9

A long-run structural model of the UK

In this chapter, we describe the estimation and testing of the core long-run model of the UK economy set out in Chapter 4. This involves the estimation of a VECM of the form defined in equation (4.46) which for convenience we reproduce here:

$$\Delta \mathbf{y}_t = \mathbf{a}_y + \boldsymbol{\alpha}_y \mathbf{b}_0 - \boldsymbol{\alpha}_y \boldsymbol{\beta}' [\mathbf{z}_{t-1} - \boldsymbol{\gamma}(t-1)] + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_{yi} \Delta \mathbf{z}_{t-i} + \boldsymbol{\psi}_{yo} \Delta p_t^o + \mathbf{u}_{yt}, \tag{9.1}$$

where $\boldsymbol{\beta}'\boldsymbol{\gamma} = \mathbf{b}_1$ in (4.46). In this specification, $\mathbf{z}_t$ is partitioned as $\mathbf{z}_t = (p_t^o, \mathbf{y}_t')'$, where $\mathbf{y}_t = (e_t, r_t^*, r_t, \Delta\tilde{p}_t, y_t, p_t - p_t^*, h_t - y_t, y_t^*)'$, $\mathbf{a}_y$ is an 8×1 vector of fixed intercepts, $\boldsymbol{\alpha}_y$ is an 8×5 matrix of error correction coefficients (also known as the loading coefficient matrix), $\boldsymbol{\Gamma}_{yi}$, $i = 1, 2, \ldots, p-1$, are 8×9 matrices of short-run coefficients, $\boldsymbol{\psi}_{yo}$ is an 8×1 vector representing the impact effects of changes in oil prices on $\Delta \mathbf{y}_t$, $\mathbf{u}_{yt}$ is an 8×1 vector of disturbances assumed to be $i.i.d.(0, \boldsymbol{\Sigma}_y)$, with $\boldsymbol{\Sigma}_y$ being a positive definite matrix, and by construction uncorrelated with u_{ot}, and $\boldsymbol{\beta}'(\mathbf{z}_{t-1} - \boldsymbol{\gamma}(t-1))$ is an $r \times 1$ vector of error correction terms. The long-run theory suggests that $r = 5$, but our approach tests the hypothesis of $r = 5$ against alternative values for r.

The above specification embodies the economic theory's long-run predictions by construction, in contrast to the more usual approach where the starting point is an unrestricted VAR model, with some vague priors about the nature of the long-run relations. By including the trend inside the error correction term, the deterministic trend properties of the model do not change with the number of cointegrating vectors, r.

9.1 The different stages of estimation and testing

As a general guide to the application of the econometric techniques described in Chapter 6, and as a precursor to our own analysis of the core model, we now describe the sequence of steps we followed in our empirical work. Note that in order to incorporate the long-run relationships into a suitable model, as defined above, it is important that the variables used in the empirical analysis can be reasonably argued to be $I(1)$. Hence the preliminary stage in any analysis is to establish the orders of integration of the variables in the vector z_t and we do this in the next section.[1] Following on from this, we can identify five stages of the estimation procedure.

First, a sequence of unrestricted VAR(p), $p = 0, 1, 2, \ldots, 6$ models are estimated over the same sample period, 1965q1–1999q4. The maximum lag order, 6, is in some sense arbitrary, but is chosen *a priori* bearing in mind the quarterly nature of the observations, and the size of the available sample (namely, 140 quarterly observations). The order of VAR model to be used in the analysis is then selected in the light of the Akaike Information Criterion (AIC) and the Schwarz Bayesian Criterion (SBC).

Second, having established the appropriate order of the VAR model, cointegration tests are carried out using the trace and the maximum eigenvalue statistics, reviewed in Chapter 6. The results of these tests can be inconclusive. So the test results need to be carefully interpreted in conjunction with the theory's prediction described in Chapter 4, before a decision is made concerning the number of the cointegrating relations that are most likely to exist among the variables under investigation.

Third, having decided that there exist, say, r cointegrating vectors among the variables, we are in a position to estimate an exactly identified set of long-run relations, in which r^2 restrictions are imposed on the cointegrating vectors (r restrictions on each of the r vectors). In one sense, the choice of the exactly identifying restrictions is arbitrary: the maximised value of the log-likelihood of the system will be the same irrespective of how the long-run relations are exactly identified. In another sense, however, the choice of exactly identifying restrictions is crucial, as it provides the basis for the development of an econometric model with economically meaningful long-run properties. It is therefore important that the cointegrating

[1] It is, however, important to note that in testing the rank of the cointegrating space, it is not necessary that the underlying variables should all be $I(1)$. The problem arises in interpreting the long-run relations; since an $I(0)$ variable can be viewed trivially as forming a cointegrating relationship with the other variables using $\beta = (0, \ldots, 0, 1, 0, \ldots, 0)$ as a cointegrating vector, with the non-zero element attached to the $I(0)$ variable in question.

relations are exactly identified by imposing restrictions that are a subset of those suggested by economic theory. It is also a good practice to avoid using doubtful theory restrictions as exact identifying restrictions. Estimation of the parameters of the core model, (9.1), can be carried out using the long-run structural modelling approach in described in Chapter 6.

The *fourth* step in the analysis considers the imposition and testing of over-identifying restrictions on the cointegrating vectors, as predicted by economic theory. This analysis is carried out along the lines set out in Pesaran and Shin (2002) and Pesaran, Shin and Smith (2000) and involves the ML estimation of the model subject to the exactly and over-identifying restrictions. The tests of over-identifying restrictions will now be in the form of the familiar χ^2 tests with degrees of freedom equal to the number of the over-identifying restrictions. It is worth noting that this is a system-estimation procedure, and the likelihood function in terms of the cointegrating vectors can be quite complicated, so that the existence of local maxima cannot be ruled out, and the search for the global maximum might be difficult. To avoid convergence problems, it is often advisable to impose over-identifying restrictions one-at-a-time and, as far as possible, in a sequence that can be meaningfully interpreted so that information can be obtained on which of the restrictions is more or less likely to be accepted by the data.[2] Another possibility would be to start from fully specified long-run relationships and then relax some of the theory restrictions one at a time.

The *fifth* step in the analysis concerns the interpretation of the results. The imposition of long-run, theory-based restrictions yield error correction terms that can be interpreted as characterising disequilibria in particular markets, and the associated error correction regressions show the short-run evolution of the variables in the model in response to deviations from equilibrium and to past changes in the variables of the model. The error correction regressions are also subjected to diagnostic tests for residual serial correlation, non-normal errors, functional form misspecification, and heteroscedasticity as is usual in the case of standard regression analysis. The magnitudes of some of the estimated regression coefficients provide useful information on the dynamics of the system, highlighting which of the variables have large and statistically significant effects on each other, although care needs to be exercised in the interpretation of

[2] The interpretation of this sequence of restriction tests should also be sensitive to the fact that asymptotic critical values of over-identifying restrictions tend to over-reject when applied to small samples, and in some cases by a large amount, as discussed earlier in Section 6.4.

the coefficients on the error correction terms as far as the stability of the system as a whole is concerned (as discussed in Chapter 6).

As part of assessing the model we would also need to analyse its dynamics. This involves the use of persistence profiles, impulse responses and probability forecasting which we discuss in Chapters 10 and 11.

