This debate has many dimensions across a number of different literatures—big versus small models, structural versus reduced-form models, theoretical versus data-intensive models,
forward-looking versus backward-looking models, static versus dynamic models, and models that impose inter-temporal budget constraints versus those that do not. Paul Krugman opens this issue of the *Oxford Review of Economic Policy* with a passionate ‘keep it simple, stupid’ plea for his preferred resolution to this debate.

The present paper continues the debate. In our view, small, simple, reduced-form macroeconomic models are inadequate tools for understanding many of the major events in the global economy over the past two decades. We think that this is true both for the ‘theoretically elegant’ optimizing models which dominate the academic literature, and for the ‘ad-hoc’ models which Krugman advocates. In our view there is an important role for large-scale, dynamic, inter-temporally optimizing, general-economy models which combine rigorous theoretical foundations with a degree of stickiness and ad-hoc short-run behaviour.

This paper first sets out our preferred approach to modelling, as it is embodied in the MSG2 (see McKibbin and Sachs, 1991) and G-Cubed (see McKibbin and Wilcoxen, 1998a) multi-country models. The key feature of these models is that they combine the modern inter-temporal optimization approach to modelling economic behaviour (as found in Sargent (1987), Blanchard and Fischer (1989), and Obstfeld and Rogoff (1996)) with short-run rule-of-thumb behaviour. In doing this they bring together real-business-cycle models, modern macroeconometric models, and traditional computable general equilibrium (CGE) models into a single framework. That framework can be used for the analysis of a wide range of macroeconomic issues: both the responses to shocks, and the design of monetary and fiscal policy to deal with these shocks. The G-Cubed model can also be used to analyse sectoral policy issues (trade liberalization, tax policy, environmental policy, etc.), but that will not be our focus of attention here. A key feature of both of these models is the way they model asset markets and the integration of these markets with real economic activity, in particular the meshing together of highly mobile international flows of financial capital and short-run fixity of physical capital. Other key features of the models are set out in section II.

Over the past decade, these models have proved to be powerful aids to the understanding of major global events. Some of the lessons learned are summarized in section III of the paper. In our view these examples clearly demonstrate why it is important to incorporate both inter-temporal optimizing effects and stickiness in a policy-relevant economic model.

In section IV, we review where these models fit into the ongoing debate on the usefulness of models. In this section we also argue that some of the usual arguments for dismissing the application of rational expectations as a modelling approach in a real-world economic model are misguided.

Section V argues that an approach using a well-defined theoretical specification, but at the same time introducing real-world rigidities, can help to solve the ‘six major puzzles in international macroeconomics’ recently highlighted by Maurice Obstfeld and Ken Rogoff (2000). In this section a stylized two-country/two-sector version of the G-Cubed model is used to explore some of these puzzles and to show that the G-Cubed approach goes a long way towards offering an explanation. What the current paper cannot explain is the seventh puzzle in international macroeconomics. Why has it taken so long for theorists to understand what applied economists who build large-scale, theoretically based, simulation models have known for over a decade? We actually can build into an empirical general equilibrium framework real-world features which complement the underlying theory, in a way that better fits the data.

II. THE THEORETICAL AND EMPIRICAL BASIS OF THE MSG2 AND G-CUBED MODELS

(i) Core Structures

The MSG2 and G-Cubed suite of models are dynamic inter-temporal general equilibrium models that attempt to integrate the best features of traditional CGE models, real-business-cycle models, and Keynesian macroeconometric models.

A summary of the key features of these models is contained in Table 1 and the country and sector...
The five main features of both of these models are as follows.

(a) Both the MSG2 and G-Cubed models are based on explicit optimization by the agents (consumers and firms) in each economy. Where these models differ from static CGE models is in the assumption of inter-temporal optimization by economic agents, subject to explicit inter-temporal budget constraints. Thus, in contrast to static CGE models, time and dynamics are of fundamental importance in the G-Cubed and MSG2 models. This makes their core theoretical structures like those of real-business-cycle models.

(b) In order to track the macro time series, the behaviour of agents is modified to allow for short-run deviations from optimal behaviour, either due to myopia or to restrictions on the ability of households and firms to borrow at the risk-free bond rate on government debt. For both households and firms, deviations from inter-temporal optimizing behaviour take the form of rules of thumb, which are consistent with an optimizing agent that does not update predictions based on new information about future events. These rules of thumb are chosen to generate the same steady-state behaviour as optimizing agents so that, in the long run, there is only a single inter-temporal optimizing equilibrium of the model. In the short run, actual behaviour is assumed to be a weighted average of the optimizing and the rule-of-thumb assumptions. Thus, aggregate consumption is a weighted average of consumption based on wealth (current asset valuation and expected future after-tax labour income) and consumption based on current disposable income. This is consistent with the econometric results in Campbell and Mankiw (1987) and Hayashi (1982). Similarly, aggregate investment is a weighted average of investment based on Tobin’s q (a market valuation of the expected future change in the marginal product of capital relative to the cost) and investment based on current firm profit-income.

2 The G-Cubed model was constructed with grants from the United States Environmental Protection Agency, the US National Science Foundation, and the Brookings Institution.
Table 2

Global Inter-temporal General Equilibrium Models

<table>
<thead>
<tr>
<th>MSG2</th>
<th>G-Cubed</th>
<th>G-Cubed (Asia Pacific)</th>
<th>G-Cubed (Agriculture)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Countries:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>United States</td>
<td>United States</td>
<td>United States</td>
</tr>
<tr>
<td>Japan</td>
<td>Japan</td>
<td>Japan</td>
<td>Japan</td>
</tr>
<tr>
<td>Canada</td>
<td>Canada</td>
<td>Australia</td>
<td>Canada</td>
</tr>
<tr>
<td>Germany</td>
<td>New Zealand</td>
<td>New Zealand</td>
<td>Australia</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>Australia</td>
<td>Korea</td>
<td>Canada</td>
</tr>
<tr>
<td>France</td>
<td>Rest of OECD</td>
<td>Rest of OECD</td>
<td>Mexico</td>
</tr>
<tr>
<td>Italy</td>
<td>China</td>
<td>China</td>
<td>Korea</td>
</tr>
<tr>
<td>Austria</td>
<td>EEFSU</td>
<td>India</td>
<td>EU12</td>
</tr>
<tr>
<td>Australia</td>
<td>OPEC</td>
<td>Thailand</td>
<td>Rest of OECD</td>
</tr>
<tr>
<td>Mexico</td>
<td>Rest of world</td>
<td>Malaysia</td>
<td>ASEAN</td>
</tr>
<tr>
<td>Korea</td>
<td></td>
<td>Singapore</td>
<td>Taiwan</td>
</tr>
<tr>
<td>High-income Asia</td>
<td></td>
<td>Indonesia</td>
<td>China/Hong Kong</td>
</tr>
<tr>
<td>Low-income Asia</td>
<td></td>
<td>Hong Kong</td>
<td>Rest of world</td>
</tr>
<tr>
<td>Rest of the EMS</td>
<td></td>
<td>Taiwan</td>
<td></td>
</tr>
<tr>
<td>Rest of the OECD</td>
<td></td>
<td>Philippines</td>
<td></td>
</tr>
<tr>
<td>OPEC</td>
<td>OPEC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EEFSU</td>
<td>EEFSU</td>
<td>Rest of world</td>
<td></td>
</tr>
<tr>
<td>Rest of world</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sectors:

<table>
<thead>
<tr>
<th>MSG2</th>
<th>G-Cubed</th>
<th>G-Cubed (Asia Pacific)</th>
<th>G-Cubed (Agriculture)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single sector</td>
<td>Electric utilities</td>
<td>Energy</td>
<td>Food grains</td>
</tr>
<tr>
<td></td>
<td>Gas utilities</td>
<td>Mining</td>
<td>Feed grains</td>
</tr>
<tr>
<td></td>
<td>Petroleum refining</td>
<td>Agriculture</td>
<td>Non-grain crops</td>
</tr>
<tr>
<td></td>
<td>Coal mining</td>
<td>Durable manufacturing</td>
<td>Livestock</td>
</tr>
<tr>
<td></td>
<td>Crude oil &amp; gas</td>
<td>Non-durable manufacturing</td>
<td>Processed food</td>
</tr>
<tr>
<td></td>
<td>extraction</td>
<td>Services</td>
<td>Forest and fishery</td>
</tr>
<tr>
<td></td>
<td>Other mining</td>
<td>Textile &amp; clothing</td>
<td>Mining</td>
</tr>
<tr>
<td></td>
<td>Agriculture, fishing</td>
<td></td>
<td>Energy</td>
</tr>
<tr>
<td></td>
<td>Forestry &amp; wood products</td>
<td></td>
<td>Textile &amp; clothing</td>
</tr>
<tr>
<td></td>
<td>Durable manufacturing</td>
<td></td>
<td>Non-durable manufacturing</td>
</tr>
<tr>
<td></td>
<td>Non-durable manufacturing</td>
<td></td>
<td>Durable manufacturing</td>
</tr>
<tr>
<td></td>
<td>Transportation</td>
<td>Services</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Services</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: EMS is European Monetary System, ASEAN is Association of South-east Asian Nations, OPEC is Organization of Petroleum-Exporting Countries, EEFSU is Eastern Europe and the former Soviet Union.

(c) As in all policy-relevant macroeconomic models (but unlike in many CGE models), there is an explicit treatment of the holding of financial assets including money. Money is explicitly introduced into the model through a restriction that households require money to purchase goods. This assumption gives money an explicit role.

(d) Both the MSG2 and G-Cubed models allow for short-run nominal wage rigidity (by different degrees in different countries) and therefore allow for significant periods of unemployment depending on the labour-market institutions in each country. This assumption, when taken together with the explicit role for money,
is what gives the models their ‘macroeconomic’ characteristics. (Here again the models’ assumptions differ from the standard market-clearing assumption in most CGE models.)

(e) The models distinguish between the stickiness of physical capital within sectors and within countries and the flexibility of financial capital, which immediately flows to where expected returns are highest. This important distinction leads to a critical difference between the quantity of physical capital that is available at any time to produce goods and services, and the valuation of that capital as a result of decisions about the allocation of financial capital.

As a result of this structure, the models contain rich dynamic behaviour, driven on the one hand by asset accumulation and, on the other hand, by wage adjustment to a neoclassical steady state.

In more detail, the workings of the models are as follows (see also Table 1). In addition, the reader is referred to the Appendix for a skeletal summary of the main relevant equations.

(i) Both the demand and the supply features of the main economies are explicitly featured.

(ii) Demand equations are based on a combination of inter-temporal optimizing behaviour and backward-looking behaviour. The former kind of behaviour enables one to capture the kinds of insights which are the subject of study by real-business-cycle theorists. The latter kind of behaviour means that there are income feedbacks of the kind identified in the kind of simple IS curve treatment of aggregate demand commended by Krugman in his paper in this issue. The treatment of investment by firms is outlined in equations (1)–(16) of the Appendix. Firms attempt to maximize the present discounted value of their present and expected future profits. At any point in time, in which the capital stock is given, they employ the profit-maximizing amount of labour. Some firms choose the amount of investment which will add most to their present discounted value, at the prevailing value of Tobin’s q, given that the more they invest the larger will be marginal adjustment costs for the installation of each additional unit of capital. Other firms invest an amount which is constrained by their current profitability. It is normally assumed that 30 per cent of firms are of the former kind. The treatment of consumption, and investment, by households is spelled out in equations (17)–(31). Some households attempt to maximize the present discounted utility of consumption. This is done, given the current value of financial wealth owned by these consumers, by first calculating the value of human capital (the present discounted value of expected future labour earnings) and then distributing the consumption, to which financial wealth and human capital would entitle these consumers, optimally across the present and all future time periods. Other households consume an amount which is constrained by their current income. It is normally assumed that 30 per cent of consumers are of the former kind. Because the models of each country are models of open economies, there are also explicit demand equations for both exports and imports. These depend on foreign demand and relative prices (for exports) and on home expenditures and relative prices (for imports) in a standard manner.

