ELSEVIER

Economic Modelling 21 (2004) 735–758

Economic Modelling

www.elsevier.com/locate/econbase

Comparing empirical models of the euro economy

Kenneth F. Wallis*

Department of Economics, University of Warwick, Coventry CV4 7AL, UK

Abstract

This article presents a comparative analysis of four macroeconometric models whose proprietors participated in a model comparison conference focused on the new euro area economy. One model, the Area-Wide Model recently developed at the European Central Bank, treats the whole area as a single economy. The other three, MULTIMOD, NIGEM and QUEST, are established multicountry models that provide disaggregated analysis of questions of economic policy in Europe. Their structural characteristics and the results of two policy simulations are compared and contrasted. The principal source of simulation differences is the different degree of forward-looking behaviour incorporated in the models.

JEL classifications: C5; E1

Keywords: Macroeconometric models; Multicountry models; Model comparisons; Simulation; Euro-area economy

1. Introduction

On 1 January 1999 the third and final stage of European Monetary Union began, with the irrevocable fixing of exchange rates between the currencies of the eleven EU members initially participating in Monetary Union, and with the conduct of a single monetary policy under the responsibility of the European Central Bank. The primary objective of monetary policy is to maintain price stability in the euro area as a whole, hence new challenges were faced by the econometric models employed in forecasting and policy analysis.

*Tel.: +44-24-7652-3055; fax: +44-24-7652-3032.
E-mail address: k.f.wallis@warwick.ac.uk.

0264-9993/04/$ - see front matter
doi:10.1016/j.econmod.2003.10.003

One response was to provide an area-wide focus directly, by developing a macroeconometric model that treats the whole area as a single economy. This precludes the consideration of the possibly different impacts of a common policy on different countries, and a system of linked national-economy models is under development within the Eurosystem to provide this. In the meantime existing multicountry models are an alternative source of such analysis, and the objective of the present model comparison exercise is to evaluate the advantages and disadvantages of the different approaches. Along with the Area-Wide Model (AWM) of the European Central Bank (Fagan et al., 2001), the model comparison exercise features three well-established multicountry models, namely

- MULTIMOD Mark III, developed in the Research Department of the International Monetary Fund in Washington, DC (Laxton et al., 1998), with further European extensions for this exercise,
- NIGEM, developed at the National Institute of Economic and Social Research, London (Barrell et al., 2001), and
- the QUEST model of the European Commission (Roeger and in't Veld, 1997).

The present exercise is in the style of a model comparison conference based on material supplied by the model proprietors, describing model exercises that they have carried out, as distinct from a third-party comparative research project, based on implementation of the models by independent researchers. The differences between the two styles of model comparison are discussed by Mitchell et al. (1998) in introducing the results of their independent third-party research on three multicountry models, among which are earlier versions of MULTIMOD and NIGEM.

Structural change and aggregation questions are at the heart of the modelling challenges presented by the final stage of European Monetary Union, although both have a long history, in the context of EMU and more generally. Structural change in models reflects changes in the problems they address and changes in the economies they represent. For potential EMU members the approach to monetary union increased the attention paid to the various convergence criteria, leading to substantial changes in fiscal policy. The final move to a single currency and a single monetary policy was the culmination of a series of policy changes over the previous two decades, as quantitative controls were abandoned, financial liberalisation proceeded apace, and monetary policy focussed more directly on the control of inflation. With exchange rate risk among member countries removed and transactions costs reduced the consequences continue, as the whole spectrum of interest rates converges and financial markets deepen. A structural rather than reduced form modelling approach is clearly needed to reflect these changes and their consequences, as is argued with detailed examples from the present context by Maclennan et al. (2000), for example.

Aggregation issues are central to the comparison of area-wide and individual country modelling, and have a theoretical literature that starts in the middle of the last century (Klein, 1946; Nataf, 1948; Gorman, 1953; Theil, 1954). The technical literature considers such questions as the conditions for 'perfect' aggregation and tests thereof, and statistical discrimination between aggregate and disaggregate models. A typical result is that, given linear regression models for an aggregate variable and each of its several components, the goodness of fit of the aggregate variable obtained by summing the fitted values of the component models is no worse than that of the aggregate model, and is potentially much better the more diverse are the component models. However, such results are usually obtained under the assumption that the various models are correctly specified, whereas specification uncertainty is endemic, and little is known about its impact. The more components there are to be modelled, the greater the opportunities for specification error. With respect to model specification in the non-linear case, which is more relevant for practical macroeconometric models, an important observation by Stoker (1986) is that models estimated with aggregate data contain distributional biases that are not measurable using aggregate data alone; time series of distributional data are required to correct these effects. Distributional movements are typically small, hence their effects may tend to persist in aggregate data over many periods, increasing the apparent explanatory power of lagged variables if they are ignored.

There is also much literature on country differences across Europe and their implications; see again Maclennan et al. (2000) and references therein. Only a small part of this is quantitative or model-based, and here a general conclusion seems to be that, for problems for which an aggregate measure of outcomes is appropriate, the differences between using an aggregate analysis and adding together the results of disaggregated analyses are not dramatic. If distributional effects appear in the loss function, however, then only a disaggregated analysis can provide an answer. They do not feature in decision-making by the European Central Bank, which always adopts an area-wide perspective, but they may become a concern of national policy agencies, should differences in economic developments across countries threaten to become too large. Distributional matters may also affect the area-wide aggregates that the ECB targets, as noted above, hence they may become a matter of concern for the ECB even though they do not feature directly in its objectives. The present model comparison exercise provides a first model-based evaluation of the nature and scale of such effects in a euro area setting.

The remainder of this article is organised as follows. Section 2 contains a comparative account of the structural characteristics of the four models, as determined by their underlying theoretical foundations and their empirical implementation, and two key areas are emphasised. The first is the supply side and inflation-unemployment nexus of the models, where a core supply-side framework is used to organise the discussion. The second area of the models that is emphasised is their treatment of consumption and investment decisions. Section 3 is concerned with two simulation experiments that demonstrate the responses of the models to fiscal and monetary policy shocks. Issues of experimental design are first discussed, and then the simulation results are compared across models and across countries. It is seen that some of the differences that emerge can be related to structural features of the models elucidated in the preceding section. Section 4 contains concluding comments.

2. Four models of the euro economy

2.1. General characteristics

The Area-Wide Model (AWM) of the European Central Bank treats the euro area as a single economy, as noted above. It corresponds to a conventional national-economy macro model, with the rest of the world treated as exogenous. Although this includes the three EU members that have not adopted the single currency, and Greece, which did not join the third stage of EMU until 1 January 2001, in the current model the rest of the world is proxied by a four-country aggregate, comprising the United States, Japan, the United Kingdom and Switzerland. The euro area dataset used in the model is based on reconstructed euro area national accounts data consistent with individual member-country data, carefully aggregated. The main discrepancy between the constructed euro area accounts and conventional national accounts arises in the trade data, where the constructed data include both trade within the area − between member countries – and trade with the rest of the world.