9.2 Unit root properties of the core variables

Before the estimation of the model can begin, it is important that the unit root properties of the variables under investigation are established to enable sensible interpretation of the long-run relations. The limitations of the standard tests for unit roots (such as the Dickey and Fuller (1979) or the Phillips and Perron (1988) tests) are well-known, but they nevertheless provide important information on the nature of the persistence of the time series under investigation. For example, it might be difficult to come to a clear-cut conclusion over whether the effects of a shock to a particular variable take a long while to die away (for an $I(0)$ variable) or whether they will never die away (for an $I(1)$ variable) using existing tests and given the limited data available. Even such an ambiguous conclusion can be helpful, however, as it suggests that certain variables are on the borderline of being $I(0)/I(1)$ or $I(1)/I(2)$. For example, one might assume that a given variable is $I(1)$, perhaps on the basis of *a priori* economic reasoning, and subsequently carry out tests to establish the number of cointegrating relations between this and other $I(1)$ variables. The knowledge that this variable is close to being stationary, when considered in isolation, means that the tests of the number of cointegrating relationships are likely to support the presence of a higher number of cointegrating relations than would be the case if the variable in question was clearly $I(1)$.

The results of the Augmented Dickey–Fuller (ADF) and Phillips–Perron (PP) tests, computed over the sample period for the levels and first differences of the core variables, are reported in Tables 9.1a and 9.1b.

Both sets of tests provide relatively strong support for the view that y_t, y_t^*,r_t, r_t^*, e_t, $(h_t - y_t)$ and p_t^o are $I(1)$ series. The unit root hypothesis is clearly rejected when applied to the first differences of these variables, but there is no evidence with which to reject the unit root hypothesis when the tests are applied to the levels. There is, however, some ambiguity regarding the order of integration of the price variables. Application of the ADF test to Δp_t, $\Delta\tilde{p}_t$ and Δp_t^* yields mixed results: the hypothesis that there is a unit root in the domestic and foreign inflation rates

Table 9.1a Augmented Dickey–Fuller unit root tests applied to variables in the core model, 1965q1–1999q4.

Variable	ADF(0)	ADF(1)	ADF(2)	ADF(3)	ADF(4)
(i) For the first differences					
Δy_t	−11.94	−8.06	-5.40^a	−5.18	−4.81
Δy_t^*	−7.43	-5.28^a	−4.53	−4.22	−4.11
Δr_t	−10.54	−7.79	−7.49	−6.08	-6.30^a
Δr_t^*	-7.06^a	−6.19	−4.89	−4.85	−4.54
Δe_t	-9.81^a	−7.89	−6.45	−5.52	−5.39
$\Delta(h_t - y_t)$	-12.21^a	−8.16	−5.79	−4.82	−3.13
Δp_t	-3.50^a	−3.19	−2.67	−2.44	−2.43
$\Delta\tilde{p}_t$	−4.41	-3.05^a	−2.97	−2.42	−2.23
Δp_t^*	−5.06	−3.47	-2.73^a	−2.75	−2.90
Δp_t^o	-11.05^a	−8.71	−6.41	−5.68	−5.71
$\Delta^2 p_t$	−13.32	-10.74^a	−8.80	−7.15	−6.95
$\Delta^2\tilde{p}_t$	−17.37	−10.22	−9.48	−8.06	−7.43
$\Delta^2 p_t^*$	−17.82	-12.75^a	−8.63	−6.65	−6.43
$\Delta(p_t - p_t^*)$	−6.69	−4.91	-3.72^a	−3.60	−3.32
(ii) For the levels					
y_t	−2.32	−2.33	−2.46	-3.14^a	−3.06
y_t^*	−3.37	−3.18	-3.22^a	−3.24	−3.22
r_t	−2.23	-2.57^a	−2.65	−2.32	−2.46
r_t^*	−1.24	-2.47^a	−2.46	−2.86	−2.71
e_t	−1.03	-1.45^a	−1.32	−1.33	−1.37
$h_t - y_t$	1.41	1.82^a	2.00	1.83	1.86
p_t	2.21	-0.39^a	−0.48	−0.76	−0.90
$\tilde{p}_t$	2.12	−0.03	-0.61^a	−0.57	−0.88
p_t^*	1.83	−0.07	−0.73	-1.20^a	−1.13
p_t^o	-1.43^a	−1.53	−1.38	−1.49	−1.44
$p_t - p_t^*$	0.47	−0.40	−0.66	-1.01^a	−0.96

Note: When applied to the first differences, augmented Dickey–Fuller (1979, ADF) test statistics are computed using ADF regressions with an intercept and p lagged first differences of dependent variable, while when applied to the levels, ADF statistics are computed using ADF regressions with an intercept, a linear time trend and p lagged first differences of dependent variable, with the exception of the following variables: r_t and r_t^* where only an intercept was included in the underlying ADF regressions. The relevant lower 5% critical values for the ADF tests are −2.88 for the former and −3.45 for the latter. The symbol 'a' denotes the order of augmentation in the Dickey–Fuller regressions chosen using the Akaike Information Criterion, with a maximum lag order of four.

is rejected for low orders of augmentation (namely, for $p = 0$ and 1), but not for higher orders. The application of the PP test rejects the unit root hypothesis when applied to Δp_t^*, Δp_t and $\Delta\tilde{p}_t$. Overall the available data is not informative as to whether domestic and foreign prices are $I(1)$ or $I(2)$.

These preliminary results regarding the unit roots properties of the core variables raise interesting issues concerning the use of economic theory

Table 9.1b Phillips and Perron unit root tests applied to variables in the core model, 1965q1–1999q4.

Variable	PP(0)	PP(5)	PP(10)	PP(15)	PP(20)
(i) For the first differences					
Δy_t	−11.94	−12.00	−12.02	−11.95	−11.98
Δy_t^*	−7.43	−7.68	−7.73	−7.72	−7.80
Δr_t	−10.54	−10.50	−10.51	−10.67	−11.41
Δr_t^*	−7.06	−7.14	−6.84	−6.43	−6.27
Δe_t	−9.81	−9.75	−9.76	−9.81	−9.73
$\Delta(h_t - y_t)$	−12.21	−12.28	−12.55	−12.86	−13.22
Δp_t	−3.05	−3.30	−3.38	−3.64	−3.78
$\Delta \tilde{p}_t$	−4.41	−4.22	−4.70	−5.04	−5.32
Δp_t^*	−5.06	−5.07	−5.52	−5.91	−6.16
Δp_t^o	−11.05	−11.03	−11.03	−11.03	−11.03
$\Delta^2 p_t$	−13.32	−5.99	−4.49	−4.03	−2.76
$\Delta^2 \tilde{p}_t$	−17.37	−19.84	−21.27	−23.83	−25.96
$\Delta^2 p_t^*$	−17.82	−20.07	−22.92	−24.82	−28.66
$\Delta(p_t - p_t^*)$	−6.69	−6.96	−7.53	−7.87	−8.01
(ii) For the levels					
y_t	−2.32	−2.70	−2.84	−2.70	−2.47
y_t^*	−3.37	−3.07	−3.08	−3.14	−3.22
r_t	−2.23	−2.45	−2.37	−2.24	−2.02
r_t^*	−1.24	−2.13	−2.03	−1.72	−1.54
e_t	−1.03	−1.29	−1.29	−1.35	−1.17
$h_t - y_t$	1.41	1.90	1.85	1.87	1.83
p_t	2.21	0.43	0.01	−0.22	−0.36
$\tilde{p}_t$	2.12	0.45	0.02	−0.18	−0.31
p_t^*	1.83	−1.45	−1.43	−1.46	−1.45
p_t^o	−1.43	−1.45	−1.43	−1.46	−1.45
$p_t - p_t^*$	0.47	−0.43	−0.69	−0.78	−0.79