(iii) On the supply side, productive potential is given by a constant elasticity of substitution (CES) production function (see equation (1) in the Appendix) which explicitly allows for the use of intermediate capital goods inputs (whose role in international trade is explicitly modelled). This specification of supply explicitly considers the effects of the accumulation of physical capital on the production function over time; in long-run equilibrium this accumulation takes place along a Swan/Solow/Ramsey neoclassical-growth-model path.

(iv) Prices adjust within any period to equate the supply of and demand for produced goods. Pricing behaviour takes explicit account of the effect of exchange rate changes on the costs of imports and thus on the costs of producing goods domestically. By contrast, wages adjust sluggishly to any imbalance between the supply of and demand for labour (see equation (32) in the Appendix). Wages are set one period in advance. Thus wage inflation between any one period and the next period depends on: price inflation between the previous period and the current one; the expectation, now, of price inflation between the current period and the next period; and
the ratio of current employment to full employment. This gives the model something of a fixed-price/ flexible-output character but only in the short run.

(v) Major flows of real and financial assets, such as private physical investment and public physical investment, fiscal deficits, and current-account imbalances cumulate into stocks of capital (and equity), infrastructure capital, government debt, and net external debt respectively. As a result, the level and the composition of national wealth changes over time.

(vi) Such wealth adjustment determines stock equilibrium in the long run, but also feeds back into short-run conditions through the effect of expected future economic conditions on prices in forward-looking share markets, bond markets, and foreign-exchange markets.

(vii) Financial assets are perfect substitutes both within economies and internationally. Asset prices adjust to equate rates of return on all assets, except for risk premia which are exogenous, and so asset prices are linked both within and between economies. In each national economy there are four kinds of financial asset: money, short-term bonds, long-term bonds, and equity. Money bears no interest and the short-term interest rate clears the money market. The values of long-dated government bonds are determined within economies by arbitrage equations linking the expected rate of change in bond prices to the difference between the return on long-term bonds and that on short-term bonds. Equities are treated in a similar way. Exchange rates are determined through uncovered interest parity equations (adjusted by risk premia) linking the rate of exchange-rate change to the difference between home and foreign interest rates and risk premia; eventually the differences between asset returns must converge to amounts equal to exogenous risk premia.3

(viii) Money is the numeraire in each economy. The nominal anchor can be made to vary across model closures, but in all of the experiments reported in this paper, the central bank in each country is assumed to fix the money supply.4

In summary, both the MSG2 and G-Cubed models embody a wide range of assumptions about individual behaviour and empirical regularities in a general equilibrium framework. The complex interdependencies are then solved out using a computer. It is important to stress that the term ‘general equilibrium’ is used here to signify that as many interactions as possible are captured, not that the economy is in a full market-clearing equilibrium at each point in time. Although it is assumed that market forces eventually drive the world economy to a neoclassical steady-state growth-equilibrium, unemployment does emerge for long periods owing to wage stickiness, to an extent which differs between countries owing to differences in labour-market institutions.

A stylized two-country/two-sector G-Cubed model (in which almost all of the inter-country and inter-sectoral detail is suppressed) is presented in the Appendix and is used for simulations in section V of the paper below.

(ii) Additional Features of the G-Cubed Model

The G-Cubed model combines the approach of the MSG2 model with the disaggregated econometric approach of Jorgenson and Wilcoxen (1990). It was constructed to contribute to the current policy debate on environmental policy and international trade with a focus on global-warming policies, but it has many features that make it useful for answering a range of issues in environmental regulation and microeconomic and macroeconomic policy questions. Like the MSG2 model it is a world model with substantial country disaggregation, but unlike the MSG2 model it contains considerable sectoral detail. With its sectoral detail and clear macroeconomic structure, the G-Cubed model has been designed to provide a bridge between CGE models that traditionally ignore the adjustment path between equilibria, real-business-cycle models that have a well-articulated supply side and a dynamic structure, and macroeconometric models that tend to ignore individual behaviour and the sectoral composition of economies that have a well-defined demand structure with rigidities in various markets.

3 These risk premia are determined empirically in the process of fitting the model to country data.
4 Experiments with different kinds of monetary rules are discussed in Henderson and McKibbin (1993)
The G-Cubed model is in the process of continual development, but it is already a large model. In its current form it contains over 8,000 equations and 150 inter-temporal costate variables. The country and sectoral coverage of the models are summarized in Table 2 (column 2). The G-Cubed (Asia Pacific) model draws on the theoretical approach of the G-Cubed model, but focuses on a country and sectoral disaggregation of relevance for the Asia Pacific region (see McKibbin, 1998a). It was developed originally to focus on trade liberalization and financial liberalization issues, but has proved useful in analysing the causes and consequences of the Asian crisis. The country coverage and sectoral detail is set out in Table 2 (column 3). The G-Cubed (Agriculture) model was developed for the United States Department of Agriculture to analyse the impact of changes in global macroeconomic conditions on US agriculture. It is outlined in McKibbin and Wang (1998).

III. INSIGHTS FROM THESE MODELS FOR UNDERSTANDING HISTORICAL EPISODES

In this section, a number of historical experiences since the early 1980s are interpreted using the published results from studies using the MSG2 and G-Cubed models undertaken at the time of the shocks. These episodes are examined to show what the models predicted at the time, in order to demonstrate how simple macroeconomic models would not have been able to give a full understanding of the impact of the shocks.

(i) The Effects of Reaganomics on Europe in the Early 1980s

One insight from the model is that the short-term fiscal multipliers are positive in most countries, but much less than unity, and turn negative after a few years, remaining negative in the longer run in the face of sustained increases in government expenditure. The reason for the small short-term multipliers is that the positive short-term demand stimulus from higher government spending is offset by the negatives operating through the inter-temporal budget constraints facing governments and countries: inter-temporal budget constraints project the future budgetary factors into the present through bond, equity, and foreign-exchange equity markets. Thus, even if solvency on the part of the government requires that a fiscal deficit resulting from higher government spending is ultimately financed by means of tax increases, the model must still generate a rise in interest rates in the long run, so as to keep consumption and/or investment sufficiently in check in the long term (because the fiscal expansion is sustained). That increases the long-term interest rate in the present, depresses equity valuations, and crowds out private investment in the short term. In addition, there will be an immediate appreciation of the real exchange rate, because of the higher interest-rate profile over time. That enables a transfer of resources from abroad into the domestic economy, so as partly to finance the fiscal expansion without sacrificing domestic consumption. Over time, resources must be transferred back to foreigners to service the higher foreign debt, so that eventually the real exchange rate will depreciate. In addition, nominal wages gradually rise, which eventually unwinds the temporary stimulus to output as a result of pressures from the supply side. Thus, to summarize, the fiscal stimulus raises demand directly through higher government spending; however, there is an immediate offset through weak investment induced by higher long-term real interest rates, and a decline in net exports.

The small size and temporary nature of fiscal multipliers may seem like conventional wisdom now, but in the mid-1980s when the first papers using the MSG2 model were written, most models had fiscal multipliers well in excess of unity (and sometimes greater than two).

In addition, in this early work with the MSG2 model, the transmission of a fiscal stimulus from the United
States across national borders tended to be negative. This was because the financing impacts through a rise in global long-term real interest rates tended to more than dominate any positive stimulus through a short-term increase in exports to the US economy. The comparative volume by Ralph Bryant and others (1988) is a good illustration of what an outlier the MSG2 model appeared to be in the 1980s. Other models displayed positive transmission abroad of a US fiscal shock, in the manner described in a conventional Mundell–Fleming model. The key reason driving the results on transmission in the MSG2 model was again the way in which the inter-temporal budget constraints projected the future budgetary factors into the present through bond, exchange-rate, and equity markets. The dominant effect abroad of this financing shadow from the future was through the impact of sustained future US government dis-saving on future short-term real interest rates, worldwide, and thus on current long-term real interest rates in foreign countries, immediately reducing investment there. The model showed clearly that the financing constraint effect was likely to dominate the positive spill-over of temporarily higher US demand.

As shown in McKibbin and Sachs (1991), the MSG2 model tracked many features (across a range of variables) of the 1980s quite well. In the US economy in the early 1980s, there was a monetary contraction at the turn of the decade which induced an economic slow-down in the USA. This was coupled with an announced fiscal expansion associated with Reaganomics which caused crowding-out of private spending through high long-term real interest rates and a strong US dollar before the spending actually began in late 1981. Thus, according to the account given by the MSG2 model, the US recession of the early 1980s was due both to a monetary contraction in the late 1970s and an expected fiscal expansion in 1980, raising real interest rates and the US dollar well before the actually spending increase occurred in 1981. Between 1982 and 1985 fiscal expansions in the USA in each year were larger than expected (see the forecasts of the OECD and the actual outcomes over this period, given in Table 5.5 of McKibbin and Sachs (1991, p. 108). This provides an explanation of why the US dollar and long-term real interest rates kept ratcheting up into 1985. The model also argued that the impact of the US fiscal stimulus on Europe was negative because the capital inflow into the USA to finance the US fiscal deficit raised European long-term real interest rates and crowded out investment in Europe. At the same time, Europe was undertaking spending cuts. In 1985, the US dollar swung around. This the model explained by means of: (i) lagged dynamics from the earlier fiscal expansions (as noted, the model suggests that a depreciation would follow the initial appreciation), plus (ii) Gramm–Rudman announcements of future fiscal contractions in the USA, and (iii) a relaxation of the monetary policy stance in the United States.

The important insight from the model in explaining this period of history was the need for differential treatment of short-term real interest rates driven by monetary policy, and long-term real interest rates driven by expected future fiscal policy and by longer-term productivity trends. The latter kind of analysis is not possible without the kind of full inter-temporal model being discussed here. The adjustment of asset prices in the short term which causes expected future events to be projected into the current economy is an important part of the story. Europe was stuck in a period of slow growth not only because of its own policies, but also because of the negative transmission of the US fiscal policy of the 1980s.

(ii) German Reunification

The MSG2 model was used for a number of papers on German unification in the early 1990s. The projections given by the MSG2 model, as well as those coming from two other similar models from the Federal Reserve and IMF, were compared with the actual outcomes in a paper by Gagnon et al. (1996) several years later. This was a case where the inter-temporal aspects of the MSG2 model, the Federal Reserve’s MX3 model, and the IMF’s Multimod models, all made a difference to the understanding of the adjustment to unification. The MSG2 model projected that the large future fiscal deficits in Germany required to finance the unifica-

---

6 These results were very much in accord with claims put forward at the time by Phelps and Fitoussi (1986) about ‘Stagflation in Europe’.
tion process would drive up German long-term real interest rates and shift capital out of the rest of Europe and the world into financing the German fiscal deficit. This would cause long-term real interest rates to rise globally and crowd out private investment in Germany as well as in the rest of Europe. In addition, the strengthening of the German mark as capital flowed into Germany, would cause short-term interest rates to rise throughout Europe owing to the Exchange Rate Mechanism (ERM) of the European Monetary System (EMS). The rise in short-term interest rates would further slow other European economies. Thus, it was predicted that there would be pressure on the ERM and a slow-down in non-German Europe. The negative impact of German policy on the rest of Europe is probably conventional wisdom now, but at the time it was argued by many that the boom in Germany from rebuilding of the East would cause a rise in demand throughout Europe and a boom in Europe.