In contrast to national-economy models, global economic models focus on macroeconomic interactions among nations. Although the three global models under consideration describe the same world they differ in respect of which countries are modelled individually and how the remaining countries are grouped together. The original Mark III version of MULTIMOD (Laxton et al., 1998) contains separate models for each of the Group of Seven (G7) countries – the US, Canada, Japan, France, Germany, Italy, UK – and for an aggregate grouping of 14 smaller industrial countries. The remaining economies of the world are then aggregated into two separate blocks of developing and transition economies. Each of the industrial-country models has the structure of a complete national-economy model, whereas the two remaining blocks are modelled in much less detail, focussing on foreign trade and the evolution of net foreign assets in a more reduced form than structural manner. The first modified version of the model used in the present exercise has a separate subgroup of ‘other euro area’ countries, whose results can be combined with those of France, Germany and Italy to obtain euro area macroeconomic outcomes. The second modified version (Mark IIIB) has a single euro area block and is estimated on more recent data.

NIGEM and QUEST are more disaggregated, each having complete country models for each EU member (taking the Belgium-Luxembourg Economic Union as a single entity) and a finer division of the rest of the world. After separate country models for the remaining three G7 members, NIGEM has six intermediate-size models consisting of a very basic description of the domestic economy (output and prices) together with trade volume and price equations and the balance of payments; these relate to Australasia, Mexico, Norway, S. Korea, Switzerland and the Visegrad economies (Czech Republic, Hungary, Poland). Finally, there are seven simpler trade-and-payments models for China, OPEC, Developing Europe, Africa, Latin America, other LDCs, and the rest of E. Asia. QUEST separately models the US and Japan, then describes the rest of the world by 11 trade-and-payments models. Four of these relate to the larger remaining OECD countries (Australia, Canada, Norway, Switzerland) and seven to various groups of countries, very much as above (Central and Eastern Europe, the rest of the OECD, OPEC, the former Soviet Union, 'dynamic' Asian economies, the rest of Asia, and the rest of Africa and Latin America).

The theoretical foundation of the complete national-economy submodels is common to all countries in each case, and also has common elements across the four models. All follow the currently prevailing paradigm in which a broadly neoclassical view of macroeconomic equilibrium coexists with a new Keynesian view of short-to-medium-term adjustment. The simple Mundell-Fleming model remains the basic underlying model of interdependence, but it is considerably extended, by adding dynamics, modelling the processes generating aggregate supply, accounting for the behaviour of asset and debt stocks as well as financial flows, and endogenizing fiscal and monetary policy. All the models incorporate forward-looking behaviour in financial markets and, in some cases discussed below, in decision-making by firms and households; they all treat future expectations as 'rational' or 'model-consistent' expectations. Some important differences across models and/or countries that nevertheless appear in two key areas are described in the following sub-sections.

Two potential sources of such differences are different relative weights on theory and data in the specification of the models, and different approaches to quantification. MULTIMOD and QUEST have more thoroughly elaborated theoretical foundations, particularly with respect to the steady state of the models, than AWM and NIGEM, whose steady state properties are based more on the use of cointegration techniques in estimation, with long-run coefficients sometimes imposed following pre-testing. MULTIMOD is an annual model, and sometimes compensates for the resulting lack of degrees of freedom in estimation by using pooled estimation methods, whereupon not only the equation specifications are common to all countries but also their coefficient estimates. The three 'European' models are quarterly and make less use of pooled estimation, NIGEM allowing different dynamic specifications across countries to emerge in estimation. However, QUEST's dynamic responses also have a more theoretical basis, and its emphasis on theoretical foundations sometimes results in the calibration of key parameters rather than their estimation from aggregate time series data.

The terms equilibrium, long run, and steady state are used interchangeably in this discussion, and all refer to the steady-state properties of a system of dynamic equations, designed in the present context for short-to-medium-term forecasting and policy analysis. Thus the range of issues considered under the heading of 'long-run implications …' is only a subset of what economists more generally might wish to consider as long-run issues. Here the long-run equilibrium is a steady-state growth path for the aggregate real and nominal variables of the model, with constant inflation. The steady-state real growth rate is equal to the natural growth rate of the standard neoclassical growth model, equal to the rate of (labour-augmenting) technical progress plus the growth rate of the population, both treated as exogenous. Nominal equilibrium is 'anchored' by specifying a target, again exogenously, for a nominal variable such as inflation, with the target value typically achieved in steady state via an appropriately specified feedback rule for nominal interest rates, as illustrated in our simulations.

The study of steady-state properties proceeds in various ways. In MULTIMOD each of the dynamic equations has an explicit steady-state analogue equation, and the complete system of steady-state analogue equations is maintained separately from the system of dynamic equations, to provide an interpretative device for understanding long-run comparative statics and to facilitate model handling by providing terminal conditions for model-consistent expectations solutions. In the remaining models there is no formal description of the steady state, although its theoretical properties are well understood, and it is studied by simulation methods. In a deterministic solution of the complete system of non-linear equations each individual equation holds exactly (to a numerical approximation) and hence has an implicit steady-state counterpart, at least in the neighbourhood of the solution. For some equations it may be possible to obtain an explicit steady-state representation, whereupon partial or sub-system analysis can proceed by analytical methods, as in some examples below.

2.2. The supply side and the inflation-unemployment nexus

The prevailing paradigm noted above implies that at least in respect of the long-run equilibrium, the level of real activity is independent of the price level and the steady-state inflation rate, whereas in the short run there is considerable real and nominal inertia. Adjustment costs and contractual arrangements imply that markets do not clear instantaneously; instead there is a relatively slow process of dynamic adjustment to equilibrium. This is by no means a full-employment equilibrium, however, and important questions are whether a model possesses a non-accelerating-inflation rate of unemployment (NAIRU) and, if so, what are its determinants. If these determinants do not include the steady-state inflation rate then the NAIRU coincides with the 'natural' rate of unemployment.

The underlying theoretical framework of the three European models, like that of many national-economy models in Europe, is one of imperfectly competitive goods and labour markets, with prices set by firms as a mark-up on costs, given demand. Wages are determined in a bargaining process and if firms have the 'right to manage' they set employment, given the wage, although wage behaviour is relatively insensitive to the particular specification of the bargaining model. The preceding questions about the NAIRU or, equivalently, whether the aggregate supply schedule is vertical and what causes it to shift, can then be addressed by analysing a core supply-side model consisting of the steady-state versions of the wage and price equations together with, in an open-economy context, the response of the exchange rate (Joyce and Wren-Lewis, 1991; Turner, 1991). We first present a simple algebraic representation of this 'core' system (Wallis, 2000), and then use it as a framework for comparison of the four models. We note that MULTIMOD, in contrast to the European models, but like several North American models, assumes that firms are perfectly competitive and have no mark-up opportunities.