Note: $PP(\ell)$ represents Phillips and Perron (1988) unit root statistic based on the Bartlett window of size ℓ. In the first difference equations, *PP* test statistics are obtained including only an intercept in the underlying *DF* regressions; in the levels equations, *PP* test statistics are obtained including an intercept and a time trend in the underlying *DF* regressions, with the exception of the following variables; r_t and r_t^* where no trend is included. The relevant lower 5% critical values are −2.88 for the first difference equations, and −3.45 for the levels equations.

and statistical evidence in macroeconometric modelling. Starting from the long-run theory set out in Chapter 4, the validity of the Fisher equation requires that inflation and interest rates have the same order of integration. The theoretical literature generally assumes that these series are $I(0)$, but as we have seen above the empirical evidence is mixed with the interest rate behaving as an $I(1)$ variable and the inflation rate being a borderline case.[3] There is, therefore, a trade-off between the demands of

[3] In this book we are confining the modelling exercise to log-linear specifications and a more complicated non-linear model might be needed for interest rates and inflation, as argued,

theory and econometrics. Our approach to this dilemma is a pragmatic one, aiming to adequately capture the statistical properties of the data in a modelling framework which, at the same time, is coherent with our underlying analytic account of how the economy operates. For these reasons, in our work, we treat r_t, r_t^*, Δp_t, $\Delta \tilde{p}_t$ and Δp_t^* as $I(1)$ variables. This allows the empirical model to adequately represent the statistical features of the series over the sample period and provides the scope for accommodating in the model the long-run relationships described in Chapter 4.

Of course, domestic and foreign prices appear in their *level* in the PPP relationship of (4.35) and this raises the potential difficulty of mixing $I(1)$ and $I(2)$ variables. Haldrup's (1998) review of the econometric analysis of $I(2)$ variables warns of the dangers of the inappropriate application of econometric methods designed for use with $I(1)$ variables and suggests that it is often useful to transform time series *a priori* to obtain variables that are unambiguously $I(1)$ rather than dealing with mixtures of $I(1)$ and $I(2)$ variables directly. In the case of the core variables under consideration, this is achieved by working with the relative price variables $p_t - p_t^*$ rather than the two price levels p_t and p_t^* separately. As shown in Table 9.1a, the relative price term is unambiguously $I(1)$ according to the ADF statistics.

The decision to include domestic prices in the model in two forms, $(p_t - p_t^*)$ and $\Delta \tilde{p}_t$ does not create difficulties of inconsistency either algebraically or economically (and would not do so even if we used Δp_t in place of $\Delta \tilde{p}_t$ in the model). Ignoring the distinction between p_t and $\tilde{p}_t$ for the moment, we note that the associated structural model of (5.2) contains nine equations in eight endogenous variables. One of the nine equations corresponds to the determination of domestic prices p_t and one corresponds to the determination of foreign prices p_t^* and this is entirely consistent with the fact that the domestic price variable influences the relative price variable and the inflation variable when the model is estimated. Further, there is considerable evidence, both on the basis of our own analysis and elsewhere, that the various alternative measures of inflation that are available are pairwise cointegrated with a cointegrating vector of $(1, -1)$ and a zero constant. The use of two measures of prices, p_t and $\tilde{p}_t$, in the analysis has no impact on the long-run properties of the

for example, in Pesaran, Timmermann and Pettenuzzo (2004). Such an approach is worth considering but lies outside the scope of the present work.

model, therefore, but is likely to capture the short-run dynamics more accurately.

In summary, then, we can say that it seems appropriate to view all nine variables of $z_t = (p_t^o, e_t, r_t^*, r_t, \Delta\widetilde{p}_t, y_t, p_t - p_t^*, h_t - y_t, y_t^*)'$ as approximately $I(1)$ on the basis of the unit root statistics reported. We therefore conducted our analysis on this basis, although the ambiguity regarding the $\Delta\widetilde{p}_t$ variable needs to be borne in mind in interpreting the subsequent results.

9.3 Testing and estimating of the long-run relations

The first stage of our modelling sequence is to select the order of the underlying VAR using AIC and SBC reported in Table 9.2.

Here we find that a VAR of order two appears to be appropriate when using the AIC as the model selection criterion, but not surprisingly that the SBC favours a VAR of order one. We proceed with the cointegration analysis using a VAR(2), on the grounds that the consequences of overestimation of the order of the VAR are much less serious than underestimating it; see Kilian (2002).[4]

Using a VAR(2) model with unrestricted intercepts and restricted trend coefficients, and treating the oil price variable, p_t^o, as a weakly exogenous $I(1)$, or long-run forcing, variable, we computed Johansen's 'trace'

Table 9.2 Akaike and Schwarz Information Criteria for lag order selection.

Lag length	Log likelihood	AIC	SBC
6	4641.2	4155.2	3440.4
5	4538.7	4133.7	3538.0
4	4459.0	4135.0	3658.4
3	4389.0	4146.0	3788.5
2	4326.0	4164.0	3925.7
1	4222.3	4141.3	4022.2
0	1775.6	1775.6	1775.6

[4] Note that, if the dimension of the VAR is large, then a relatively low lag order can be selected and still accommodate rich dynamic specifications at the level of individual series. Specifically, if the $m \times 1$ vector z_t follows a p-order autoregression, then in general the individual elements follow an $ARMA\,(mp, mp-p)$ process. See Hamilton (1994, p. 349). In our application where $m = 9$ and $p = 2$, the univariate representation of the individual series could be $ARMA(18, 16)$.

Table 9.3 Cointegration rank test statistics for the core model, $(p_t - p_t^*, e_t, r_t, r_t^*, y_t, y_t^*, h_t - y_t, \Delta\widetilde{p}_t, p_t^o)$.

H_0	H_1	Test statistic	95% Critical values	90% Critical values
(a) Trace statistic				
$r=0$	$r=1$	324.75	199.12	192.80
$r\leq 1$	$r=2$	221.16	163.01	157.02
$r\leq 2$	$r=3$	161.88	128.79	123.33
$r\leq 3$	$r=4$	116.14	97.83	93.13
$r\leq 4$	$r=5$	78.94	72.10	68.04
$r\leq 5$	$r=6$	48.71	49.36	46.00
$r\leq 6$	$r=7$	22.46	30.77	27.96
$r\leq 7$	$r=8$	6.70	15.44	13.31
(b) Maximum eigenvalue statistic				
$r=0$	$r=1$	103.59	58.08	55.25
$r\leq 1$	$r=2$	59.27	52.62	49.70
$r\leq 2$	$r=3$	45.75	46.97	44.01
$r\leq 3$	$r=4$	37.20	40.89	37.92
$r\leq 4$	$r=5$	30.23	34.70	32.12
$r\leq 5$	$r=6$	26.25	28.72	26.10
$r\leq 6$	$r=7$	15.76	22.16	19.79
$r\leq 7$	$r=8$	6.70	15.44	13.31

Note: The underlying VAR model is of order 2 and contains unrestricted intercepts and restricted trend coefficients, with p_t^o treated as an exogenous $I(1)$ variable. The statistics refer to Johansen's log-likelihood-based trace and maximal eigenvalue statistics and are computed using 140 observations for the period 1965q1–1999q4. The asymptotic critical values are taken from Pesaran, Shin and Smith (2000).

and 'maximal eigenvalue' statistics.[5] These statistics, together with their associated 90% and 95% critical values, are reported in Table 9.3.