What the MSG2 model showed was that the long-term inter-temporal budget constraints acting through capital flows and adjustment to real long-term interest rates acted as a shadow over the short-term effects of fiscal spending in Germany. The model also suggested that the problems of real resource transfers throughout Europe and inter-temporally, were what was driving long-term real interest rates and that, in the main, this was independent of the details of the reaction of the Bundesbank, with its setting of short-term nominal interest rates. Nevertheless, the reactions of German monetary policy did affect whether this real adjustment happened through price adjustment or quantity adjustment. As it turned out, the Bundesbank tightened monetary policy significantly, which caused a greater slowdown in the rest of Europe than might have been achieved if policy had been looser. What the studies using the three models did not predict well was the extent of this additional monetary tightening in Germany. Therefore, the three models under-predicted the size of the negative effects of unification, but captured the key adjustment mechanisms.

(iii) The Impacts of NAFTA

In a study for the US Congressional Budget Office report on the North Atlantic Free Trade Area (NAFTA) (CBO, 1993, subsequently published as Manchester and McKibbin, 1995), the MSG2 model was used to assess the impact of the trade agreement between Canada, the USA, and Mexico. The conventional wisdom at the time, as well as results from most (if not all) early CGE studies, was that NAFTA was expected to lead to a flood of cheap goods into the US economy and, as Ross Perot argued in the 1992 election campaign, the giant sucking sound of US jobs to Mexico.7 The MSG2 results showed the opposite.

The study was based on the proposition that the key macroeconomic aspect of the agreement was not actually the removal of tariffs in the USA on Mexican goods, but the impact of NAFTA on (i) expected future productivity in Mexico and (ii) the reduction in the risk premium attached to holding of Mexican assets. The model predicted that NAFTA would lead to a large flow of financial capital from the rest of the world into the Mexican economy in response to the induced rise in the expected return to capital, and the reduction in the risk premium, in the Mexican economy. This was predicted to cause the Mexican real exchange rate to appreciate and crowd out net exports. In the short run, the capital inflow into Mexico would be matched, not by a surge of Mexican exports into the US economy, but by a rise in the Mexican current account deficit. This was how the resources were to be transferred into the Mexican economy to take advantage of the increase in the risk-adjusted rate of return to investment there. As it turned out, the model prediction of a giant sucking sound of capital to Mexico and an associated trade deficit with the rest of the world—rather than a trade surplus through higher exports of low-cost goods to the United States—is what actually happened. Also the study warned of an inappropriate monetary tightening in Mexico in response to a widening current account deficit.

The key features of the model that captured the essence of NAFTA were that it predicted a big impact of expected long-term productivity improvements, and that it showed how, through the operation of inter-temporal forces, this stimulated short-term capital inflows into Mexico. This completely dwarfed the static effect of the tariff changes between the USA and Mexico, which was the focus of the CGE studies of the effects of NAFTA.

7 See for example the survey by Brown (1992) and other studies in Lustig et al. (1992).
described above. Only later, after the event, did the CGE literature begin to overlay its static story with the kinds of capital-flows effects on which the MSG2 model focused from the start.

(iv) The Effect of Fiscal Consolidation in Europe

The MSG2 model has been used for a number of papers on European issues such as the impact of the Maastricht targets (McKibbin, 1994) and the Stability and Growth Pact (Allsopp et al., 1999). The key results from these papers were that a permanent fiscal contraction implemented upon announcement would have a small negative effect on GDP in the year of the cut in spending, but a positive effect thereafter. The key point was the same as in the US fiscal policy story in which the asset-price adjustment (exchange rate, long-term interest rates), adjusting owing to the inter-temporal budget constraints, would, after a delay of just one year, come to dominate any short-term Keynesian stimulus sustained by short-term rigidities in labour markets and by liquidity constraints. In the case of a phased cut in fiscal deficits, which is closer to the Maastricht arrangements as announced in 1991, the issue is more interesting. The phasing, by its nature, imposes a small short-term fall in fiscal spending. But the long-term financing issues are the same as for the unphased permanent fiscal cuts. Thus, in the short term, a smaller short-term cut combined with the same positive financing effects of lower expected future fiscal deficits, implies that a phased fiscal contraction would tend to be less contractionary in the short run (and could be expansionary) relative to an unphased permanent fiscal contraction. Applied to the announced phased fiscal contraction in Europe under the Maastricht Treaty, the model predicted a positive effect on European GDP throughout the fiscal consolidation (see McKibbin, 1994). In the short term, there is crowding in of investment through lower long-term real interest rates and a rise in net exports owing to a depreciation of the European currencies. Yet the spending cuts occur more in the future than the present, so that by the time the actual cuts occur, there is enough stimulus induced by the forward-looking response and inter-temporal budget constraints that the adjustment costs and wage stickiness do not cause a slow-down in GDP.

This story was also predominant in the results from the MSG2 model in Allsopp et al. (1999) on the impacts of the Stability and Growth Pact. In those results, the positive effects of phased-in future budget cuts on long-term interest rates projecting the long-term supply-side gains through investment into short-run demand gains, were sufficient (because of timing) to offset the standard contractionary effect of a fiscal contraction. This was very different from the results of other studies done at the time, because of the role of inter-temporal financing constraints in the MSG2 model and the way these were projected into short-term decisions through asset markets.

To summarize: a permanent announced fiscal contraction is more contractionary in the short run than an announced and gradually phased-in fiscal contraction. The critical issue in the case of European fiscal policy during the 1990s was one of the timing of the changes in spending and taxes. This timing had little impact on the financing consequences (which were positive), but had a big impact on the extent to which there would be negative short-run Keynesian effects.

(v) The Asian Crisis

Perhaps the most controversial results obtained from the models were produced in 1997, when they were used to predict the causes and consequences of the Asian crisis that began in 1997. In McKibbin and Martin (1998), the G-Cubed (Asia Pacific) model was used to simulate the crisis. That paper used data from the key crisis economies of Thailand, Korea, and Indonesia as inputs into model simulations to see if the model could generate the scales of adjustment in asset markets, in particular stock markets and foreign-exchange rates, as well as the sharp declines in economic activity, that actually occurred.

That paper first looked for evidence of traditional causes of crisis in emerging economies, such as terms-of-trade shocks, world interest-rate shocks,
or the fact that the crisis countries had exchange rates tied in various ways to the US dollar. Although perhaps each played some role, none of these explanations seemed large enough to explain the magnitude of the crisis.

The authors then used the models to explain the main features of the crisis as the outcome of just three shocks, which are straightforward to handle in the framework which we are discussing.

A fall in expected future growth

The authors examined the stock-market data for Thailand, which indicated a significant fall in share prices through 1996 and into 1997. They then translated that into a revision of expectations about future growth prospects in Thailand and used this as input into simulations of the G-Cubed Asia Pacific Multi-country Model. The model suggested that a revision in expected growth in 1997 of 0.5 per cent per year for 5 years (with a permanent reduction in the level of GDP after that) would lead to a 20 per cent depreciation of the Thai baht, a collapse in real investment of 10 per cent in the year of the revision, and a collapse of private consumption of 40 per cent in the first year. The consumption collapse was primarily due to a large capital loss associated with a sharp rise in US-dollar-denominated foreign debt and the collapse in the value of physical capital. These wealth effects were exacerbated in the short run by the Keynesian multiplier. Similar magnitudes of revisions to growth were also shown to cause problems for Korea and Indonesia as well as other economies in the region. This explanation for Korea is also plausible because it was consistent with falling stock markets in Korea before the crisis. However, it was less plausible as an explanation for what happened in Indonesia because there were no indications of a fall in financial prices reflecting revisions to growth expectations until after the crisis emerged.

One interesting aspect of the simulation of this shock (in Thailand alone) was that, despite the severity of the shock and the size of real exchange-rate changes, the spill-overs through trade and capital flows to the rest of the region were reasonably small. This was because of offsetting effects of losses from trade flows but gains from capital inflows. (In fact, as discussed below, capital also flowed out of these countries, which suggested that shocks happened in synchronization.)

Risk premia

One of the important shocks that faced Thailand, Korea, and Indonesia was an increase in the risk premia on assets of these economies. The authors used the actual data on risk spreads in Eurocurrency markets on assets from these economies to impose a shock on these economies in the model. The impact of the increase in risk spreads was very similar to the results for a fall in expected productivity. Capital flowed out of crisis economies causing a sharp real and nominal exchange-rate depreciation. This reduced the value of capital, which together with a significant revaluation of US-dollar-denominated foreign debt, caused a sharp fall in wealth and a large collapse of private consumption expenditure. The fall in the return to capital and large rise in real long-term interest rates caused a fall in private investment. According to the model, this risk shock was crucial to understanding the Asian crisis. The importance of whether the risk shock is permanent or temporary is explored in McKibbin (1998b).

Actual and expected monetary responses

The third source of shocks considered by the authors was the impact of the actual and expected monetary responses in the crisis countries. Both Thailand and Korea reacted with a tightening of macroeconomic policy, whereas Indonesia had a dramatic monetary expansion in late 1997. The model results suggest that if a monetary loosening was expected in crisis countries, then the actual tightening that was administered would have larger real negative output consequences. In the model, this operated through nominal wages which fell by less than needed given the shock, in anticipation of a monetary relaxation. In the model, wages one period ahead are a weighted average of the change in the consumer price index (CPI) between the current period and last period and the expected change in the CPI between the current period and next period. Thus wages one period ahead, set now, depend on current expectations of prices one period ahead. Consequently, an anticipation of monetary relaxation would lead to a larger wage rise in the expectation of higher prices. But tighter monetary policy than expected would lead to a smaller price
rise than expected and so to a rise in real wages. This would lower real output even more. The model suggests that the impact of a monetary loosening to the extent experienced in Indonesia would be sufficient to further dramatically devalue the rupiah, with the consequence that the impacts on US-dollar-denominated foreign debt were exacerbated.

From this study, it appears that revisions of growth prospects, changes in risk perceptions, and policy responses in the individual countries can explain the qualitative and quantitative events that unfolded in the crisis economies.

It is clear that we require the kind of model being discussed here if we are to analyse the effects of these three features properly. First, expected growth revisions operated through changing current asset prices: these had income effects and wealth effects that were important. Second, in the risk shocks, the extent to which financial markets responded through inter-temporal arbitrage relations was crucial. Finally, being able to model the anticipated policy responses, both through price-setting and through asset-market adjustments, and being able to track the differences between expectations and the actual realizations of expected outcomes, was crucial to an understanding of the outcomes.

Early in the debate on the Asian crisis, the G-Cubed model was giving results which were counter to popular commentary on the global consequences, both in Australia and in the USA. The model showed that although the international trade effects were negative for countries that export to Asia, the international capital outflow would push down world interest rates and stimulate non-traded sectors of economies that were not also affected by changes in risk assessment. Thus the model suggested that Australia would only slow slightly in the short run and that the United States would actually experience stronger growth as a result of the capital reallocation. This is now conventional wisdom. Second, for Australia in particular, the existence of markets outside Asia and changes in relative competitiveness meant that substitution was possible for Australian exports. Thus, models with an aggregate world growth variable or a single exchange-rate variable would not capture this international substitution effect, which was an important part of the story.