Long run: Given wage and price equations expressed in error-correction form, the first step is to solve the long-run or error-correction parts of the equations for the NAIRU, which requires their static homogeneity. The long run of a typical wage equation may be written

$$w = p + pr - \alpha u + z^{w} \tag{1}$$

where w, p and pr are (log) nominal earnings, producer prices and average labour productivity, respectively, u is the unemployment rate (entered linearly for simplicity) and z^{w} is a set of wage pressure variables. A typical assumption about pricing behaviour is that producer prices are determined as a mark-up on import and unit labour costs, with the mark-up increasing with capacity utilization (cu), hence

$$p = \beta p_{m} + (1-\beta)(w - pr) + \gamma cu, \tag{2}$$

where p_{m} is the import price. The capacity utilization term would not appear if an equilibrium state were defined as one in which output coincides with 'normal' output; it may also be related to unemployment. Static homogeneity is implicit in the unit coefficients in Eq. (1) and the unit sum of the first two coefficients in Eq. (2); these restrictions are usually found to be data-admissible.

The real exchange rate is defined as

$$q = p - p_{m}. \tag{3}$$

Substituting for p_{m} from Eq. (3) and then solving Eqs. (1) and (2) gives the equilibrium rate of unemployment as

$$u^{*} = \frac{1}{\alpha}\left[z^{w} - \frac{\beta}{1-\beta}q + \frac{\gamma}{1-\beta}cu\right].$$

Dynamic equations: The error-correction formulation expresses the dynamic equations in terms of the first differences (Δs) of the (log) levels of wages and prices, that is, the wage and price inflation rates, together with lagged values of these and other variables expressed in similar form, except for the levels variables that appear in the error-correction term itself. Thus, deleting dynamic terms in other variables and the time subscript for convenience, we have

$$\phi_{1}(L)\Delta w = \theta_{1}(L)\Delta p + \psi_{1}(L)\Delta pr - \delta_{1} ec(w) \tag{4}$$

$$\phi_{2}(L)\Delta p = \theta_{2}(L)\Delta cost - \delta_{2} ec(p) \tag{5}$$

where $\phi(L)$ etc. are polynomials in the lag operator L, which may include forward terms, $ec(w)$ and $ec(p)$ are the error-correction terms given as the (lagged) residuals from Eqs. (1) and (2), respectively, and the cost terms have been combined into a single variable.

We consider a steady state with a constant rate of price inflation π in all countries (and hence constant nominal exchange rates), and constant real wage growth equal to the rate of growth of productivity, that is

$$\Delta p = \Delta p_m = \Delta cost = \pi$$

$$\Delta w = \pi + \Delta pr = \pi + g.$$

Other first-difference terms that appear are set to zero, for example, the wage-pressure variables are assumed constant. In such a steady state, Eqs. (4) and (5) then give the NAIRU as

$$u^{\dagger} = u^{*} + \frac{1}{\alpha}\frac{\psi_1(1) - \phi_1(1)}{\delta_1}g + \frac{1}{\alpha}\left\{\frac{\theta_1(1) - \phi_1(1)}{\delta_1} + \frac{\theta_2(1) - \phi_2(1)}{\delta_2}\right\}\pi.$$

This is the natural rate of unemployment, independent of π, if the expression in curly brackets is zero. A sufficient condition is that each of the two terms in the curly brackets is zero, that is, that the wage and price equations are dynamically homogeneous or inflation neutral. In empirical work on the European economies this is more commonly found to be the case for wage equations than for price equations.

It is found in some countries that the wage equation is not dynamically homogeneous with respect to productivity, with $\psi_1(1) < \phi_1(1)$, so that in a dynamic steady state with real wage growth occurring at the rate of growth of productivity, this steady-state productivity growth rate affects the *level* of real wages and the NAIRU: lower productivity growth gives a higher natural rate of unemployment. This is consistent with the evidence of Manning (1992) that the slowdown in productivity growth was an important explanation of the increase in unemployment in many OECD countries. It suggests that, of the two competing effects of growth on unemployment identified by Aghion and Howitt (1994), the ‘capitalisation’ effect, whereby an increase in growth raises the capitalised returns from creating jobs and consequently reduces the equilibrium unemployment rate, empirically dominates the ‘creative destruction’ effect.

Taxation: If the relevant taxes are identified in the model, then different real wage concepts are relevant to the objectives of the two parties to the wage bargain. For employers, what matters are their real wage costs $(w + t_e - p)$, namely nominal wages plus employment taxes deflated by producer prices, whereas employees focus on the real consumption wage $(w - t_d - p_c)$, namely nominal wages less direct taxes deflated by consumer prices (p_c, in logs, with average tax rates t_e and t_d expressed as proportions). The difference between the two is the tax wedge

$$(w + t_e - p) - (w - t_d - p_c) = t_d + t_e + p_c - p.$$

Let λ be the weight of domestic goods in (pre-tax) consumer prices, so that the (post-tax) consumer price variable is

$$p_c = \lambda p + (1-\lambda) p_m + t_i$$

where t_i is the average indirect tax rate. Then the wedge is given as

$$wedge = t_d + t_i + t_e + t_m,$$

where $t_m = (1-\lambda)(p_m - p)$ is the 'tax' imposed by high import prices; these definitions imply that $p_c = p + t_i + t_m$.

This formulation has two implications for the wage equation. The first is that all four tax variables, treated as wage-pressure variables, have the same effect; it is customary to combine them into a single tax wedge variable, which also avoids the statistical difficulties of accurately estimating the separate effects of variables which typically show relatively little variation over the sample. The second implication is that, in order to assess the incidence of taxation, the wage equation should be specified with one or other of the two after-tax real wage variables as dependent variable. Choosing employers' real wage costs, for example, Eq. (1) becomes

$$(w + t_e - p) = pr - \alpha u + \rho(t_d + t_i + t_e + t_m) + \text{other } z^w \text{ variables},$$

where ρ is a measure of the incidence of taxation. A null hypothesis of interest might then be that $\rho = 0$, conforming to the view that, in the long run, rises in the wedge are borne entirely by labour. Short-run wedge effects may nevertheless be important and rather persistent, with $\Delta(wedge)$ appearing in the error-correction form of the wage equation.

A final observation on this analysis is that the answers to some important 'long-run' questions are seen to depend on the specification of the short-run or dynamic adjustment parts of the equations, in contrast to the very sharp separation found in the cointegration literature between the long-run and short-run implications of statistical models. Constant terms have not featured in the analysis, whereas they are often present in estimated equations. Johansen (1991) draws the important distinction between the two roles of a constant term in the vector error-correction model: one contributes to the intercept in the cointegrating relation while the other determines a linear trend (exponential trend in the case of a log-linear model). He provides a procedure for testing that the trend is absent, although this is not applicable in the single-equation case, where the allocation between the two roles is at the disposition of the modeller, as in some examples below.

AWM: The key wage and price equations of AWM are similar in appearance to the above core model, but differ from it in important ways. The wage equation is

$$\Delta(w - p_c - pr) = 0.27\Delta(w - p_c - pr)_{-1} + 0.2(\pi^e - \Delta p_{c,-1}) + \theta_1(L)\Delta^2 p_c + \psi(L)\Delta^2 pr$$

$$-0.015(u - utr)_{-1} - 0.10\left[\overline{w - pr} - p - \ln(1-\beta)\right]_{-1}$$

where p_c is the consumption deflator, p is the GDP deflator at factor cost, u is the unemployment rate, utr its trend, and $\overline{w-pr}$ denotes trend unit labour costs (all in logs); π^e is inflation expectations. Despite using different price deflators in the long-run and short-run parts of the equation there is no consideration of wedge effects, nor any other wage pressure variable. In the long run real unit labour costs are equal to labour's share, in turn equal to the labour elasticity in the production function, $1-\beta$. Dynamic homogeneity holds, but since the unemployment term is the deviation of unemployment from the exogenous NAIRU the statement that 'the long-run Phillips curve is vertical in the model' has less empirical content here than in our core model.