The maximal eigenvalue statistic indicates the presence of just two cointegrating relationships at the 5% significance level, which does not support our *a priori* expectations of five cointegrating vectors. However, as shown by Cheung and Lai (1993), the maximum eigenvalue test is generally less robust to the presence of skewness and excess kurtosis in the errors than the trace test. Given that we have evidence of non-normality in the residuals of the VAR model used to compute the test statistics, we therefore believe it is more appropriate to base our cointegration tests on the trace statistics. As it happens the trace statistics reject the null hypotheses that $r = 0, 1, 2, 3$ and 4 at the 5% level of significance but cannot reject the null hypothesis that $r = 5$. This is in line with our *a priori* expectations based on the long-run theory of Chapter 4, which suggests the

[5] An account of the algorithms used for the computation of cointegration test statistics in the presence of $I(1)$ exogenous variables can be found, for example, in Pesaran, Shin and Smith (2000).

existence of five possible long-run relations, reproduced below for ease of exposition:

$$p_t - p_t^* - e_t = b_{10} + b_{11}t + \xi_{1,t+1} \tag{9.2}$$

$$r_t - r_t^* = b_{20} + \xi_{2,t+1} \tag{9.3}$$

$$y_t - y_t^* = b_{30} + \xi_{3,t+1} \tag{9.4}$$

$$h_t - y_t = b_{40} + b_{41}t + \beta_{44}r_t + \beta_{46}y_t + \xi_{4,t+1} \tag{9.5}$$

$$r_t - \Delta p_t = b_{50} + \xi_{5,t+1}. \tag{9.6}$$

Proceeding under the assumption that there are five cointegrating vectors, the five long-run relations of the core model, (9.2)–(9.6), can be written more compactly as

$$\boldsymbol{\xi}_t = \boldsymbol{\beta}_{TH}' \mathbf{z}_{t-1} - \mathbf{b}_0 - \mathbf{b}_1(t-1), \tag{9.7}$$

where

$$\mathbf{b}_0 = (b_{10}, b_{20}, b_{30}, b_{40}, b_{50})',$$
$$\mathbf{b}_1 = (b_{11}, 0, 0, b_{41}, 0)',$$
$$\boldsymbol{\xi}_t = (\xi_{1t}, \xi_{2t}, \xi_{3t}, \xi_{4t}, \xi_{5t})',$$

and

$$\boldsymbol{\beta}_{TH}' = \begin{pmatrix} 0 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & -\beta_{44} & 0 & -\beta_{46} & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{9.8}$$

The matrix $\boldsymbol{\beta}_{TH}'$, as described in equation (9.8), imposes all the restrictions necessary to correspond to the long-run relationships and as such is *over-identified*. However, the first step in the estimation is to exactly identify the long run, which with five cointegrating relations requires five restrictions on each relationship. In view of the underlying long-run theory as encapsulated in the relations (9.2)–(9.6), we impose 25 *exactly* identifying restrictions on the cointegrating matrix (in the form of five restrictions on each of the five cointegrating vectors) so that the exactly identified

cointegrating matrix is given by:

$$\boldsymbol{\beta}_{EX}' = \begin{pmatrix} \beta_{11} & \beta_{12} & 0 & 0 & \beta_{15} & 0 & 1 & \beta_{18} & 0 \\ \beta_{21} & 0 & \beta_{23} & 1 & \beta_{25} & 0 & 0 & 0 & \beta_{29} \\ \beta_{31} & 0 & 0 & 0 & 0 & 1 & \beta_{37} & \beta_{38} & \beta_{39} \\ \beta_{41} & 0 & 0 & -\beta_{44} & \beta_{45} & -\beta_{46} & 0 & 1 & 0 \\ \beta_{51} & 0 & 0 & \beta_{54} & -1 & 0 & 0 & \beta_{58} & \beta_{59} \end{pmatrix}. \tag{9.9}$$

The first vector (the first row of $\boldsymbol{\beta}_{EX}'$) relates to the purchasing power parity (*PPP*) relationship defined by (9.2) and is normalised on $p_t - p_t^*$; the second relates to the interest rate parity (*IRP*) relationship defined by (9.3) and is normalised on r_t; the third relates to the 'output gap' (*OG*) relationship defined by (9.4) and is normalised on y_t;[6] the fourth is the money market equilibrium condition (*MME*) defined by (9.5) and is normalised on $h_t - y_t$.; and the fifth is the real interest rate relationship (*FIP*) defined by (9.6), normalised on $\Delta\tilde{p}_t$.

Having exactly identified the long-run relations, we then tested the over-identifying restrictions predicted by the long-run theory. There are 20 unrestricted parameters in (9.9) and, based on the theory restrictions as set out in (9.8), there are 18 theory-based over-identifying restrictions that could be tested. Note that the theory does not restrict two of the parameters of the money demand equation (β_{44} and β_{46}) in the fourth row of $\boldsymbol{\beta}_{TH}$ defined by (9.8). In addition, working with a cointegrating VAR with restricted trend coefficients (as described in Sections 6.2.1 and 6.2.3), there are potentially five further parameters on the trend terms in the five co-integrating relationships. There is no economic rationale for including time trends in the *IRP*, *FIP* or *OG* relationships, and the imposition of zeros on the trend coefficients in these relationships provides a further three over-identifying restrictions. The absence of a trend in the *PPP* relationship is also consistent with the theory of Chapter 4, as is the restriction that $\beta_{46} = 0$ (so that equation (9.5) is effectively a relationship explaining the velocity of circulation of money). Hence, once the long-run theory is fully imposed, there are just two parameters to be freely estimated in the cointegrating relationships, and there are a total of 23 over-identifying restrictions on which the core model is based and with which the validity of the long-run economic theory can be tested.

[6] Our use of the term 'output gap relationship' to describe (9.4) should not be confused with the more usual use of the term which relates more specifically to the difference between a country's actual and potential output levels (although clearly the two uses of the term are related and, for some open economies, the foreign output variable might provide a good proxy for potential output).

9.3.1 *Small sample properties of the tests of restrictions on the cointegrating vectors*

When testing the linear restrictions implied by our long-run theory, we need to take account of the relatively small sample size available. This issue arises in our example despite having 140 quarterly observations as we are investigating the properties of a large dimensioned VARX model subject to a large number of over-identifying restrictions. In order to deal with the small sample bias, we apply the methods described in Section 6.4. These methods involve a bootstrapping exercise to investigate and accommodate the small sample properties of the log-likelihood ratio (LR) test of over-identifying restrictions, generating a simulated distribution for the test statistic when only a small sample is available and using this to derive appropriate critical values against which to compare the estimated test statistic.

Specifically, the LR test for jointly testing the 23 over-identifying restrictions described above and implied by our long-run theory takes the value 71.49. To compute appropriate small sample critical values, we adopt a bootstrap procedure based on 3000 replications of the LR statistic testing the 23 restrictions. For each replication, an artificial dataset is generated (of the same length as the original dataset) on the assumption that the estimated version of the core model is the true data-generating process, using the observed initial values of each variable, the estimated model, and a set of random innovations. These innovations can be obtained as draws from a multivariate normal distribution chosen to match the observed correlation of the estimated reduced form errors (termed a 'parametric bootstrap') or by re-sampling with replacement from the estimated residuals (a 'non-parametric bootstrap'). In the light of the evidence of non-normality of residuals that we found in estimation, we apply the non-parametric bootstrap in this exercise (see Chapter 7 for further details). For each simulated dataset, the cointegrating VAR is estimated first subject to the exactly identifying restrictions of (9.9) and then subject to the over-identifying restrictions of (9.8).[7] The LR test of the over-identifying restrictions is carried out on each of the replicated datasets and the empirical distribution of the test statistic is derived across all replications.