The small spill-over effects of Asia in these studies is in contrast with the OECD analysis (OECD, 1997), in which results for economic spill-overs in 1997 for the USA were presented suggesting that growth was likely to be reduced by up to 0.3 per cent in 1997 and 0.7 per cent by 1998. Those results largely ignored the capital-flow effects, owing to the nature of the model used by the OECD (which focused largely on trade linkages between economies and on the demand side). Understanding the international capital-flow consequences of the Asian crisis, and the impact of these on supply potential in countries outside Asia, has become an important part of our understanding of the crisis. The G-Cubed model has made a useful contribution to this.

A key insight gained from the model relates to the difference between the fixity of physical capital in the short run and the flexibility of financial capital. In the Asian crisis, the value of capital fell when financial capital flowed overseas, but the physical capital remained in place. This had enormous wealth effects on consumption and direct rate-of-return implications for private investment. But the models suggested that recovery would be rapid when confidence returned, because the physical capital was still in place. Subsequently, recoveries in the affected economies have followed the paths suggested in the temporary risk shock simulations described in the above papers.

It is worth stressing a point about how the adjustment costs in physical capital affect the allocation of financial capital globally. Physical investment is assumed to be subject to rising costs of adjustment that are quadratic in the rate of investment ($I/K$). Thus if $100$ billion was injected into global capital markets, it would tend to flow into the largest economies first, since a dollar in the USA faces a smaller adjustment cost in being turned into physical capital than does a dollar flowing into New Zealand.

---

9 This assumption has a similar importance in the discussion below of the US share market in the late 1990s.
10 This claim is substantiated in McKibbin and Stoeckel (1999).
11 This result also shows up importantly in the climate-change research using the model (see McKibbin et al., 1999).
precisely because the capital stock in New Zealand is much smaller than that in the United States. The adjustment costs will enter into the financial decision because of the inter-temporal arbitrage between financial and physical assets. The marginal changes due to adjustment costs will tend to be equalized and thus more dollars of capital will flow into the countries with bigger capital stocks. This is a very important aspect of the model and captures some of the observed features of global capital markets well.

(vi) US Productivity Growth/Strong US Share Market

One of the interesting issues facing the world economy during the late 1990s and into 2000 has been whether the high US equity prices are a bubble or the result of some fundamental factors. An obvious related question is what would be the impact of a sharp decline in these prices on the US and world economies?

The first issue to understand is the rise in US equity markets. According to the models there are three key factors at work.

The first factor is a rise in global and US public-sector saving. An analysis of the implications of an unanticipated change in saving was set out in a theoretical paper by McKibbin and Wilcoxen (1998b). That paper pointed out that in a closed economy with adjustment costs in physical capital accumulation (but without the complication of wage stickiness and so with continual full employment of resources), a surprise rise in saving would not initially be matched by a sharp rise in physical investment. This is because firms are trying to minimize adjustment costs and to smooth investment over time. The first effect is, instead, a sharp rise in the value of existing capital. Asset prices jump sharply, yielding capital gains to owners of existing capital. They must jump enough to make saving fall back again, given the difficulty of raising investment. Gradually, firms are induced to respond to the saving increase by building more physical structures to transform current saving into future output. But it may take a long time for such desired increases in wealth to be effected. Thus, for example, a large unanticipated fiscal contraction, such as has been experienced in most OECD countries during the second half of the 1990s, will tend to push up equity prices sharply before there is much of a response in physical investment.

The second factor is that productivity growth in the US computer sector has risen sharply and is expected to continue to rise. In these models, such a future productivity gain causes a rise in the expected return to capital, a sharp rise in equity prices, and over time, a boom in investment. Through wealth effects, it also causes a rise in private consumption as forward-looking consumers attempt to smooth consumption by consuming some of the future income gains immediately. Thus a surge in economic activity develops in the short run. This is magnified by Keynesian features, as backward-looking firms and households experience a strong current economy. The initial response is driven by the responses of forward-looking households and firms to changes in asset market prices and the expected future real growth prospects: the presence of rule-of-thumb consumers and firms causes Keynesian multiplier effects on consumption and accelerator effects on investment which induce some overshooting in the real economy.

The two factors of a sharp rise in global saving and a rise in actual and expected productivity growth go a significant way towards explaining the US equity price rise. However, even with these factors, there is still a large rise in US equity prices from 1996 through to the end of 1999 to be explained. A third factor is also needed.

This third factor is a fall in the equity risk premium—that is the difference between the expected return on holding government bonds and on holding equity. In the model, this premium is actually calculated internally to the model as the wedge in the arbitrage equations between equity and bonds that is required to make the actual returns across asset classes and the model-generated returns in a given base year (1996) be equal (see Bagnoli et al., 1996, for a description). Subject to the variety of ways to measure it, the equity risk premium has fallen from around 8 per cent in the mid- to late 1980s to close to zero by 1999. Adding this to the model is sufficient to explain the change in the US equity market (see McKibbin and Wilcoxen, 2001). This effect works in the model through the inter-temporal arbitrage condition for Tobin’s q. The current value of q is calculated as the present value of expected future
increments to the physical marginal product of capital relative to the profile of future short-term real interest rates on government bonds, plus a risk premium: a fall in the risk premium thus directly pushes \( q \) upwards and encourages investment.

The change in the equity risk premium is the most important of the three factors with which the models explain the high levels of equity prices in the USA. Again, we can see that it is essential to have a model which contains all of inter-temporal optimization, adjustment costs, and wage stickiness, if the effects of these three factors are to be disentangled in the way just described. A model containing only a simple IS curve is unlikely to be good enough for us to be able to do this adequately.

Once the factors driving up US equity prices are explained in the context of the model, it is interesting to use the model for the counterfactual experiment of a rise in the equity risk premium back towards historical levels. This has been done in McKibbin and Stoeckel (1999). In this simulation, we assume the premium on equities that normally existed prior to the late 1980s returns partly back to that level, rising by 5 per cent overnight. The first thing that happens as a result of a return in the equity risk premium in the United States is that investors take their money out of equities and redistribute it over other assets, both internationally and within the United States. The consequence is that real interest rates drop in the United States as people buy government bonds. As foreign assets now look relatively more attractive, there is a large capital outflow. As a result there is a sharp depreciation of the US dollar until the point at which an expected re-appreciation of the dollar provides the required higher yields in the USA relative to abroad. This causes a large improvement in the trade balance, as exports rise while imports fall. Along with the capital outflow, there is a large initial drop in real investment in the United States, a fall in real consumption, and a fall in real GDP, relative to what it would otherwise have been. But there is not a recession in the USA. The reason is the stabilizing effects of lower real interest rates and a weaker US dollar, both of which dampen the decline in investment and stimulate exports. Thus the inter-temporal factors create automatic stabilizers that dampen the impact of an exogenous change in risk.

What impact does this deterioration in the United States economy have on other economies? There are two effects operating. One set of effects operates through the capital side and the other set works through the trade side. As demand in the USA weakens, countries that sell products to the United States experience a decline in their net exports and therefore experience a negative demand shock. On the other hand, the reallocation of global financial capital causes a fall in real interest rates globally, which stimulates net investment outside the USA. In a sense this part of the story is very similar to the effect of a fiscal contraction in the USA (which is the obverse of what was discussed above when we consider the effects of Reaganomics). What is going on is also similar in a way to the Asian crisis story, except that in Asia the rise in risk was on all assets in the Asian economies (imposed via the interest parity condition) whereas in this US stock-market scenario the rise in risk is only on US equities (imposed only in the arbitrage equation between bonds and equity within the United States). Thus, in the Asian-crisis simulations, agents substituted out of all assets within the Asian economies, and so the prices of all assets fell within these economies, including the price of government bonds. That is, interest rates rose. Thus in that case the rise in risk premia caused a rise in real interest rates in Asia, which caused a slow-down there, and, since the shock was large, the slow-down was severe. In the case of the shock to US equities, people substitute out of equities into other US assets as well as overseas. Bond prices rise in the USA. Thus there is a fall in real interest rates. As a result there is strong domestic dampening of the effects. The model clearly illustrates an important difference between the two types of confidence/risk shocks, depending on how widely the shock to confidence is distributed. It is important to have inter-temporal arbitrage relationships in a model, in order to be able to disentangle these different types of shocks to risk levels.

Perhaps the most unsatisfactory part of the analysis of risk shocks using the model is that the model does not explain these shocks. They simply involve exogenous changes to the arbitrage equations between various financial assets and between financial assets and physical capital stocks. None the less, the analysis does provide an insight into the adjustment of economies to a very common type of shock.
IV. STRENGTHS OF INTER-TEMPORAL MODELS WITH SOME STICKINESS

As outlined above, the MSG2 and G-Cubed models have been successfully used for a wide range of policy issues and scenario analysis. We have argued that a number of important insights have been acquired in using these models that would not be available from models which did not combine inter-temporal optimization and stickiness in the way that these models do. Here we briefly review the relative advantage of these models, within the context of debates about the usefulness of macroeconomic models in general.

(i) Big versus Small

Krugman’s paper in this issue reviews the effect of the ‘modern macroeconomic purists’ who have controlled most of macroeconomics for the last 25 years. According to these writers, a model cannot, and must not, be taken seriously unless it has well-articulated, and fully specified, micro-foundations. The approach of many researchers in the real-business-cycle tradition has been to take this requirement absolutely seriously, and to build simple models which meet it, with characteristics that can be used to address a specific issue. The trouble with this tradition is that the models usually have to be contorted to fit empirical data, essentially because a number of features that need to be taken into account for a model to track macroeconomic data are ignored in the process of simplification. These are one-issue models that are not much use for policy purposes.

There is now also a new policy-related macroeconomics, producing small models for the analysis of monetary policy in small open economies (see the paper in this volume by Taylor and the references which it contains). Central banks in countries such as Australia have developed small models of this kind in recent years. The small number of variables in these models reflects the recent focus of central banks on inflation targeting and inflation forecasting rather than being concerned with a wide range of variables. Unfortunately, the danger with this tradition is that by not considering the general equilibrium nature of economies—in particular the inter-temporal issues on which we have focused in this paper—critical mistakes may be made with monetary policy in misunderstanding the source and implications of a shock.

One core argument for smallness used to be that smallness makes the results easier to understand. Interpretability is now a necessary condition for a model: nobody any longer defends black-box policy models. However, a model does not need to be small to be easy to use or to understand. The critical issue is the extent to which the model is constrained by economic theory, and thus has properties which can be understood through theoretical insights. Even the outputs of very large models can be understood in this way: one of the advantages of very large CGE models such as the GTAP model (see Hertel, 1997) is that they can be relatively easy to understand because of their reliance on a clear theoretical structure. Thus, the analysis of the impact of a policy change on literally thousands of individual commodities or households can be understood through the effects on relative prices and incomes. The issue of size need not be a real constraint on intelligibility.

A second argument for smallness used to be the difficulty of solving large models without this requiring hours of computer time per run. That was particularly true once the use of rational expectations became a working hypothesis. This increased the degree of numerical complexity, and initially rational expectations models were restricted to a few equations purely because this was all that could be solved. Over the past two decades, the degree of technical innovation in this area has almost completely removed this constraint on model builders.

By most criteria, the MSG2 model is a large model which was initially challenging to solve. But there is only one sector in each country, and the model of each country essentially boils down to only a small number of behavioural equations (for production, consumption, investment, exports, imports, and wages), plus equations for the determination of asset prices, and, although some parameters differ across countries, the behaviour of all countries in the model is generically similar. Thus, the model is relatively easy to understand. It is also true that the model contains a large number of forward-looking or ‘jumping’ variables. (For each country there are jumping variables for human capital, Tobin’s q, wages, and the exchange rate; thus with 18 coun-
tries there are over 70 such variables.) In the mid-1980s, when the model was being developed, most other modelling solution systems could cope with only two or three jumping variables. But the numerical solution technique used to solve this model (McKibbin and Sachs, 1991, Appendix C) enables this whole system to be solved relatively easily.