The price equation contains the same error-correction term (whose sign is changed above to maintain the usual convention of a negative coefficient), as follows:

$$\Delta p = 0.0039 + 0.23\Delta p_{-1} + 0.2(\pi^e - \Delta p)_{-1} + 0.03 ygap_{-1} + 0.031\Delta p_{m,-1}$$
$$+ \theta_2(L)\Delta(\overline{w-pr}) - 0.045\left[p - (\overline{w-pr}) + \ln(1-\beta)\right]_{-1}$$

Dynamic homogeneity is strongly rejected by the data, which is interpreted as implying that 'in principle the mark-up in the long run depends on steady state inflation' (Fagan et al., 2001, p. 17). Solving the steady-state version of the estimated equation for a common steady-state rate of inflation of domestic and foreign prices and unit labour costs gives a rate of $0.0039/(1-0.746)=0.015$ per quarter or 6% per annum, which is close to the average inflation rate of the GDP (at factor cost) deflator over the sample period. In simulation exercises the constant term is adjusted to ensure that 'the long-run real equilibrium of the model coincides with the theoretical steady state' at an inflation rate, in the present case, of 2% per annum.

MULTIMOD: Like its predecessors, Mark III has no labour market − no employment and wage equations. Wage setting is subsumed with price setting into a 'Phillips curve' described as a reduced-form relationship between (price) inflation and unemployment. The latter variable is of course endogenous in the model; the inflation equation is in reduced form in the sense that it can be derived from a two-equation system as above by substituting the wage equation into the price equation (Laxton et al., 1998, Box 6; Sgherri, 2000, Appendix). Unemployment enters non-linearly, to reflect increases in the degree of real wage rigidity as the unemployment rate rises. The equation is

$$\pi = \lambda \pi^e_{+1} + (1-\lambda)\pi_{-1} - \gamma\left(\frac{u-u^*}{u-\phi}\right)$$

where π^e, u^* and ϕ are constructed proxy series, u^* representing a NAIRU concept.

Coefficient estimates for the three main euro-area countries are

France	$\lambda = 0.75$	$\gamma = 0.011$
Germany	$\lambda = 0.74$	$\gamma = 0.008$
Italy	$\lambda = 0.91$	$\gamma = 0.023$

The non-linearity of the equation implies that country responses to an aggregate demand shock may depend not only on these coefficient differences but also on differences in their relative initial positions. A supply-side shock can be implemented by directly perturbing the NAIRU.

The inflation-unemployment nexus is closed by an unemployment equation that translates output gaps $y - \bar{y}$, where $\bar{y}$ is an estimate of potential output, into unemployment gaps $u - \bar{u}$, where $\bar{u}$ is an estimate of the natural rate. The equation is

$$(u - \bar{u}) = \gamma_1 (y - \bar{y}) + \gamma_2 (u - \bar{u})_{-1}$$

France	$\gamma_1 = -0.30$	$\gamma_2 = 0.44$
Germany	$\gamma_1 = -0.33$	$\gamma_2 = 0.18$
Italy	$\gamma_1 = -0.09$	$\gamma_2 = 0.79$

It is sometimes described as a derived demand for labour function, but it has no price effects and assumes constant labour supply.

NIGEM: The general structure of the wage and price equations of NIGEM and their delivery of the NAIRU is very close to our core model. The wage equations exhibit static and dynamic homogeneity. They contain a mixture of forward-looking and backward-looking terms in price inflation, and in terms of the degree of forward-lookingness the three main euro-area economies are ranked in the same order as in MULTIMOD's Phillips curve, namely Italy, France, Germany. (In both models the corresponding UK equation is entirely forward-looking.)

Price setting refers to a CES cost function, hence producer prices are driven by the user cost of capital as well as import prices and unit labour costs. Then the NAIRU in turn depends on the equilibrium real interest rate. The price equations in all countries have static homogeneity, but dynamic homogeneity is less widespread. In particular, the UK price equation is dynamically homogeneous whereas those of France, Germany and Italy are not, as in the corresponding euro area specification of AWM. Again the interpretation is that the mark-up depends on the inflation rate.

QUEST: The labour market specification in QUEST is based on theoretical search models, nevertheless wage bargaining results in a wage equation of the general form discussed above. A specific wage pressure variable is the reservation wage, equal to the pre-tax value of unemployment benefits, and so both the income tax rate and the benefit rate affect the NAIRU. A supply-side shock that affects the long-term unemployment rate can be implemented by perturbing the reservation wage, whereas AWM and MULTIMOD tell no story about what it is that alters the NAIRU and what other effects it might have in the model. Dynamics in QUEST are explicitly based on staggered contracts, hence expected future prices for the duration of the contract appear in the equation.

The price equation in the present version of the model is a version of the 'hybrid' equation with forward-looking and backward-looking pricing behaviour estimated on the AWM dataset by Gali et al. (2001). The inflation dynamics are common to all countries, and the forward coefficient of 0.7 is close to the corresponding estimates for France and Germany in MULTIMOD, although in that model other countries take different values and the change of data frequency should be noted. Sensitivity to the output gap is lowest in Germany, again as in MULTIMOD. The work of Gali et al. (2001) is an extension to the euro area of the 'New Keynesian Phillips Curve' estimated on US data by Gali and Gertler (1999). The original US equation is subjected to rigorous testing, which it does not survive, by Rudd and Whelan (2001, 2002). Its European counterpart suffers a similar fate at the hands of Bardsen et al. (2002), who also report similar findings for Norway and the United Kingdom. Likewise, Balakrishnan and Lopez-Salido (2002) find that the New Phillips Curve performs badly in the United Kingdom.

2.3. Consumption and investment decisions

Consumption: The general framework for the analysis of consumers' expenditure decisions is the Blanchard-Yaari life-cycle model of intertemporal optimization by finitely-lived households, moderated by a proportion of households who are unable to implement full intertemporal optimization because of financial constraints and so consume their current disposable income. The models differ in their relative weights on theory and data, as noted above. MULTIMOD and QUEST pay more detailed attention to the underlying framework and its 'deep' parameters, whereas AWM and NIGEM simply report an aggregate 'solved out' consumption function, using Muellbauer and Lattimore's (1995) term; this is in error-correction form.

AWM: The short-run part of the error-correction equation in the present version of the AWM model includes current and lagged income growth, together with a direct real interest rate term. In the long run consumers' expenditure is related to disposable income and wealth, with elasticities 0.65 and 0.35, respectively. Wealth is defined as cumulated savings, assuming that households own all the assets in the economy, namely public debt, net foreign assets and the private capital stock, hence there are no asset price effects.