[7] Given the complexity of the likelihood in the over-identified case, the choice of the optimisation algorithm to be used in maximising the likelihood may be important in this exercise. We found the Simulated Annealing routine by Goffe *et al.* (1994) to be useful. The simulated annealing algorithm explores a function's entire surface and tries to optimise the function while moving both uphill and downhill. It is therefore largely independent of starting values, and it can escape local minima and go on to find the global optimum by the uphill and downhill moves. Simulated annealing also makes less stringent assumptions on the form of the function than conventional algorithms and can therefore deal more easily with functions that have ridges and plateaux. Hence it is less likely to fail on difficult functions and is more robust than conventional Newton–Raphson and David–Fletcher–Powell uphill-only algorithms.

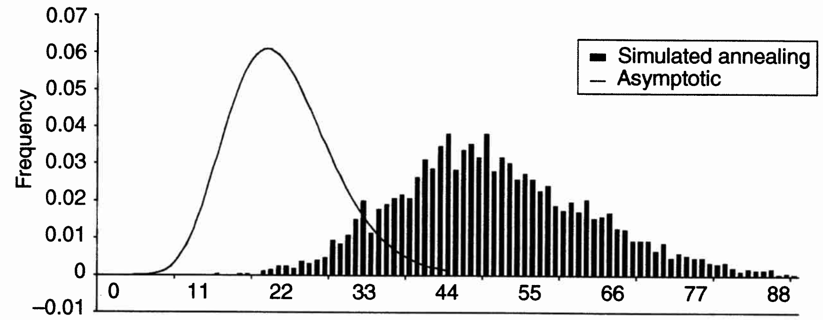

Figure 9.1 Asymptotic and empirical distribution generated by the simulated annealing algorithm of the test of the long-run over-identifying restrictions.

Figure 9.1 illustrates the empirical distribution obtained in this way, plotting this alongside the corresponding asymptotic χ^2_{23} distribution. The figure shows the empirical distribution of the test statistic lies substantially to the right of its asymptotic counterpart, demonstrating clearly the need for taking into account the small sample in this instance.

The bootstrapped critical values for the joint tests of the 23 over-identifying restrictions are 67.51 at the 10% significance level and 73.19 at the 5% level. Using these bootstrapped critical values, the 23 theory restrictions cannot be rejected at the conventional 5% level. Moreover, it is worth noting that the simulation is used to find the probability of rejection for one point in H_0, taking the estimated parameters of the core model as given. The classical significance level is the *maximum* of the rejection probabilities over H_0. So, by using a single point, the observed critical values potentially understate the true rejection level. The fact that we (almost) fail to reject at the 5% level might provide more compelling evidence to support the validity of the restrictions than it first appears therefore.

9.4 The vector error correction model

9.4.1 *The long-run estimates*

The estimates of the long-run relations and the reduced form error correction specification are provided in Table 9.4 below. The long-run relations, which incorporate all the restrictions suggested by the theory in Chapter 4, are summarised below:

$$(p_t - p_t^*) - e_t = 4.588 + \widehat{\xi}_{1,t+1} \tag{9.10}$$

$$r_t - r_t^* = 0.0058 + \widehat{\xi}_{2,t+1} \tag{9.11}$$

$$y_t - y_t^* = -0.0377 + \widehat{\xi}_{3,t+1} \tag{9.12}$$

$$h_t - y_t = -0.0538 - \underset{(22.2844)}{56.0975}\, r_t - \underset{(0.0012)}{0.0073}\, t + \widehat{\xi}_{4,t+1} \tag{9.13}$$

$$r_t - \Delta\widetilde{p}_t = 0.0036 + \widehat{\xi}_{5,t+1}. \tag{9.14}$$

The bracketed figures are asymptotic standard errors. The first equation, (9.10), describes the *PPP* relationship and the failure to reject this in the context of our core model provides an interesting empirical finding. Of course, there has been considerable interest in the literature examining the co-movements of exchange rates and relative prices, and the empirical evidence on *PPP* appears to be sensitive to the dataset used and the way in which the analysis is conducted. For example, the evidence of a unit root in the real exchange rate found by Darby (1983) and Huizinga (1988) contradicts *PPP* as a long-run relationship, while Grilli and Kaminsky (1991) and Lothian and Taylor (1996) have obtained evidence in favour of rejecting the unit root hypothesis in real exchange rates using longer annual series. In work investigating *PPP* using cointegration analysis, the results seem to be sensitive to whether the model is a trivariate one (including e_t, p_t and p_t^* in the VAR as separate variables) or a bivariate one (including e_t and $(p_t - p_t^*)$ as two separate variables). The null of no cointegration is rejected more frequently in trivariate than in bivariate analyses.[8] The finding here that *PPP* can be readily incorporated into the model is a useful contribution to this literature, indicating that the empirical evidence to support the relationship is stronger in a more complete model of the macroeconomy incorporating feedbacks and interactions omitted from more partial analyses.

The second cointegrating relation, defined by (9.11), is the *IRP* condition. This includes an intercept, which can be interpreted as the deterministic component of the risk premia associated with bonds and foreign exchange uncertainties. Its value is estimated at 0.0058, implying a risk premium of approximately 2.3% per annum. The empirical support we find for the *IRP* condition is in accordance with the results obtained in the literature, and is compatible with *UIP*, defined by (4.14). However, under

[8] See Taylor (1988) and Mark (1990) for illustrations of further work in this area, and Froot and Rogoff (1995) and MacDonald (1995) for a review of the literature.

the *UIP* hypothesis it is also required that a regression of $r_t - r_t^*$ on $\Delta \ln(E_{t+1})$ has a unit coefficient, but this is not supported by the data.

The third long-run relationship, given by (9.12), is the *OG* relationship with per capita domestic and foreign output (measured by the total OECD output) levels moving in tandem in the long run. It is noteworthy that the co-trending hypothesis cannot be rejected; *i.e.* the coefficient of the deterministic trend in the output gap equation is zero. This suggests that average long-run growth rate for the UK is the same as that in the rest of the OECD. This finding seems, in the first instance, to contradict some of the results obtained in the literature on the cointegrating properties of real output across countries. Campbell and Mankiw (1989), Cogley (1990) and Bernard and Durlauf (1995), for example, consider cointegration among international output series and find little evidence that outputs of different *pairs* of countries are cointegrated. However, our empirical analysis, being based on a single foreign output index, does not necessarily contradict this literature, which focuses on pairwise cointegration of output levels. The hypothesis advanced here, that y_t and y_t^* are cointegrated, is much less restrictive than the hypothesis considered in the literature that *all* pairs of output variables in the OECD are cointegrated.[9]

For the *MME* condition, given by (9.13), we could not reject the hypothesis that the elasticity of real money balances with respect to real output is equal to unity, and therefore (9.13) in fact represents an M0 velocity equation. The *MME* condition, however, contains a deterministic downward trend, representing the steady decline in the money–income ratio experienced in the UK over most of the period 1965–1999, arising primarily from the technological innovations in financial intermediation. There is also strong statistical evidence of a negative interest rate effect on real money balances. This long-run specification is comparable to recent research on the determinants of the UK narrow money velocity reported in, for example, Breedon and Fisher (1996).