The G-Cubed model is a larger model still, with over 8,000 equations and 150 jumping variables. However, like the MSG2 model, it is relatively easy to understand, as a result of having a clear theoretical structure and tight theoretical restrictions.

Thus we are now able to combine the imposition of inter-temporal budget constraints and rational expectations with enough stickiness to track the data, in a large system describing the interaction of different countries. And we can even, as in G-Cubed, add sectoral detail and still end up with a model whose properties are easy to understand, which is also straightforward to solve.

(ii) Structural versus Reduced Form

The big versus small debate is partly related to the issue of structural versus reduced-form models. In structural models there is a clear economic structure with behaviour based on maximizing objective functions of agents (such as utility functions or profit functions) subject to budget constraints (either within a period, or inter-temporal). In structural models, deep parameters of the utility function or technology can be more easily identified, something which is a precondition for the kind of the interpretability discussed in the previous section. One of the problems with VAR models, for example, is that it is never clear exactly how to implement a particular policy rule or shock. Usually, a policy change is defined as an innovation in one or more of the model’s equations, but it is never clear what this actually means (see Sims, 1988). Furthermore, structural models can be made less prone to the ‘Lucas Critique’ (Lucas, 1973). It has been an important advance that, with models such as those discussed in this paper, we can now analyse a range of policy regimes without worrying about parameter endogeneity (see Ericsson and Irons, 1995).

G-Cubed and MSG2 are structural models with a great deal of detail. Behaviour is traced back to the underlying deep structural parameters, such as tastes and technology, and to inter-temporal optimizing behaviour. Enough examples have been given in the previous section to show how difficult it is to get good analysis of policy rules and shocks without this structure. This issue will be thrown into even sharper relief in the next section when we explicitly consider the difference in the short-run effects of temporary and permanent shocks—differences which cannot be analysed at all in backward-looking reduced-form models. In addition, institutional details are imposed in the models, such as through the operation of monetary policy, and in the structure of the labour markets in different economies. Further institutional details appear in the mix of short-run inter-temporal-optimizing behaviour and rule-of-thumb behaviour which is assumed for both consumers and producers. To a purist, these institutional features of the models might not have a clear theoretical basis. However, extensive empirical investigation—the outcomes of which were described in the historical discussions in the previous section—has shown that these features are essential if the models are to replicate actual historical experience.12

(iii) Theoretical versus Data Intensive

There remains a vast amount of disagreement on the extent to which theoretical restrictions should be imposed on a model and the extent to which the data should be ‘allowed to speak for themselves’. Computable general equilibrium (CGE) models tend to be based on simple theoretical relations with data used to calibrate key ratios, but little in the way of econometric estimation.13 Other ways of producing calibrated models can be found in the real-business-cycle literature, in which the authors tend to choose parameters based on generating variances and covariances consistent with a given data set. Structural models in the macroeconometric modelling literature tend to contain equations obtained from

---

12 There are arguments that could be made in support of the approach in MSG2 and G-Cubed, such as the existence of monitoring costs which prevent agents from re-optimizing and gathering information continually. Although this might seem a fairly weak straw to clutch for those who insist on an optimizing reason for all observed behaviour, it is the best that we have at present. (See Mankiw (2000) for a discussion of this general issue in the particular context of why wages are sticky in the short run.)

13 Those of Dale Jorgenson and his colleagues at Harvard are an exception to this rule.
time-series estimation of equations, but with the imposition of theoretical restrictions where these are accepted by data (and even sometimes when they are not). VAR models are at the other end of this spectrum.

The G-Cubed model stands in the middle of this spectrum. It uses econometric estimation techniques to estimate the key substitution elasticities in production and consumption using a time series of US input–output data. It is then assumed that all countries have the same shaped isoquants and indifference curves, but they are assumed to have different initial levels of technology in each sector and different initial endowments of inputs, and so different relative factor prices, and so different relative goods prices. The dynamics of the model’s adjustment over time are partly driven by the dynamics of the wage equation, which is based on estimation by researchers outside the modelling group. Also important are assumptions about the structure and size of adjustment costs in the installation of capital, about the parameters of consumers’ utility functions, and about the extent to which consumers and investors face liquidity constraints; all of these assumptions are based on the relevant empirical literature.

It is possible to use the technique of McKibbin et al. (1998) to estimate an error-correction model around the theoretical specification of the G-Cubed model, in order to get a better dynamic fit of the model, but this has not been done yet for the G-Cubed model. The dynamic fit of the G-Cubed model is one area where the model is less well suited than other models for short-term forecasting purposes, although it does appear to do reasonably well over horizons of several years, primarily because the calibration of adjustment costs, and of the shares of optimizing and rule-of-thumb agents, is performed with attention to the dynamic properties of the model.

It is critical in the approach taken in the MSG2 and G-Cubed models that the theoretical restrictions are imposed so that the mechanisms of adjustment can be understood clearly. However, in many cases, the quantitative story will be driven by the relative size of parameters and, therefore, in this modelling approach considerable attention is required to best quantify the parameters in the model. Thus both theoretical restrictions, including the imposition of inter-temporal accounting, as well as a significant attention to the magnitudes of parameters, are a critical part of the modelling strategy behind the models.

(iv) Forward-looking versus Autoregressive

The debate on the formation of expectations is an active one, and no consensus has been reached among modellers. Nevertheless, the use of rational expectations in large-scale models is widespread. McCallum (1994) makes a strong argument in favour of rational expectations as an operating assumption in economy-wide models. In most models, proprietors have an option to switch between alternative expectations assumptions for sensitivity analysis. In the G-Cubed model, we have a mix of assumptions about expectations. All agents are forward-looking in asset markets. A share of consumers and producers are forward looking (usually 30 per cent). These forward-looking agents are assumed to be rational in the sense that they use the model for their forecast of the future. The remaining agents use a steady-state rule of thumb for decisions. Similarly, wage inflation is partly forward looking and partly backward looking. These rules of thumb are rational in the steady state. But in the short run, these agents do not update the future variables in their information set, although they do respond to changes in contemporaneous variables. Such an assumption can be interpreted in a number of ways. It could be that there are two types of agents in the model—those that are rational and those that are not—or it can be interpreted as being a single agent who has partially rational expectations. Or it could be that the rule-of-thumb agents would like to act rationally if they could, but instead their behaviour is pinned down by liquidity constraints. These assumptions on expectations mean that asset prices tend to be more volatile, because they need to move to sustain equilibrium in the model when there is a good deal of persistence in the model from sluggish adjustment of non-optimizing agents.

As we have seen, this persistence, and the volatility which it generates, contributes much to our understanding of a number of issues.

A clear additional advantage of applying rational expectations in a structural model such as MSG2 and G-Cubed is that one can analyse anticipated shocks or policy announcements. It is also possible to analyse the impact of a policy announcement that
is believed in one period but then not implemented in subsequent periods.

There seems to be a perception that the use of rational expectations in a model always rules out agents making errors in their expectations formation. We can deal with this vast range of interesting issues by changing the information sets of agents over time. Indeed, in chapter 5 of McKibbin and Sachs (1991) we illustrate how to use this type of large-scale model with rational expectations to track the world economy from 1980 to 1990, using new information sets each year. Essentially this involves solving the model given previous decisions and a given information set, rolling the model solution forward by one period, and then solving again, given a new information set. Being able to do this was certainly important in our understanding of the Asian crisis. A recent paper justifying adaptive expectations as a basis for a dynamic version of GTAP (Ianchovichina et al., 1999) criticized the approach of the G-Cubed model for analysing the Asian crisis, on the argument that the agents that invested in Asia were clearly irrational before 1996 because they pulled their money out of Asia in 1997. But perhaps what the events in Asia more plausibly indicate is that the information sets of agents changed quickly and dramatically, and that this is precisely why so much money was pulled out. Indeed, as we have argued above, a rational expectations model with variable risk premia is a very good model for analysing this particular issue. In our view, an adaptive expectations model could not explain the Asian crisis or the subsequent adjustments.

(v) Static versus Dynamic

There is a great deal of policy analysis, particularly that using CGE models, which is of a comparative-static kind. CGE models are calibrated to a particular year, a change in policy is imposed, and the new hypothetical equilibrium for that year is calculated. The model then shows the longer-term impacts of policy, with a primary focus on the effect of the policy on the efficiency of allocation of resources across the economy. A static analysis might be reasonable for some policies, such as tariff changes or tax policy changes, in order to evaluate the longer-term effects of the policy. However, a model such as G-Cubed, with a detailed dynamic profile, enables one to see quickly why the adjustment path might be just as important as the long-term equilibrium effects of a policy. There are two key issues.

The first issue concerns the short-run impact of a policy. A model such as G-Cubed shows that this may often be of opposite sign to that of the longer-term outcome. An example of this is a tariff reduction. In the long run this stimulates exports, but in the short run can cause a large rise in the return to capital in the export sector which induces a capital inflow and causes an appreciation of the exchange rate. The appreciation of the exchange rate can reduce exports in the short run, until the additional productive capacity is in place. Only then do exports expand.14

The second issue is one of political economy. The adjustment path may be important in determining whether the longer-term outcome can be reached. In McKibbin (1998a) it was shown that trade liberalization can have an effect on the GDP of the liberalizing country which is negative for several years, but which becomes positive in the longer run. It was also shown that trade liberalization in foreign countries tends to have positive effects on home GDP in the short run, but that in the long term its contribution to GDP is small, relative to the gains from home liberalization. Thus coordinated trade liberalization could offset the short-run cost of home liberalization. This could help policy-makers to seek out the longer-term gains from liberalization which CGE models repeatedly find in the trade literature. In all evaluations of policy changes like this, understanding the dynamic path of adjustment is crucial. This is because policy-makers may become disaffected with the policy change (here liberalization) if the resulting short-run economic adjustments are perceived to be politically costly, or have effects which are misunderstood, in the period before the benefits come through. That may lead to pressure for counterproductive policy changes. For example, large capital inflows in response to trade liberalization announcements could cause significant deterioration in the current account which, if misunderstood, could cause a counterproductive change in macroeconomic policy.

14 See the discussion above of the effects of NAFTA.
It is clear where the G-Cubed and MSG2 models are located in the debates outlined above. The next section will illustrate some areas where a large structural model such as G-Cubed makes a contribution to our understanding of some currently active puzzles in the theoretical macroeconomics literature.

V. EXPLORING SIX PUZZLES IN INTERNATIONAL MACROECONOMICS

In a recent paper, Obstfeld and Rogoff (2000) introduced six puzzles that have permeated the theoretical literature in international macroeconomics and asked how they might be solved. Those six puzzles are: (i) the bias in trade towards consuming home goods; (ii) the own-country bias in ownership of financial assets; (iii) the Feldstein–Horioka result that there is a high correlation between national saving and national investment spending; (iv) the international consumption-correlations puzzle—the low correlation between growth in consumption across countries—which is also expressed as the puzzle that output growth seems to be more highly correlated than consumption growth across countries; (v) the apparent breakdown of purchasing power parity (PPP) in the short to medium term or the persistence of changes in real exchange rates; and (vi) the ‘exchange rate disconnect puzzle’—shown by the apparent disconnect between exchange rates and underlying macroeconomic variables. This section shows that the focus of the models which we have described here—with an emphasis on forward-lookingness, adjustment costs and liquidity constraints, and wage stickiness—can help us to explain four out of these six puzzles in a way that complements the insights of Obstfeld and Rogoff.