MULTIMOD: The theoretical framework in MULTIMOD incorporates a concave life-cycle profile of labour income, rising with age and experience when individuals are young, and eventually declining with retirement. This is estimated from data on labour income and employment by age cohort for the US, and these estimates are then used in all other industrial countries. Laxton et al. (1998, Table 10) report a G7 pooled estimate of 0.41 for the intertemporal elasticity of substitution, and country-specific estimates of the sensitivity of consumption to disposable income. These imply unusually large estimates of the share of income-constrained consumption, and the new estimates in Mark IIIB reduce this to 44% for the euro area.

NIGEM: The consumption function in NIGEM has the same long-run form as AWM, but adjustment behaviour is forward looking. Panel estimation of common parameters for Europe (but not the UK) yields a greater long-run elasticity with respect to income of 0.83 in Europe and 0.86 in the UK. The wealth measure is financial wealth.

QUEST: The consumption model in QUEST, like that in MULTIMOD, is formulated in terms of deeper parameters. The share of liquidity-constrained consumption is set at 0.3 in all countries, since individual country estimates based on Euler equations do not differ significantly from this. The intertemporal elasticity of substitution for unconstrained consumers is calibrated at 0.5, having in mind estimates in the literature from both aggregate time series data and household survey data.

Investment: In AWM and NIGEM the behaviour of investment is a process of dynamic adjustment to a steady state described by factor demand equations consistent with the underlying production function. A Cobb–Douglas production function is assumed in AWM, and the investment/output ratio responds to deviations of the ratio of investment to potential output from its implied steady-state value. This is a function of the user cost of capital, the variable component of which is represented in the model equation by the short-term real interest rate. The underlying technology in NIGEM is CES, as noted above, and estimates of the elasticity of substitution for each country are obtained from the labour demand functions. These determine the steady-state capital/output ratio, with the user cost of capital depending on the long real rate of interest. The response to a temporary shock to nominal interest rates might then be expected to be more muted than in AWM. There is a little variation in the dynamics of adjustment across countries in NIGEM.

In contrast MULTIMOD and QUEST adopt the model of intertemporal optimization by firms based on Tobin's q theory. While theoretically attractive, in particular specifying a role for profit taxes, this is difficult to implement econometrically. In MULTIMOD a common equation is based on a panel of the industrial countries, whereas in QUEST, as in NIGEM, there is some variation in the dynamics of adjustment across countries.

3. Two simulation experiments

3.1. Experimental design

Our comparative analysis is based on the results of the fiscal policy and monetary policy simulations, focusing on the outcomes for GDP and inflation. The simulations were carried out by the model proprietors under conditions that were standardised as far as possible, and full results are presented in their separate articles (Barrell et al., 2004; Dieppe and Henry, 2004; Hunt and Laxton, 2004; Roeger and in't Veld, 2004).

The fiscal shock is a reduction in government spending of 1% of GDP, and the monetary shock is an increase of 1% point in the nominal short-term interest rate. Both are unanticipated temporary shocks, maintained for 1 year (four quarters). During this period the models' normal policy reaction functions are switched off, and then reintroduced once the temporary perturbation is over. In the fiscal shock this implies that the relevant policy instruments, namely a fiscal tax rate and the nominal short-term interest rate, are held at their base-run values in the first year; in the monetary shock the fiscal instrument is again held at base, whereas it is the monetary policy instrument that is perturbed.

Standardisation of policy regimes remains a contentious issue. Mitchell et al. (2000) argue, in the light of their own analysis and the experience of the second Brookings model comparison project (Bryant et al., 1993), that it is not enough to ask modellers to run simulations 'with their normal tax-rate reaction functions in place'. They show that the 'normal' fiscal policy rules on some leading models differ substantially in their impact on simulation results. In the present exercise the variation among rules is much less, however, since all four models operate tax-difference rules that target either the debt/GDP ratio (MULTIMOD, QUEST) or the deficit/GDP ratio (AWM, NIGEM), and Mitchell et al. (2000) show that it is immaterial whether the exogenous targets are expressed in terms of the debt or the deficit, provided that the choices are mutually consistent. The parameterisation of the rules determines the nature and speed of the dynamic adjustment of their target variables, and here too different rules can be adjusted to give equivalent outcomes on a given model. These are a function not only of the rule but also of the model's intrinsic dynamics, however, and standardisation of the performance of different rule-model combinations may be necessary to obtain a sharp comparative focus on other areas of the models. But this would require experimentation that is beyond the scope of the present exercise, and the normal tax-rate reaction functions were accordingly left in place. With respect to monetary policy, there has been much convergence in recent years towards interest-rate rules that target inflation. Although differences in their normal rules remain across the models there was interest among the modellers in standardising on a 'Taylor' rule that also targets the output gap, with coefficients 1.0 and 0.25 on the inflation and output terms, respectively. In the event additional results under a variety of monetary policy rules are also reported for all models except MULTIMOD, which only considers this standard Taylor rule.

The organisers of the first Brookings multicountry model comparisons (Bryant et al., 1988) also attempted to standardise the simulation baseline path, in order to promote comparability of simulation results by eliminating potential base dependence. Non-linearity in models is usually cited as a possible source of base dependence, a practical example being provided by the study of the impact of starting point conditions on fiscal multipliers in the accompanying article by Hunt and Laxton. One response is to apply relatively small shocks in simulations and rely on local linearity. Even in the absence of complicated non-linear functions, however, stock-flow effects can be seriously influenced by differences in baseline stocks. The example cited by Bryant et al. (1988, p. 30) is the stock of government debt, where 'differences in baseline debt stocks can alter significantly the simulated effects of a change in interest rates'. This example is relevant to the present exercise where, in particular, the different models' base solutions incorporate different assumptions about the evolution of public debt in Italy towards the Maastricht Treaty target of 60% of GDP. In the event the Brookings project failed to achieve a common base solution across models, and it has not been attempted here. The solution period typically begins in the recent past, and differences become more pronounced further out into the future. The responses to temporary shocks applied at the beginning of the period, on which this article focuses, are less likely to be vulnerable to differences in the baseline than the results of permanent shocks.

To analyse possible differences across countries and the consequences of aggregation the fiscal shock is applied, where possible, to the major EU economies one at a time and to the euro area as a whole, and in both cases results are presented, where possible, for individual countries and the aggregate. The area-wide shock and its area-wide consequences are all that AWM and MULTIMOD Mark IIIB can consider, while MULTIMOD Mark III describes the consequences of an area-wide shock for France, Germany, Italy and 'other' euro area, which are cumulated to give the area-wide responses. Mark III also considers country-specific shocks, and NIGEM and QUEST are operated mostly in this mode. Reflecting the introduction of a single monetary policy, the monetary shock is applied area-wide, and contrasted with a worldwide perturbation to interest rates. Again, where possible, results are presented for individual countries and the euro area as a whole.

Tables 1–4 provide a summary comparison of the GDP and inflation responses in fiscal and monetary policy simulations under the 'standard' Taylor rule discussed above. For each model the inflation measure reported is the variable that appears in the rule, either the GDP deflator (PGDP) or the consumers' expenditure deflator (CED). Other material discussed below is extracted from the modelling groups' tabulations as required, without detailed cross-referencing.