Finally, the fifth equation, (9.14), defines the *FIP* relationship, where the estimated constant implies an annual real rate of return of approximately 1.44%. While the presence of this relationship might appear relatively uncontentious, there is empirical work in which the relationship does not seem to be supported by the empirical evidence; see, for example MacDonald and Murphy (1989) and Mishkin (1992). In La Cour and

[9] See Lee (1998) for further discussion of cross-country interdependence in growth dynamics. Pesaran (2004a) also provides an analysis of pairwise output gaps, showing that output convergence is not generally supported by the time series observations.

MacDonald (2000), evidence of a cointegrating relationship between interest rates and inflation was obtained in an analysis of financial data series from the euro area and US. However, the *FIP* relationship itself, with coefficients of $(1, -1)$ on the interest rate and inflation, was observed in the two zones only when the financial variables were incorporated into a larger macrosystem. Our results support the *FIP* relationship and again highlight the important role played by the *FIP* relationship in a model of the macroeconomy which can incorporate interactions between variables omitted from more partial analyses.

9.4.2 *Error correction specifications*

The short-run dynamics of the model are characterised by the eight error correction specifications given in Table 9.4.

The estimates of the error correction coefficients show that the long-run relations make an important contribution in most equations and that the error correction terms provide for a complex and statistically significant set of interactions and feedbacks across commodity, money and foreign exchange markets. The results in Table 9.4 also show that the core model fits the historical data well and has satisfactory diagnostic statistics. The diagnostic statistics of the equations in Table 9.4 are generally satisfactory as far as the tests of the residual serial correlation, functional form and heteroscedasticity are concerned. The assumption of normally distributed errors is rejected in all the error correction equations which is understandable if we consider the three major hikes in oil prices experienced during the estimation period and the special events that have afflicted the UK economy such as the three-day week, coal miners' strikes, the stock market crash of 1987 just to mention a few.

Figures 9.2a–9.2h plot the actual and fitted values for the reduced form error correction equations reported in Table 9.4.

These figures illustrate the extent to which the model fit the historical series. As might be expected, the exchange rate and domestic interest rate equations appear to have least explanatory power, with $\bar{R}^2$ of 0.07 and 0.12 respectively, and the model struggles to fit the observations of variables associated with the unusual events described above and during the volatile periods of the 1970s. But significant equilibrating pressures are found even in the Δe_t and Δr_t equations and, by-and-large, the fitted values seem to perform well in terms of tracking the main movements of all the dependent variables, reflecting the fact that the remaining $\bar{R}^2$ are relatively high and

Table 9.4 Reduced form error correction specification for the core model.

Equation	$\Delta(p_t - p_t^*)$	Δe_t	Δr_t	Δr_t^*	Δy_t	Δy_t^*	$\Delta(h_t - y_t)$	$\Delta(\Delta\tilde{p}_t)$
$\hat{\xi}_{1,t}$	−0.015†	0.060†	0.002	0.002	0.017†	0.021†	−0.024*	−0.005
	(0.007)	(0.029)	(0.002)	(0.001)	(0.008)	(0.004)	(0.013)	(0.004)
$\hat{\xi}_{2,t}$	−0.840†	1.42	0.049	0.130*	1.34†	0.891†	−0.721	−0.811†
	(0.301)	(1.28)	(0.107)	(0.043)	(0.353)	(0.181)	(0.576)	(0.297)
$\hat{\xi}_{3,t}$	0.062†	−0.210*	−0.013	−0.006	−0.165†	−0.021	0.106*	0.034
	(0.029)	(0.121)	(0.010)	(0.004)	(0.034)	(0.017)	(0.055)	(0.028)
$\hat{\xi}_{4,t}$	0.018†	−0.029	−0.003*	−0.001*	−0.027†	−0.016†	−0.003	0.009*
	(0.005)	(0.020)	(0.002)	(0.001)	(0.005)	(0.003)	(0.009)	(0.005)
$\hat{\xi}_{5,t}$	−0.149*	−0.244	−0.054*	−0.024†	−0.099	−0.119†	0.408†	0.451†
	(0.083)	(0.353)	(0.028)	(0.012)	(0.098)	(0.050)	(0.159)	(0.082)
$\Delta(p_{t-1} - p_{t-1}^*)$	0.459†	0.150	−0.039	−0.028†	−0.136	−0.013	0.046	0.436†
	(0.095)	(0.404)	(0.032)	(0.014)	(0.111)	(0.057)	(0.182)	(0.094)
Δe_{t-1}	0.051†	0.216†	−0.005	−0.001	0.021	0.013	0.007	−0.022
	(0.022)	(0.092)	(0.007)	(0.003)	(0.025)	(0.013)	(0.042)	(0.021)
Δr_{t-1}	0.416†	−1.31	0.125	−0.067	0.467	0.204	−0.677	0.974†
	(0.294)	(1.25)	(0.098)	(0.042)	(0.345)	(0.177)	(0.562)	(0.290)
Δr_{t-1}^*	−0.810	2.75	−0.606†	0.430†	0.306	0.573	−0.267	0.166
	(0.617)	(2.62)	(0.205)	(0.088)	(0.723)	(0.371)	(1.18)	(0.606)
Δy_{t-1}	0.083	0.072	0.017	0.015	−0.044	0.031	−0.168	0.356†
	(0.089)	(0.381)	(0.030)	(0.013)	(0.105)	(0.053)	(0.172)	(0.089)
Δy_{t-1}^*	0.010	−0.630	−0.050	0.040*	−0.073	0.069	0.602*	−0.010
	(0.161)	(0.683)	(0.054)	(0.023)	(0.188)	(0.097)	(0.307)	(0.158)
$\Delta(h_{t-1} - y_{t-1})$	0.116	0.331	0.026	0.006	0.069	−0.014	−0.253†	0.140†
	(0.054)	(0.228)	(0.018)	(0.008)	(0.063)	(0.032)	(0.103)	(0.053)
$\Delta(\Delta\tilde{p}_{t-1})$	−0.151†	0.321	0.016	0.010	0.125	−0.082*	0.012	−0.244†
	(0.073)	(0.302)	(0.024)	(0.011)	(0.086)	(0.044)	(0.140)	(0.072)
Δp_t^o	−0.018†	−0.024	0.001	0.001†	−0.010†	0.0001	0.024†	0.003
	(0.004)	(0.018)	(0.001)	(0.0005)	(0.005)	(0.002)	(0.008)	(0.004)
Δp_{t-1}^o	0.010†	−0.013	−0.002	−0.0001	0.006	0.002	−0.011	0.016†
	(0.005)	(0.019)	(0.002)	(0.0001)	(0.005)	(0.003)	(0.009)	(0.004)
$\bar{R}^2$	0.484	0.070	0.115	0.345	0.260	0.367	0.257	0.445
Benchmark $\bar{R}^2$	0.316	0.026	0.007	0.213	0.022	0.194	0.00	0.191
$\hat{\sigma}$	0.007	0.032	0.002	0.001	0.009	0.004	0.014	0.007
$\chi^2_{SC}[4]$	2.79	0.96	2.43	17.13†	6.71	0.79	8.37†	5.63
$\chi^2_{FF}[1]$	8.57†	0.13	4.34†	6.70†	0.04	5.28†	0.033	0.01
$\chi^2_{N}[2]$	12.53†	13.98†	17.15†	19.9†	112.4†	10.84	31.45†	118.9†
$\chi^2_{H}[1]$	6.13†	1.97	4.53†	5.2†	0.88	0.93	0.19	4.55†

Note: The five error correction terms are given by

$$\hat{\xi}_{1,t+1} = p_t - p_t^* - e_t - 4.588,$$

$$\hat{\xi}_{2,t+1} = r_t - r_t^* - 0.0058,$$

$$\hat{\xi}_{3,t+1} = y_t - y_t^* + 0.0377,$$

$$\hat{\xi}_{4,t+1} = h_t - y_t + \underset{(22.2844)}{56.0975} r_t + \underset{(0.0012)}{0.0073} t + 0.05379,$$

$$\hat{\xi}_{5,t+1} = r_t - \Delta\tilde{p}_t - 0.0036.$$

Standard errors are given in parentheses. '*' indicates significance at the 10% level, and '†' indicates significance at the 5% level. The diagnostics are chi-squared statistics for serial correlation (SC), functional form (FF), normality (N) and heteroscedasticity (H). The benchmark $\bar{R}^2$ statistics are computed based on univariate ARMA(s, q), $s, q = 0, 1, \ldots, 4$ specifications with the s and q orders selected by AIC; see text for details.