The issues raised by the first two of the puzzles have little to do with the main questions addressed in this paper and these can be quickly dealt with. First, it is worth noting that the G-Cubed model directly addresses the empirical evidence that agents within countries tend to consume a large proportion of goods produced within their national borders, relative to the consumption of imports. This is done by making the Armington (1969) assumption (for which there is very good empirical evidence) that goods from different countries are different goods, and then simply calibrating the model so that the initial shares of goods in the consumption bundle are equal to the actual shares in the data. The model does this directly as well as building in the critical insight of Obstfeld and Rogoff’s paper, namely that there are costs (tariffs, transport costs, etc.) in shipping goods between countries. Second, we note that the G-Cubed model does not have much to say about the home equity bias because there is assumed to be complete arbitrage between alternative financial assets, equalizing the expected returns after adjustment by an exogenous risk premium. Households therefore only hold own-country assets, except for foreign debt. Home bias in asset holdings is assumed in the structure.

To illustrate how a model such as G-Cubed performs in relation to the remaining four puzzles, we now consider results for a temporary rise and a permanent rise in total factor productivity (TFP) in the home country. We use the symmetric prototype two-country model outlined in the Appendix rather than the full G-Cubed model in order to simplify the analysis. These results are contained in Figures 1 and 2 for a subset of variables that are contained in the prototype model. All results are presented as deviations from the baseline of the model (either percentage, per cent of GDP, or percentage point). Thus, a zero in each figure would indicate no change relative to baseline.

(i) The Feldstein–Horioka Puzzle and the Consumption-correlations Puzzle

Feldstein and Horioka (1980), and numerous authors subsequently, pointed to the high correlation between saving and investment rates within countries, suggesting that international capital was not as mobile internationally as one might expect. Because of the close correlation between saving and investment, the difference, which is the current account of a country, does not change by very much over time. This is in contrast with the implications of high

---

15 What the model does not do is explain the degree of observed home biases by means of the measured levels of tariffs, transport costs, etc.
Figure 1
Temporary Rise in Home Country TFP

W. J. McKibbin and D. Vines
Figure 2
Permanent Rise in Home Country TFP

Real Output and Consumption: Home Country

Saving and Investment: Home Country

Real Private Consumption: Home and Foreign Countries

Real Output: Home and Foreign Countries

Current Account and Trade Balance: Home Country

Real and Nominal Exchange Rate: Home Country

Wages and producer prices: Home Country

Short and Long Term Interest Rates: Home Country
capital mobility in simple theoretical models, which predicts much larger changes in current accounts than those which are observed.

Consider what happens to current accounts in the G-Cubed model in response to a temporary productivity shock. Refer to Figure 1. The results show that the responses of saving and investment to such a shock are very highly correlated with each other. That initially seems surprising. One might expect that, in an inter-temporal model such as G-Cubed, a temporary productivity shock would lead to consumption smoothing, and that, with perfect capital mobility, this would happen through the acquisition of foreign assets at the given world rate of interest, rather than through the accumulation of domestic capital which is subject to diminishing returns. Costs-of-adjustment considerations would also suggest that, because the productivity shock is temporary, inter-temporally optimizing firms facing adjustment costs would be unwilling to put in place capital which is costly to install. This provides an additional reason why consumption smoothing might be expected to happen via the acquisition of foreign assets through a balance-of-payments surplus, rather than through the acquisition of extra physical assets at home.16

The main reason that this does not happen in the model is the imperfect substitution between home and foreign goods—the ‘Armington assumption’ which we have already discussed. The attempt to consumption-smooth, by saving a large proportion of a temporarily higher level of income, means, as a result of wage stickiness, that demand will tend to fall below output. As a result, output prices fall, and there is also a fall in the short-term nominal, and real, interest rate. Consequently, the real exchange rate depreciates.17 This real currency depreciation is the mechanism by which the required current-account improvement is effected; it is also what puts a floor under the extent to which the domestic interest rate falls. But the fall in domestic interest rates that does happen will lead to a rise in the value of existing capital. This valuation effect and the fall in the real interest rate both raise the relative price of future consumption relative to current consumption, thus raising current consumption and dampening the increase in savings. The interest-rate and exchange-rate responses will also serve to stimulate domestic investment: currency depreciation will raise profitability and so increase investment by backward-looking firms,18 and even optimizing firms facing adjustment costs will respond positively to the valuation effect to some extent. Thus the interest-rate and exchange-rate responses just described will both dampen the required amount of consumption smoothing and, at the same time, promote domestic investment as a means of accommodating the savings which results. This reason is remarkably straightforward, and is essentially what Obstfeld and Rogoff have re-discovered in the form of costs and rigidities of trading goods. The less substitutable are home and foreign goods for each other, the more will interest rates and the exchange rate have to fall and so the more of the increase in saving that will feed through to an increase in domestic investment in this way. Experiment with the model confirms this explanation. When the elasticities of substitution between domestic and foreign goods are raised in the model, the gap between investment and saving increases.

In the case of the permanent productivity shock, the correlation between investment and saving is also high (see Figure 2). Again this is not what might have been expected.

In a perfectly flexible inter-temporal model, one would expect that an increase in productivity would lead to an immediate increase in output. Investment would immediately surge and then gradually fall back, because improved productivity would lead to a rise in the desired stock of capital. As a result, the increase in output would become even larger in the longer term. Consumption would immediately jump upwards, in line with the improved long-run prospects for consumers. Although, in the long run, enough would need to be saved to pay for investment in a larger stock of capital, this would be gradual, and saving might even actually fall in the

16 See the discussion at the end of the section on the Asian crisis above where it is shown that the adjustment costs are minimized if any demand for extra physical capital is spread around the world rather than concentrated in a single country.

17 The reason that there is not an appreciable depreciation in the nominal exchange rate is discussed in section V(ii) below.

18 Such firms will also be increasing their investment directly as a result of the higher profitability which follows immediately from the productivity shock.
short run (since long-run prospects for consumers would have increased by more than the immediate increase in output). Savings and investment would certainly not be highly correlated, and the current-account would initially go into deficit.

What actually happens in the model is that output, and investment, go up as expected. But everything else is rather different. Although consumption does rise, backward-looking households raise their savings in response to higher incomes and output, leading to a shortfall in aggregate demand. Wages and prices fall (although in the long run, of course, the real wage will rise.) But they do not fall enough to soak up all of the increase in production. Nominal interest rates fall sharply and the real and nominal exchange rate depreciates sharply. The current account actually improves, both because domestic absorption has risen by less than domestic output and because competitiveness has improved.

There is a well-documented result that consumption growth across countries has lower correlation than would be expected in a consumption-smoothing world. This is not unrelated to the Feldstein–Horioka puzzle just discussed. In addition Backus et al. (1992) observed that internationally, output of different countries was more highly correlated than consumption across countries.

Again, refer to the results for G-Cubed in Figure 1 for the temporary shock to TFP in the home country. These results illustrate that consumption rises sharply in the home country, with very little smoothing internationally or inter-temporally. We have already discussed why this happens. In this temporary-shock case, even though the home country is large, very little happens to consumption or output in the foreign country because the short-run adjustment process is mainly bottlenecked in the home country, for reasons already discussed extensively. For a permanent shock (Figure 2), there is an effect on foreign output, essentially because of the terms-of-trade effect which results when a richer home economy buys more foreign goods. The correlation between consumption in the two countries now parallels that between output in the two countries, as would be expected in a model with no stickiness.

Nevertheless, in the short run, home consumption rises by less than in that benchmark case, for reasons which have already been explained.

(ii) The Exchange-rate Puzzles

The failure of short-run PPP is well documented. In the G-Cubed model, nominal wages are sticky, even though prices are assumed to clear the goods markets annually. As noted by Obstfeld and Rogoff, once you introduce sticky nominal variables, you can quickly generate the observed break-down of PPP. As an illustration of how important and persistent this is in the G-Cubed model, results for real and nominal exchange rates for the temporary and permanent TFP shocks are shown in Figures 1 and 2 respectively.

When there is a permanent change in productivity, the nominal and real exchange rate changes immediately, something which is obviously unsurprising to advocates of PPP. It is what happens in response to a temporary shock to TFP which is more surprising. As a result of such a shock, the real exchange rate is actually more volatile than the nominal exchange rate over several years. When productivity increases, prices in the home country fall and, with fixed nominal wages, this causes a sharp depreciation of the real exchange rate in the year of the shock. However, the nominal exchange rate only changes slightly. That is because, although the fall in goods prices increases the real quantity of money and lowers the short-term nominal interest rate, prices are expected to increase in subsequent periods, thereby creating expectations of increases in interest rates. The initial value of the nominal exchange rate only changes slightly. That is because, although the fall in goods prices increases the real quantity of money and lowers the short-term nominal interest rate, prices are expected to increase in subsequent periods, thereby creating expectations of increases in interest rates. The initial value of the nominal exchange rate is the integral of the expected interest-rate differentials between home and foreign bonds, and, averaging over the short term and the long term, these tend to cancel out.

The divergences from PPP which we have shown, or the movements in the real exchange rate, are large and persistent for the real shocks which we have considered. In the case of nominal shocks (not shown) there are, of course, also short-run deviations from PPP, due to Dornbusch-type overshooting effects. These tend to disappear fairly rapidly,

19 After the productivity shock is over, aggregate demand remains temporarily above base, creating excess demand, and forcing prices to increase.
meaning that persistent deviation from PPP is much less in response to these shocks.

The second exchange-rate issue is the apparent disconnect between exchange rates and fundamentals. But there is a problem of how actually to estimate a robust relationship between exchange rates and fundamentals which is well illustrated by the model results discussed in this paper. The relationship between these variables depends very importantly on the nature of the shocks to the system. As well as the real versus nominal distinction, and the temporary versus permanent distinction, we could have run anticipated versus unanticipated shocks. In some cases results have opposite signs in the period before the shock occurs. Unless the underlying shocks can be identified correctly, it is difficult to see how reduced-form econometric exercises, of the types that are manifest in the exchange-rate literature, will be able properly to identify the relationship between exchange rates and other macroeconomic variables. One obvious exercise would be to run a large batch of stochastic simulations using the G-Cubed model and then to estimate some reduced-form relationships from these. This would enable one to see how robust the standard econometric techniques generally applied in the exchange-rate literature actually are.

VI. CONCLUSION

This paper has given an overview of the nature of the MSG2 and G-Cubed models. It has made a case for using these large structural models for policy analysis and has summarized the insights from the models for a range of shocks experienced during the 1980s and 1990s. Simple macroeconomic models have a difficult time explaining many aspects of these shocks. We have argued that imposing inter-temporal budget constraints, incorporating their role in determining asset prices, and combining this with short-term stickiness—whether in price adjustment, or in adjustment costs in investment, or in rule-of-thumb behaviour by consumers and producers—is a powerful way of doing what is necessary to build a real-world policy model. Although for some policy issues and the analysis of some shocks, simple models may be both necessary and sufficient, we believe that there is also a need for large-scale models, which pull together theory and empirical relationships in a coherent way.