3.2. Fiscal policy simulation results

The responses of GDP to the fiscal contraction show wide variation, with impact multipliers well in excess of one on AWM and MULTIMOD, and somewhat below one on NIGEM and QUEST. These apparent pairwise similarities mask important differences, however.

Taking AWM and MULTIMOD first, we focus on the euro area concurrent responses to the area-wide shock. Expressing the components of the familiar national income identity $Y = C + I + G + (X - M)$ as percentages of baseline GDP gives the multiplier decompositions

$$\text{AWM:} -1.35 = -0.72 - 0.54 - 1.0 + 0.91$$

$$\text{MULTIMOD:} -1.48 = -0.60 - 0.01 - 1.0 + 0.13$$

With the nominal interest rate held at base during the first year, and inflation being reduced, the real short rate rises, and this has a strong direct effect on consumption and investment, both fixed investment and stockbuilding, in AWM. Each of the three endogenous components also responds directly to the fall in demand, the trade balance especially so, due to the strong response of imports to lower GDP, although the inclusion of intra-area trade raises questions about the quantification of this effect. With respect to investment, MULTIMOD's intertemporal optimisers recognise the temporary nature of the perturbation, and scarcely change their actions. Forward-

Table 1
Fiscal shock, real GDP responses (percent deviation from base)

	France	Germany	Italy	Euro area
Area-wide shock				
AWM				
Year 1				−1.35
2				−0.52
3				0.09
5				0.22
10				0.03
MULTIMOD III				
Year 1	−1.46	−1.54	−1.52	−1.48
2	0.22	0.21	0.16	0.23
3	0.24	0.21	0.17	0.24
5	0.15	0.09	0.08	0.11
10	−0.01	0.01	0.04	0.00
MULTIMOD IIIB				
Year 1				−1.14
2				0.11
3				0.10
5				0.05
10				0.00
Individual country shocks				
MULTIMOD III				
Year 1	−1.26	−1.33	−1.32	
2	0.25	0.25	0.27	
3	0.29	0.27	0.30	
5	0.19	0.16	0.18	
10	−0.05	−0.03	−0.05	
NIGEM				
Year 1	−0.78	−0.99	−0.67	
2	−0.15	0.08	−0.14	
3	0.02	0.08	0.01	
5	0.02	0.02	0.04	
10	0.01	0.02	0.00	
QUEST				
Year 1	−0.87	−0.86	−0.85	
2	0.25	0.14	0.27	
3	0.21	0.11	0.18	
5	0.06	0.04	0.04	
10	0.03	0.01	0.02	

Table 2
Fiscal shock, inflation responses (percentage point deviation from base)

	France	Germany	Italy	Euro area
Area-wide shock				
AWM (CED)				
Year 1				−0.18
2				−0.27
3				−0.09
5				−0.15
10				0.09
MULTIMOD III (PGDP)				
Year 1	−0.46	−0.36	−0.46	−0.47
2	−0.39	−0.33	−0.35	−0.36
3	−0.26	−0.22	−0.21	−0.19
5	−0.09	−0.05	−0.03	−0.01
10	0.01	0.01	0.03	0.00
MULTIMOD IIIB (PGDP)				
Year 1				−0.17
2				−0.14
3				−0.06
5				−0.01
10				0.00
Individual country shocks				
MULTIMOD III (PGDP)				
Year 1	−0.28	−0.22	−0.21	
2	−0.18	−0.16	−0.09	
3	−0.02	−0.05	0.01	
5	0.12	0.05	0.07	
10	−0.04	−0.02	−0.02	
NIGEM (CED)				
Year 1	−0.09	−0.09	−0.09	
2	−0.03	−0.26	−0.05	
3	−0.04	−0.04	−0.04	
5	0.00	0.06	0.03	
10	0.02	0.01	0.00	
QUEST (PGDP)				
Year 1	−0.19	−0.15	−0.20	
2	−0.09	−0.05	−0.09	
3	0.07	0.04	0.09	
5	0.03	0.01	0.02	
10	0.00	0.00	0.00	

looking consumers represent a relatively small share of consumption in MULTIMOD Mark III, while the share of liquidity constrained consumers, who consume directly out of current income, is much reduced in the more recent estimates of Mark IIIB.

Table 3
Euro area monetary shock, real GDP responses (percent deviation from base)

	France	Germany	Italy	Euro area
AWM				
Year 1				−0.19
2				−0.30
3				−0.28
5				−0.17
10				−0.01
MULTIMOD III				
Year 1	−0.17	−0.22	−0.19	−0.20
2	−0.13	−0.12	−0.25	−0.14
3	−0.03	−0.02	−0.08	−0.02
5	−0.02	−0.05	−0.03	−0.02
10	0.03	0.02	0.05	0.02
MULTIMOD IIIB				
Year 1				−0.39
2				−0.16
3				−0.04
5				−0.02
10				0.02
NIGEM				
Year 1	−0.12	−0.20	−0.10	−0.16
2	−0.22	−0.25	−0.17	−0.22
3	−0.13	−0.10	−0.11	−0.11
5	−0.05	−0.04	−0.06	−0.04
10	0.00	0.00	−0.02	0.00
QUEST				
Year 1	−0.52	−0.59	−0.57	−0.57
2	−0.11	−0.13	−0.07	−0.10
3	−0.03	−0.04	−0.02	−0.03
5	−0.02	−0.02	−0.02	−0.02
10	−0.01	−0.02	−0.01	−0.01

This change accounts for most of the reduction in the GDP multiplier, the corresponding decomposition being

$$\text{MULTIMOD IIIB:} -1.14 = -0.19 - 0.01 - 1.0 + 0.07.$$

Individual country shocks have smaller effects than the area-wide shocks, thanks to the net export position of the individual country with its European trading partners, who do not experience the decline in demand. In general the differences between models considerably exceed the differences across countries on a given model. To bring the remaining two models, with impact multipliers below one, into the comparison, we take the individual-country results for France to be typical, and these have the following national income decompositions:

Table 4
Euro area monetary shock, inflation responses (percent deviation from base)