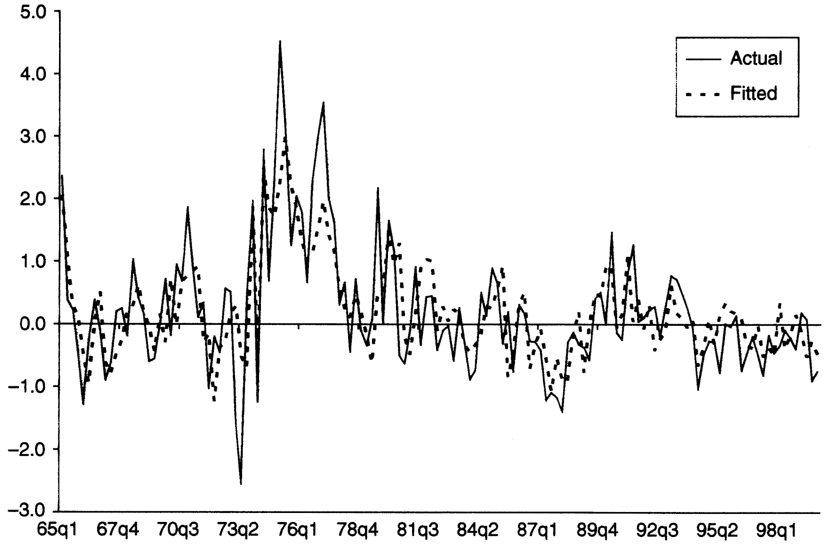

Figure 9.2a Actual and fitted values for the $\Delta(p_t - p_t^*)$ reduced form ECM equation.

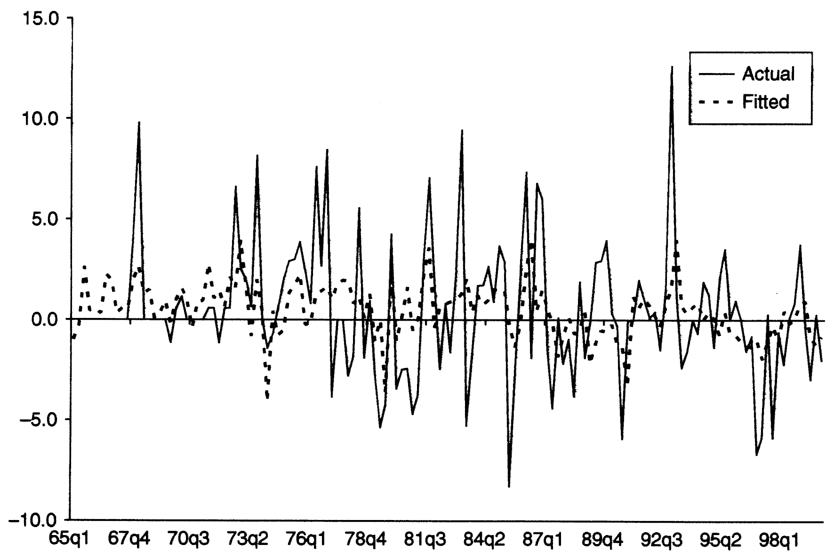

Figure 9.2b Actual and fitted values for the Δe_t reduced form ECM equation.

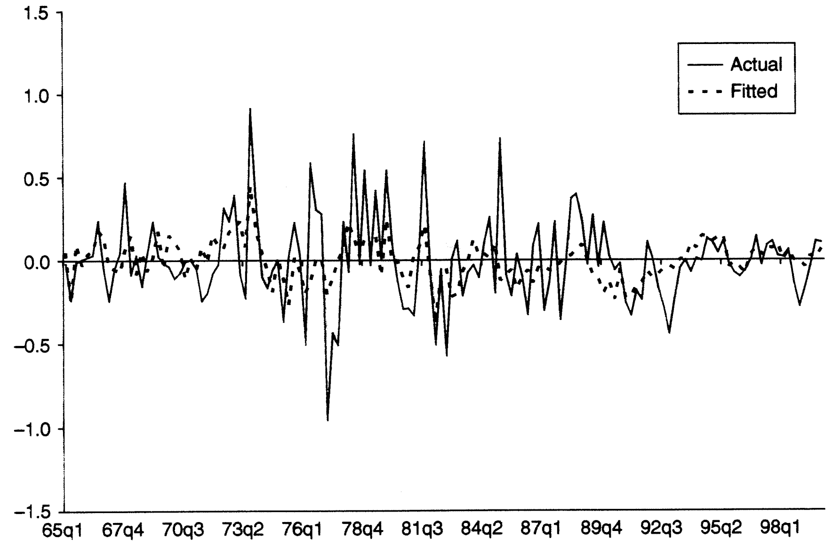

Figure 9.2c Actual and fitted values for the Δr_t reduced form ECM equation.

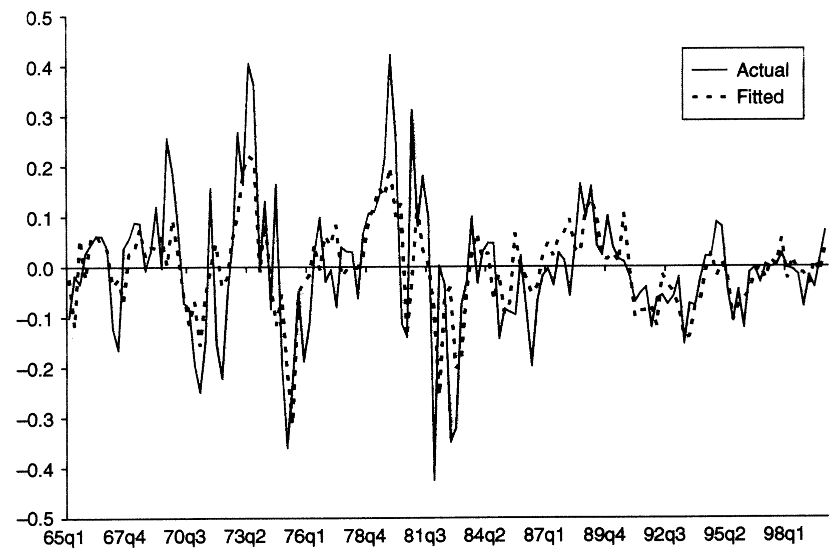

Figure 9.2d Actual and fitted values for the Δr_t^* reduced form ECM equation.