The last part of the paper focused on the issues raised by Obstfeld and Rogoff (2000). In our view, that paper contains an eloquent, and elegant, plea for something which has been happening in the applied economic modelling literature for more than a decade. The authors argue that we have a lot of relevant theory—to do with inter-temporal features and with the role of asset markets—and that this needs to be put together in a coherent and internally consistent fashion, if we are to create policy relevant models that actually fit the data. But as the earlier part of this paper makes clear, advances in computer technology and numerical algorithms have made this feasible, and the process of creating these models has already been under way for a long time. We can now build inter-temporal general equilibrium models, containing many of the core features of Obstfeld and Rogoff’s (1996) classic textbook, but also containing sticky wages, and adjustment costs in capital accumulation, and rule-of-thumb behaviour by consumers and producers. And we can make these models large enough to enable simultaneous treatment of a large number of countries and sectors. Such models are already being used for policy analysis. As we argued in section V, they enable us to explain many of Obstfeld and Rogoff’s ‘puzzles’, as well as improving our understanding of the world economy.
APPENDIX: A STYLIZED TWO-COUNTRY G-CUBED MODEL

In this section a stylized two-country model is presented which distills the essence of the G-Cubed model and, in particular, how the inter-temporal aspects of the model are handled. The reader is referred to chapters 2 and 5 of McKibbin and Wilcoxen (2001) for greater detail.

In this stylized model there are two symmetric countries (based essentially on US data adjusted to create symmetry). Each country consists of several economic agents: households, the government, the financial sector, and two firms, one each in the two production sectors. The two sectors of production are energy and non-energy (this is much like the aggregate structure of the MSG2 model). The following gives an overview of the theoretical structure of the model by describing the decisions facing these agents in one of these countries. Throughout the discussion all quantity variables will be normalized by the economy’s endowment of effective labour units. Thus, the model’s long-run steady state will represent an economy in a balanced growth equilibrium.

Firms

We assume that each of the two sectors can be represented by a price-taking firm which chooses variable inputs and its level of investment in order to maximize its stock-market value. Each firm’s production technology is represented by a constant elasticity of substitution (CES) function. Output is a function of capital, labour, energy, and materials:

\[ Q_i = A_i^{o} \left( \sum_{j=k,l,r,m} (\delta_{ij})^{1/\sigma_{ij}} X_i^{(\sigma_{ij}^{-1})/\sigma_{ij}} \right)^{\sigma_{ij}^{o} / (\sigma_{ij} - 1)} \] (1)

where \( Q_i \) is the output of industry \( i \), \( x_i \) is industry \( i \)’s use of input \( j \), and \( A_i^{o}, \delta_{ij}, \sigma_{ij}^{o} \) are parameters. \( A_i^{o} \) reflects the level of technology, \( \delta_{ij}^{o} \) is the elasticity of substitution, and the \( \delta_{ij}^{o} \) parameters reflect the weights of different inputs in production; the superscript \( o \) indicates that the parameters apply to the top, or ‘output’, tier. Without loss of generality, we constrain the \( \delta_{ij}^{o} \)’s to sum to one.

20 This approach follows Armington (1969).
21 This does not require that both sectors purchase the same amount of oil, or even that they purchase oil at all; only that they both feel the same way about the origins of oil they buy.

The goods and services purchased by firms are, in turn, aggregates of imported and domestic commodities which are taken to be imperfect substitutes. We assume that all agents in the economy have identical preferences over foreign and domestic varieties of each commodity. We represent these preferences by defining composite commodities that are produced from imported and domestic goods. Each of these commodities, \( Y_{ci} \) is a CES function of inputs of domestic output, \( Q_i \), and an aggregate of goods imported from all of the country’s trading partners, \( M_i \):

\[ y_i = A_i^{o} \left( \left( \delta_{ij}^{o} \right)^{1/\sigma_{ij}^{o}} Q_i^{(\sigma_{ij}^{o} - 1)/\sigma_{ij}^{o}} + \left( \delta_{ij}^{o} \right)^{1/\sigma_{ij}^{o}} M_i^{(\sigma_{ij}^{o} - 1)/\sigma_{ij}^{o}} \right)^{\sigma_{ij}^{o} / (\sigma_{ij}^{o} - 1)} \] (2)

where \( \sigma_{ij}^{o} \) is the elasticity of substitution between domestic and foreign goods. For example, the energy products purchased by agents in the model are a composite of imported and domestic energy. The aggregate imported good, \( M_i \), is itself a CES composite of imports from individual countries, \( M_{ic}^{o} \), where \( c \) is an index indicating the country of origin:

\[ M_i = A_i^{o} \left( \sum_{c=1}^{C} (\delta_{ic}^{o})^{1/\sigma_{ic}^{o}} M_{ic}^{(\sigma_{ic}^{o} - 1)/\sigma_{ic}^{o}} \right)^{\sigma_{ic}^{o} / (\sigma_{ic}^{o} - 1)} \] (3)

The elasticity of substitution between imports from different countries is \( \sigma_{ic}^{o} \).

By constraining all agents in the model to have the same preferences over the origin of goods we require that, for example, the agricultural and service sectors have the identical preferences over domestic oil and imported oil. This accords with the input–output data we use and allows a very convenient nesting of production, investment, and consumption decisions.

In each sector the capital stock changes according to the rate of fixed capital formation \( (J_i) \) and the rate of geometric depreciation \( (\delta) \):

\[ \dot{k}_i = J_i - \delta k_i \] (4)

Following the cost-of-adjustment models of Lucas (1967), Treadway (1969), and Uzawa (1969), we

130
assume that the investment process is subject to rising marginal costs of installation. To formalize this we adopt Uzawa’s approach by assuming that in order to install \(J\) units of capital a firm must buy a larger quantity, \(I\), that depends on its rate of investment \(\frac{J}{k}\):

\[
I_i = \left(1 + \frac{\phi_i}{2} \frac{J}{k_i} \right) J_i
\]

(5)

where \(\phi_i\) is a non-negative parameter. The difference between \(J\) and \(I\) may be interpreted in various ways; we will view it as installation services provided by the capital-goods vendor. Differences in the sector-specificity of capital in different industries will lead to differences in the value of \(\phi_i\).

The goal of each firm is to choose its investment and inputs of labour, materials, and energy to maximize inter-temporal net-of-tax profits. For analytical tractability, we assume that this problem is deterministic (equivalently, the firm could be assumed to believe its estimates of future variables with subjective certainty). Thus, the firm will maximize:

\[
\int_0^\infty (\pi_i - (1 - \tau_4) p^I I_i) e^{-(R(s)-n)(s-t)} ds
\]

(6)

where all variables are implicitly subscripted by time. The firm’s profits, \(\pi_i\), are given by:

\[
\pi_i = (1 - \tau_2) (p^1 Q_i - w_x x_d - p_x^1 x_{ed} - p_m^1 x_{em})
\]

(7)

where \(\tau_2\) is the corporate income tax, \(\tau_4\) is an investment tax credit, and \(p^1\) is the producer price of the firm’s output. \(R(s)\) is the long-term interest rate between periods \(t\) and \(s\):

\[
R(s) = \frac{1}{s-t} \int_t^s r(v) dv
\]

(8)

Because all real variables are normalized by the economy’s endowment of effective labour units, profits are discounted adjusting for the rate of growth of population plus productivity growth, \(n\). Solving the top-tier optimization problem gives the following equations characterizing the firm’s behaviour:

\[
x^o_{ij} = \delta^o_{ij} (A^o)^{o-1} Q \left( \frac{p^o_i}{p^o_j} \right)^{o-1} j \in \{l, e, m\}
\]

(9)

\[
\lambda_i = (1 + \phi_i \frac{J}{k_i}) (1 - \tau_4) p^I
\]

(10)

\[
\frac{d\lambda_i}{ds} = (r + \delta) \lambda_i - (1 - \tau_2) p^I \frac{dQ_i}{dk_i}
\]

(11)

where \(\lambda_i\) is the shadow value of an additional unit of investment in industry \(i\).

Equation (9) gives the firm’s factor demands for labour, energy, and materials, and equations (10) and (11) describe the optimal evolution of the capital stock. Integrating (11) along the optimum trajectory of investment and capital accumulation, \((J(t), k(t))\), gives the following expression for \(\lambda_i\):

\[
\lambda_i(t) = \int_t^\infty (1 - \tau_2) p^I \left( \frac{dQ_i}{dk_i} \right) ds
\]

(12)

Thus, \(\lambda_i\) is equal to the present value of the after-tax marginal product of capital in production (the first term in the integral) plus the savings in subsequent adjustment costs it generates. It is related to \(q_i\), the after-tax marginal version of Tobin’s q (Abel, 1979), as follows:

\[
q_i = \frac{\lambda_i}{(1 - \tau_4) p^I}
\]

(13)

Thus we can rewrite (10) as:

\[
\frac{J}{k_i} = \frac{1}{\phi_i} (q_i - 1)
\]

(14)

Inserting this into (5) gives total purchases of new capital goods:

\[
I_i = \frac{1}{2\phi_i} (q^2_i - 1) k_i
\]

(15)

\[\text{22 The rate of growth of the economy’s endowment of effective labour units, } n, \text{ appears in the discount factor because the quantity and value variables in the model have been scaled by the number of effective labour units. These variables must be multiplied by } \exp(nt) \text{ to convert them back to their original form.}\]
Based on Hayashi (1979), who showed that actual investment seems to be partly driven by cash flows, we modify (3) by writing $I_i$ as a function not only of $q_i$, but also of the firm’s current cash flow at time $t$, $\pi_i$, adjusted for the investment tax credit:

$$I_i = \alpha_2 \frac{1}{2\phi_i} (q_i^2 - 1)k_i + (1 - \alpha_2) \frac{\pi_i}{(1 - \tau_i)} p_i$$  \hspace{1cm} (16)

This improves the model’s ability to mimic historical data and is consistent with the existence of firms that are unable to borrow and therefore invest purely out of retained earnings.

So far we have described the demand for investment goods by each sector. Investment goods are supplied, in turn, by a third industry that combines labour and the outputs of other industries to produce raw capital goods. We assume that this firm faces an optimization problem identical to those of the other two industries: it has a nested CES production function, uses inputs of capital, labour, energy, and materials in the top tier, incurs adjustment costs when changing its capital stock, and earns zero profits. The key difference between it and the other sectors is that we use the investment column of the input–output table to estimate its production parameters.

**Households**

Households have three distinct activities in the model: they supply labour, they save, and they consume goods and services. Within each region we assume household behaviour can be modelled by a representative agent with an inter-temporal utility function of the form:

$$U_t = \int t \left( \ln c(s) + \ln g(s) \right) e^{-\theta(s-t)} ds$$  \hspace{1cm} (17)

where $c(s)$ is the household’s aggregate consumption of goods and services at time $s$, $g(s)$ is government consumption at $s$, which we take to be a measure of public goods provided, and $\theta$ is the rate of time preference.\(^{23}\) The household maximizes (1) subject to the constraint that the present value of consumption be equal to the sum of human wealth, $H$, and initial financial assets, $F$:\(^{24}\)

$$\int t\int p^c(s)c(s)e^{-(R(s)-\eta(s-t))} = H_t + F_t$$  \hspace{1cm} (18)

Human wealth is defined as the expected present value of the future stream of after-tax labour income plus transfers:

$$H_t = \int t\int (1 - \tau_i)(W(L_i^C + L_i^C + L_i^L) + \sum_{i=1}^{12} L_i^C) + TR) e^{-(R(s)-\eta(s-t))} ds$$  \hspace{1cm} (19)

where $\tau$ is the tax rate on labour income, $TR$ is the level of government transfers, $L_i^C$ is the quantity of labour used directly in final consumption, $L_i^L$ is labour used in producing the investment good, $L_i^C$ is government employment, and $L_i^L$ is employment in sector $i$. Financial wealth is the sum of real money balances, $MON/P$, real government bonds in the hand of the public, $B$, net holding of claims against foreign residents, $A$, the value of capital in each sector:

$$F = \frac{MON}{p} + B + A + q^l k^l + q^c k^c + \sum_{i=1}^{12} q^l k^l.$$

Solving this maximization problem gives the familiar result that aggregate consumption spending is equal to a constant proportion of private wealth, where private wealth is defined as financial wealth plus human wealth:

$$p^c c = \theta(F + H).$$  \hspace{1cm} (21)

However, based on the evidence cited by Campbell and Mankiw (1990) and Hayashi (1982) we assume some consumers are liquidity-constrained and consume a fixed fraction $\gamma$ of their after-tax income (INC).\(^{25}\) Denoting the share of consumers who are

---

\(^{23}\) This specification imposes the restriction that household decisions on the allocations of expenditure among different goods at different points in time be separable.