	France	Germany	Italy	Euro area
AWM (CED)				
Year 1				−0.03
2				−0.02
3				−0.06
5				−0.13
10				−0.03
MULTIMOD III (PGDP)				
Year 1	−0.20	−0.17	−0.25	−0.21
2	−0.23	−0.20	−0.25	−0.23
3	−0.21	−0.18	−0.21	−0.19
5	−0.10	−0.12	−0.09	−0.10
10	0.02	0.00	0.02	0.01
MULTIMOD IIIB (PGDP)				
Year 1				−0.12
2				−0.20
3				−0.09
5				−0.09
10				0.00
NIGEM (CED)				
Year 1	−0.04	−0.02	−0.03	
2	0.00	−0.07	−0.02	
3	−0.03	−0.10	−0.03	
5	−0.03	−0.01	−0.02	
10	0.00	0.01	0.00	
QUEST (PGDP)				
Year 1	−0.18	−0.17	−0.19	−0.18
2	−0.11	−0.10	−0.11	−0.11
3	0.01	0.00	0.00	0.01
5	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00

$$\text{MULTIMOD}: -1.26 = -0.51 + 0.00 - 1.00 + 0.25$$

$$\text{NIGEM}: -0.78 = -0.04 - 0.35 - 1.00 + 0.61$$

$$\text{QUEST}: -0.87 = -0.09 + 0.08 - 0.97 + 0.12.$$

Model features discussed in the previous paragraph maintain their relevance. The share of liquidity constrained consumers falls from 0.44 in MULTIMOD Mark IIIB to 0.3 in QUEST, while NIGEM's consumers are entirely forward-looking. Investment in NIGEM is treated in a similar way to AWM, and is mainly affected through the demand channel, as its cost of capital measure is based on the long-term real rate of interest, which rises by a smaller amount *ceteris paribus* than the short rate used in AWM. The fiscal shock in QUEST is a proportionate reduction in government purchases and government employment, the latter resulting in a sharper rise in the unemployment rate than elsewhere. This has a negative impact on real wage costs, and while expected future profitability is little affected by this temporary shock, as in MULTIMOD, the change in current profitability induces a small investment response of the opposite sign to that observed in other models.

Cross-country differences are greatest in NIGEM. Differences in trade shares and factor shares are reflected in all models, but NIGEM has greater variety in estimated coefficients across countries, in trade and factor demand equations. Its consumption function is the same across the three countries in our tables, however. Here different consumption responses occur as different income effects induced elsewhere in the model work through in the four quarters of the first year. They are nevertheless the smallest consumption responses across the models, for the reason noted above.

Once the perturbation is removed policy reaction functions come into play to correct the imbalances that have arisen. Tax rates fall in response to the budget surplus produced by the fiscal contraction, and interest rates fall by amounts proportional to the accompanying deflation. These latter effects are greater in AWM and MULTIMOD than in NIGEM and QUEST. Again, however, MULTIMOD IIIB's more recent estimates deliver smaller effects and hence smaller policy responses, in this case due to the reduction in the persistence of inflation implied by reestimation of its Phillips curve on more recent data. In general adjustment occurs quickly, and targets and instruments are back at base well within the period covered by our tables. In NIGEM and QUEST, whose Taylor rules are forward-looking ('inflation-forecast' targeting), adjustment occurs with a small amount of overshooting. The exception is AWM, whose inflation correction is much more protracted, in the absence of forward-looking behaviour in wage and price setting in this model. In addition, the response of prices to the output gap is weaker in AWM than the 'marked capacity effects in the price system' of NIGEM.

3.3. Monetary policy simulation results

Responses to the temporary increase in interest rates in the euro area are summarised in Tables 3 and 4. Exchange rate effects are stronger in this experiment than in the case of a worldwide increase in interest rates, and our comparative analysis focuses on the euro area increase.

The first-year monetary policy multipliers are greater in QUEST than in the other three models, and across countries are greater in Germany than elsewhere. In both dimensions it is the investment response that provides the main explanation for these differences. The direct shock to interest rates serves to emphasise model differences in the treatment of investment already noted above. Although MULTIMOD and QUEST are both based on Tobin's q model of intertemporally optimising firms, in MULTIMOD changes in q have both current and lagged effects on investment, with coefficients 0.03 and 0.05, respectively. Since this is an annual model, the larger lagged effect does not feature during the period of the interest rate increase. The new estimates incorporated in MULTIMOD Mark IIIB include a doubling of the contemporaneous coefficient. These coefficient differences explain the difference in the investment response between the two versions of the model, and help to bring the overall GDP response of Mark IIIB to a value intermediate between those of MULTIMOD Mark III and QUEST. The contribution of investment to the first-year GDP multiplier in AWM and NIGEM lies between that of the other models; they take similar approaches to one another, with NIGEM's smaller responses again being explained by the use of the long-term rather than the short-term interest rate in calculating the user cost of capital. Across countries, the larger responses in Germany are mostly due to its larger share of investment in GDP.

The contribution of net exports to the first-year GDP multiplier is generally small, although this temporary shock has a permanent effect on the nominal exchange rate. The effects are largest in AWM, albeit subject to the previous reservation about the inclusion of intra-area trade. Although the remaining contributions are small there is disagreement about their sign, but we note that the net effect is the difference between two much larger quantities, each susceptible to small variations in estimated income and price elasticities.

In the second year the setting of interest rates is determined by the monetary policy reaction function, and fiscal policy now reacts to the deterioration in the public finances that has resulted from the temporary increase in interest rates. This increases the cost of debt service, the more so in Italy given its larger stock of public debt. In MULTIMOD the fiscal correction works mostly through household disposable income and hence the consumption of the large share of liquidity-constrained households in this model, and the further decline in consumption is sufficient to give a further decline in the overall GDP multiplier. This does not occur in the other countries, however, and the Italian second-year decline washes out in the euro area aggregate. In contrast, AWM and NIGEM exhibit similar additional falls in the GDP multiplier, for the area-wide aggregate in AWM, and for the three major economies in NIGEM. A further decline in investment drives this result in NIGEM, indeed the second-year (negative) investment responses are larger than the first, due again to the demand channel. AWM experiences further falls in both consumption and investment, through the same mechanism. Finally, in complete contrast, the large first-year GDP effects in QUEST decline quickly and smoothly back to base, reflecting its emphasis on optimising behaviour rather than estimated dynamics.

Interest rate trajectories in general show the similarities to be expected from the use of the same reaction coefficients in the Taylor rule representation of the single monetary policy, although in implementation there are some differences in the timing of the inflation response, as noted above. These are sufficient to explain variations in the size of the reduction in interest rates in the second year, relative to the different inflation responses to the shock. Interest rates, along with inflation, then return smoothly back to base. The exception is again AWM, where the inflationary consequences are rather small but the output gap effects rather long-lived, and interest rates continue to decline for a further 3 years. In this model the investment responses to both shocks are very persistent and oscillatory, and raise a question about its dynamic specification.

4. Conclusion

The comparisons reported in this article represent a noteworthy beginning to comparative modelling research in the new, and still evolving, economic and political environment of Europe. The simulation results quantify some important policy responses, and highlight some areas of the models where lack of agreement has important consequences for these point estimates, although perhaps also indicating a range of uncertainty around them. Nevertheless the extent of disagreement is much less than in some previous model comparison projects.

With monetary policy taking an area-wide perspective, an aggregate area-wide model appears to be adequate for monetary policy making. Analysis of its consequences is undertaken by national central banks, however, and fiscal policy remains at the national level, albeit increasingly coordinated, hence multicountry modelling remains necessary. The degree of variation across the three major euro economies studied in this exercise differs across the three multicountry models, which place different relative weights on theory and data in the process of quantification. It is easy to surmise that, however large or small it is in this exercise, cross-country variation would be increased by including the smaller euro economies.