\(^{24}\) As before, $n$ appears in (1) because the model’s scaled variables must be converted back to their original basis.

\(^{25}\) There has been considerable debate about the empirical validity of the permanent-income hypothesis. In addition to the work of Campbell, Mankiw, and Hayashi, other key papers include Hall (1978), and Flavin (1981). One side effect of this specification is that it prevents us from computing equivalent variation. Since the behaviour of some of the households is inconsistent with (1), either because the households are at corner solutions or for some other reason, aggregate behaviour is inconsistent with the expenditure function derived from our utility function.
not constrained and choose consumption in accordance with (1) by \( \alpha \), total consumption expenditure is given by:

\[
p^c = \alpha_\theta (F_t + H_t) = (1 - \alpha_\theta) \gamma \text{INC}.
\]

(22)

The share of households consuming a fixed fraction of their income could also be interpreted as permanent-income behaviour in which household expectations about income are myopic.

Once the level of overall consumption has been determined, spending is allocated among goods and services according to a CES utility function. The demand equations for capital, labour, energy, and materials can be shown to be:

\[
p_i x^i = \delta_i \left( \frac{p_i}{p^c} \right)^{\sigma_i^{-1}} , i \in \{ k, l, e, m \}
\]

where \( y \) is total expenditure, \( x^i \) is household demand for good \( i \), \( \sigma_i \) is the top-tier elasticity of substitution, and the \( \delta_i \) are the input-specific parameters of the utility function. The price index for consumption, \( p^c \), is given by:

\[
p^c = \left( \sum_{j=k, l, e, m} \delta_j p_j^{\sigma_j^{-1}} \right)^{\frac{1}{\sum_{j}^{\sigma_j^{-1}}}}.
\]

(24)

Household capital services consist of the service flows of consumer durables plus residential housing. The supply of household capital services is determined by consumers themselves who invest in household capital, \( k^c \), in order to generate a desired flow of capital services, \( c^k \), according to the following production function:

\[
c^k = \alpha k^c
\]

(25)

where \( \alpha \) is a constant. Accumulation of household capital is subject to the condition:

\[
\dot{k}^c = J^c - \delta^c k^c.
\]

(26)

We assume that changing the household capital stock is subject to adjustment costs, so household spending on investment, \( F \), is related to \( J^c \) by:

\[
I^c = \left( 1 + \frac{\phi^c}{2} \right) J^c.
\]

(27)

Thus the household’s investment decision is to choose \( F \) to maximize:

\[
\int_{t}^{\infty} (p^c \alpha k^c - p^{i} J^c) e^{-\gamma (x^s - \delta^s (x^t - \cdot))} ds
\]

(28)

where \( p^c \) is the imputed rental price of household capital. This problem is nearly identical to the investment problem faced by firms and the results are very similar. The only important differences are that no variable factors are used in producing household capital services and there is no investment tax credit for household capital. Given these differences, the marginal value of a unit of household capital, \( \lambda_c \), can be shown to be:

\[
\lambda_c(t) = \int_{t}^{\infty} \left( p^c \alpha + p^{i} \frac{\phi^c}{2} \left( \frac{J^c}{k^c} \right)^2 \right) e^{-\gamma (x^s - \delta^s (x^t - \cdot))} ds
\]

(29)

where the integration is done along the optimal path of investment and capital accumulation, \((J^c(t), k^c(t))\). Marginal \( q \) is:

\[
q^c = \frac{\lambda_c}{p^c}
\]

(30)

and investment is given by:

\[
\frac{J^c}{k^c} = \frac{1}{\phi^c} (q^c - 1).
\]

(31)

The labour market

We assume that labour is perfectly mobile among sectors within each region but is immobile between regions. Thus, wages will be equal across sectors within each region, but will generally not be equal between regions. In the long run, labour supply is completely inelastic and is determined by the exogenous rate of population growth. Long-run wages adjust to move each region to full employment. In the short run, however, nominal wages are assumed to adjust slowly according to an overlapping contracts model, where wages are set based on current and expected inflation and on labour demand rela-
tive to labour supply. The equation below shows how wages in the next period depend on current wages, the current, lagged, and expected values of the consumer price level, and the ratio of current employment to full employment:

\[ W_{t+1} = W_t \left( \frac{p^c_{t+1}}{p^c_t} \right)^{\alpha_5} \left( \frac{p^c_t}{p^c_{t-1}} \right)^{-\alpha_6} \left( \frac{L_t}{L} \right)^{\alpha_6}. \] (32)

The weight that wage contracts attach to expected changes in the price level is \( \alpha_5 \), while the weight assigned to departures from full employment (\( L/\bar{L} \)) is \( \alpha_6 \). Equation (32) can lead to short-run unemployment if unexpected shocks cause the real wage to be too high to clear the labour market. At the same time, employment can temporarily exceed its long-run level if unexpected events cause the real wage to be below its long-run equilibrium.

The government

We take each region’s real government spending on goods and services to be exogenous and assume that it is allocated among inputs in fixed proportions, which we set to 1996 values. Total government outlays include purchases of goods and services plus interest payments on government debt, investment tax credits, and transfers to households. Government revenue comes from sales taxes, corporate and personal income taxes, and from sales of new government bonds. In addition, there can be taxes on externalities such as carbon dioxide emissions.

The government budget constraint may be written in terms of the accumulation of public debt as follows:

\[ \dot{B}_t = D_t + G_t + TR_t - T_t. \] (33)

where \( B \) is the stock of debt, \( D \) is the budget deficit, \( G \) is total government spending on goods and services, \( TR \) is transfer payments to households, and \( T \) is total tax revenue net of any investment tax credit.

We assume that agents will not hold government bonds unless they expect the bonds to be paid off eventually and accordingly impose the following transversality condition:

\[ \lim_{s \to \infty} B(s)e^{-(R(s)-r)s} = 0 \] (34)

This prevents per-capita government debt from growing faster than the interest rate forever. If the government is fully leveraged at all times, (34) allows (33) to be integrated to give:

\[ B_t = \int_0^\infty (T - G - TR)e^{-(R(s)-r)(s-t)} ds. \] (35)

Thus, the current level of debt will always be exactly equal to the present value of future budget surpluses.27

The implication of (35) is that a government running a budget deficit today must run an appropriate budget surplus as some point in the future. Otherwise, the government would be unable to pay interest on the debt and agents would not be willing to hold it. To ensure that (35) holds at all points in time we assume that the government levies a lump-sum tax in each period equal to the value of interest payments on the outstanding debt.28 In effect, therefore, any increase in government debt is financed by consols, and future taxes are raised enough to accommodate the increased interest costs. Other fiscal closure rules are possible, such as requiring the ratio of government debt to GDP to be unchanged in the long run. These closures have interesting implications but are beyond the scope of this paper.

Financial markets and the balance of payments

The eight regions in the model are linked by flows of goods and assets. Flows of goods are determined by the import demands described above. These demands can be summarized in a set of bilateral trade matrices which give the flows of each good between exporting and importing countries.

Trade imbalances are financed by flows of assets between countries. Each region with a current-account deficit will have a matching capital account surplus, and vice versa.29 We assume asset markets

---

27 Strictly speaking, public debt must be less than or equal to the present value of future budget surpluses. For tractability we assume that the government is initially fully leveraged so that this constraint holds with equality.

28 In the model the tax is actually levied on the difference between interest payments on the debt and what interest payments would have been if the debt had remained at its base case level. The remainder, interest payments on the base-case debt, is financed by ordinary taxes.

29 Global net flows of private capital are constrained to be zero at all times—the total of all funds borrowed exactly equals the total funds lent. As a theoretical matter this may seem obvious, but it is often violated in international financial data.
are perfectly integrated across regions. With free mobility of capital, expected returns on loans denominated in the currencies of the various regions must be equalized period to period according to a set of interest-arbitrage relations of the following form:

\[ i_k + \mu_k = i_j + \mu_j + \frac{E_k^j}{E_k} \]  

(36)

where \( i_k \) and \( i_j \) are the interest rates in countries \( k \) and \( j \), \( \mu_k \) and \( \mu_j \) are exogenous risk premiums demanded by investors (calibrated in the baseline to make the model condition hold exactly with actual data), and \( E_k^j \) is the exchange rate between the currencies of the two countries.

Capital flows may take the form of portfolio investment or direct investment, but we assume these are perfectly substitutable \textit{ex ante}, adjusting to the expected rates of return across economies and across sectors. Within each economy, the expected returns to each type of asset are equated by arbitrage, taking into account the costs of adjusting physical capital stock and allowing for exogenous risk premiums. However, because physical capital is costly to adjust, any inflow of financial capital that is invested in physical capital will also be costly to shift once it is in place. This means that unexpected events can cause windfall gains and losses to owners of physical capital and \textit{ex-post} returns can vary substantially across countries and sectors. For example, if a shock lowers profits in a particular industry, the physical capital stock in the sector will initially be unchanged but its financial value will drop immediately.

\textit{Money demand}

Finally, we assume that money enters the model via a constraint on transactions. We use a money-demand function in which the demand for real money balances is a function of the value of aggregate output and short-term nominal interest rates:

\[ MON = PYt^\epsilon \]  

(37)

where \( Y \) is aggregate output, \( P \) is a price index for \( Y \), \( i \) is the interest rate, and \( \epsilon \) is the interest elasticity of money demand. The supply of money is determined by the balance sheet of the central bank and is exogenous.

\textbf{Assessing the model}

All models have strengths and weaknesses, and G-Cubed is no exception. Its most important strength is that it distinguishes between financial and physical capital, and includes a fully integrated treatment of inter-temporal optimization by households, firms, and international portfolio holders. This allows the model to do a rigorous job of determining where physical capital ends up, both across industries and across countries, and of determining who owns the physical capital and in what currency it is valued. Overall, the key feature of G-Cubed is its treatment of capital, and that is also what most distinguishes it from other models in the macro, trade, or CGE literatures.

G-Cubed also has other strengths. All budget constraints are satisfied at all times, including both static and inter-temporal budget constraints on households, governments, and countries. Short-run behaviour captures the effects of slow wage adjustment and liquidity constraints, while long-run behaviour is consistent with full optimization and rational expectations. In addition, wherever possible the model’s behavioural parameters are determined by estimation, which is discussed further in chapter 4 of McKibbin and Wilcoxen (2001).

For the two-country model used in the remainder of this paper, the parameter estimates are aggregated from the underlying 12 sectors to two sectors using output shares as weights. Also, the initial values of foreign debt of each country are set to zero in the prototype model to preserve symmetry.

\[ ^{30} \text{The mobility of international capital is a subject of considerable debate; see Obstfeld (1995) or Feldstein and Horioka (1980).} \]

\[ ^{31} \text{Unlike other components of the model we simply assume this rather than deriving it from optimizing behaviour. Money demand can be derived from optimization under various assumptions: money gives direct utility; it is a factor of production; or it must be used to conduct transactions. The distinctions are unimportant for our purposes.} \]
REFERENCES