The principal source of simulation differences across the four models is the different degree of forward-looking behaviour they incorporate in their treatment of consumption and investment decisions and the setting of wages and prices. This affects not only the dynamics of the economic responses to the policy experiments considered here, but also, in some cases, their magnitude. It suggests an agenda for comparative econometric testing, to try to answer Christ's (1975) classic question, which (if any) are right, and to reduce the disagreements described above. In their final review of models of the UK economy undertaken by the ESRC Macroeconomic Modelling Bureau, the authors note that 'Sometimes comparative testing may lead to a preferred and/or improved specification. The sensitivity of overall model properties can then be checked by replacing the various original specifications by the preferred specification and observing the impact of this change on the comparative simulation results. Sometimes the available data cannot discriminate between competing specifications, but at least the model user is then clear about where the uncertainty lies, and can base a choice on whatever other grounds may be appropriate to the particular application. This combination of simulation analysis of overall model properties and econometric analysis of individual model equations or groups of equations in the context of cross-model comparisons [has] proved to be a productive methodological development' (Church et al., 2000). Its application to the existing and emerging models of the European economy is an important next step.

Acknowledgments

The helpful comments, advice and assistance of Ray Barrell, Alistair Dieppe, Jerome Henry, Andrew Hughes Hallett, Ben Hunt, Jan in't Veld and Susanne

Mundschenk are gratefully acknowledged. Responsibility for the contents of this article and any errors therein nevertheless remains with the author.

References

Aghion, P., Howitt, P., 1994. Growth and unemployment. Rev. Econ. Stud. 61, 477–494.

Balakrishnan, R., Lopez-Salido, J.D., 2002. Understanding UK inflation: the role of openness. Working Paper No.164, Bank of England, London.

Bardsen, G., Jansen, E.S., Nymoen, R., 2002. Testing the New Keynesian Phillips curve. Working Paper 2002/05, Norges Bank, Oslo.

Barrell, R., Dury, K., Hurst, I., Pain, N., 2001. Modelling the world economy: the National Institute Global Economic Model. Presented at an ENEPRI workshop on Simulation Properties of Macroeconometric Models, CEPII, Paris.

Barrell, R., Becker, B., Byrne, J., Gottschalk, S., Hurst, I., van Welsum, D., 2004. Macroeconomic policy in Europe: experiments with monetary responses and fiscal impulses. Economic Modelling, 21, 877–931.

Bryant, R.C., Henderson, D.W., Holtham, G., Hooper, P., Symansky, S.A. (Eds.), 1988. Empirical Macroeconomics for Interdependent Economies. Brookings Institution, Washington DC.

Bryant, R.C., Hooper, P., Mann, C.L. (Eds.), 1993. Evaluating Policy Regimes: New Research in Empirical Macroeconomics. Brookings Institution, Washington DC.

Christ, C.F., 1975. Judging the performance of econometric models of the US economy. Int. Econ. Rev. 16, 54–74.

Church, K.B., Sault, J.E., Sgherri, S., Wallis, K.F., 2000. Comparative properties of models of the UK economy. Natl. Institute Econ. Rev. 171, 106–122.

Dieppe, A., Henry, J., 2004. The euro area viewed as a single economy: how does it respond to shocks? Econ. Modelling, 21, 833–875.

Fagan, G., Henry, J., Mestre, R., 2001. An Area-Wide Model (AWM) for the Euro Area. Working Paper No.42, European Central Bank, Frankfurt.

Gali, J., Gertler, M., 1999. Inflation dynamics: a structural econometric analysis. J. Monetary Econ. 44, 195–222.

Gali, J., Gertler, M., Lopez-Salido, J.D., 2001. European inflation dynamics. Eur. Econ. Rev. 45, 1237–1270.

Gorman, W.M., 1953. Community preference fields. Econometrica 21, 63–80.

Hunt, B., Laxton, D., 2004. Some simulation properties of the major euro area economies in MULTIMOD. Economic Modelling, 21, 759–783.

Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica 59, 1551–1580.

Joyce, M., Wren-Lewis, S., 1991. The role of the exchange rate and capacity utilisation in convergence to the NAIRU. Econ. J. 101, 497–507.

Klein, L.R., 1946. Macroeconomics and the theory of rational behavior. Econometrica 14, 93–108.

Laxton, D., Isard, P., Faruqee, H., Prasad, E., Turtelboom, B., 1998. MULTIMOD Mark III: the core dynamic and steady-state models. Occasional Paper No.164, International Monetary Fund, Washington DC.

Maclennan, D., Muellbauer, J., Stephens, M., 2000. Asymmetries in housing and financial market institutions and EMU. In: Jenkinson, T. (Ed.), Readings in Macroeconomics. 2nd ed. Oxford University Press, Oxford, pp. 74–98. Updated version of an article with the same title published in Oxford Review of Economic Policy, 14(3), 54-80, Autumn 1998.

Manning, A., 1992. Productivity growth, wage setting and the equilibrium rate of unemployment. Discussion Paper No.63, Centre for Economic Performance, London School of Economics.

Mitchell, P.R., Sault, J.E., Smith, P.N., Wallis, K.F., 1998. Comparing global economic models. Econ. Modelling 15, 1–48.

Mitchell, P.R., Sault, J.E., Wallis, K.F., 2000. Fiscal policy rules in macroeconomic models: principles and practice. Econ. Modelling 17, 171–193.

Muellbauer, J., Lattimore, R., 1995. The consumption function: a theoretical and empirical overview. In: Pesaran, M.H., Wickens, M.R. (Eds.), Handbook of Applied Econometrics. Blackwell, Oxford, pp. 221–311.

Nataf, A., 1948. Sur la possibilite de construction de certaines macromodeles. Econometrica 16, 232–244.

Roeger, W., in't Veld, J., 1997. QUEST II: a multi country business cycle and growth model. Working paper, European Commission DGII, Brussels.

Roeger, W., in't Veld, J., 2004. Some selected simulation experiments with the European Commission's QUEST model. Econ. Modelling, 21, 785–832.

Rudd, J., Whelan, K., 2001. New tests of the New-Keynesian Phillips curve. Finance and Economics Discussion Series, Paper No.2001-30, Federal Reserve Board, Washington DC.

Rudd, J., Whelan, K., 2002. Does the labor share of income drive inflation? Finance and Economics Discussion Series, Paper No.2002-30, Federal Reserve Board, Washington DC.

Sgherri, S., 2000. When is labour market flexibility welcome? More on asymmetric policy impacts in Europe. Research Memorandum WO and E No.619, De Nederlandsche Bank, Amsterdam.

Stoker, T.M., 1986. Simple tests of distributional effects on macroeconomic equations. J. Polit. Econ. 94, 763–795.

Theil, H., 1954. Linear Aggregation of Economic Relations. North-Holland, Amsterdam.

Turner, D.S., 1991. The determinants of the NAIRU response in simulations on the Treasury model. Oxf. Bull. Econ. Stat. 53, 225–242.

Wallis, K.F., 2000. Macroeconometric modelling. In: Gudmundsson, M., Herbertsson, T.T., Zoega, G. (Eds.), Macroeconomic Policy: Iceland in an Era of Global Integration. University of Iceland Press, Reykjavik, pp. 399–414.