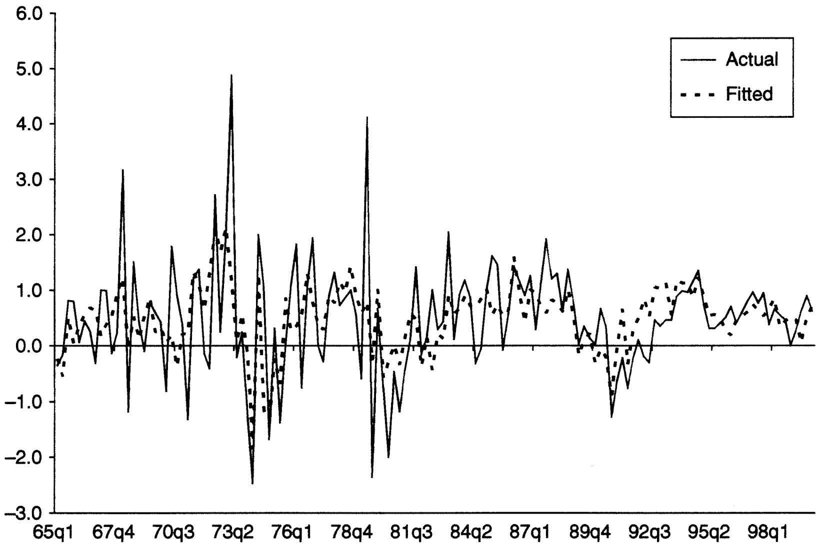

Figure 9.2e Actual and fitted values for the Δy_t reduced form ECM equation.

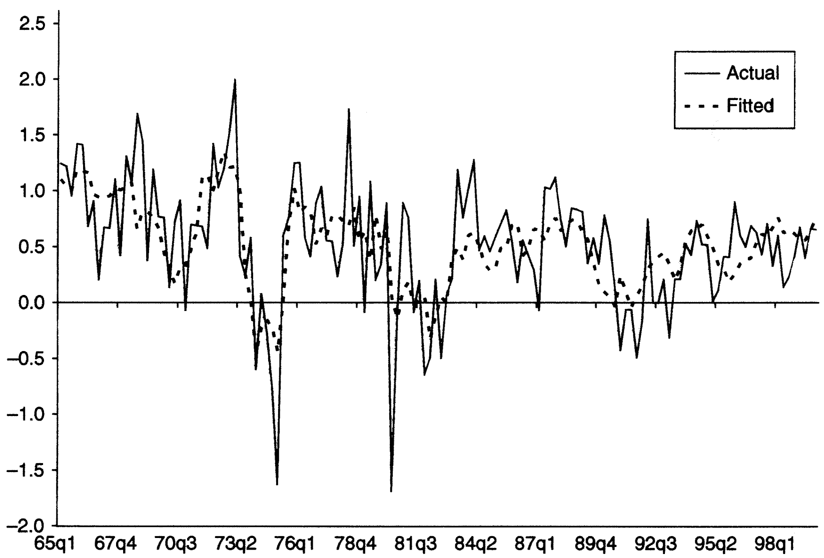

Figure 9.2f Actual and fitted values for the Δy_t^* reduced form ECM equation.

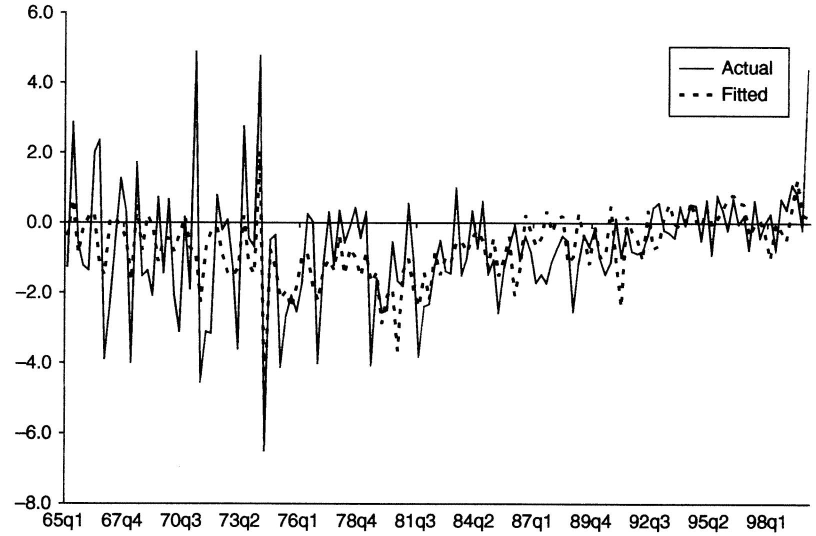

Figure 9.2g Actual and fitted values for the $\Delta(h_t - y_t)$ reduced form ECM equation.

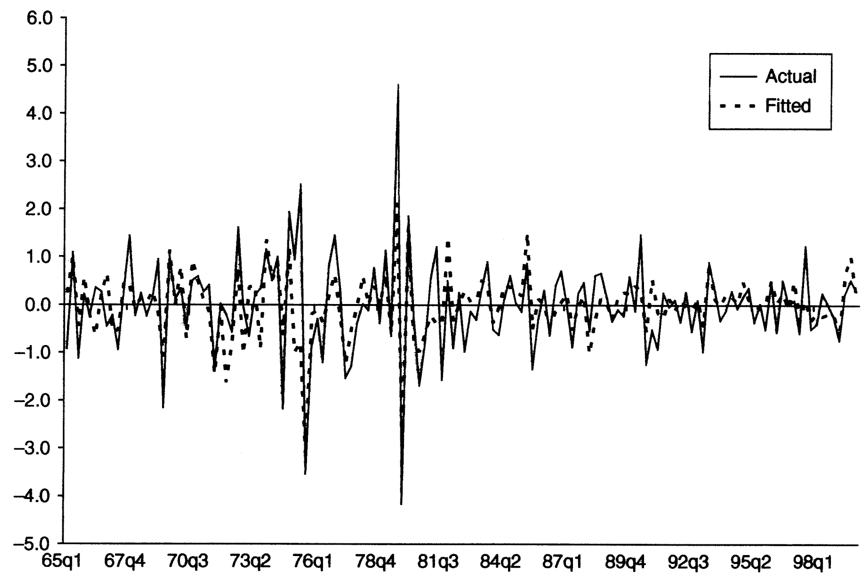

Figure 9.2h Actual and fitted values for the $\Delta(\Delta\tilde{p}_t)$ reduced form ECM equation.

lie in the range [0.25, 0.49]. Generally speaking, then, the equations of Table 9.4 appear to capture well the time series properties of the main macroeconomic aggregates in the UK over the period since the mid-1960s.

9.4.3 *Comparing the core model with benchmark univariate models*

In order to evaluate the in-sample fit of the individual equations in the core long-run structural model a little more rigorously, we can compare the ECM specifications in Table 9.4 with a set of 'benchmark' univariate time series representations. To this end, and in view of the unit root properties of the variables, we estimate ARMA(p, q) specifications applied to the first differences of each of the eight core endogenous variables in turn. These benchmark models are selected following the Box–Jenkins methodology and allow us to address the question of how much, if at all, the explanatory power and potential forecasting ability of the model has improved by the adoption of the long-run structural modelling approach.[10]

We examine a range of ARMA models for each core endogenous variable. For example, in the case of the real output variable, y_t, the ARMA(p, q) specification can be written as:

$$\Delta y_t = \alpha + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + \sum_{i=1}^{q} \gamma_i \epsilon_{t-i} + \epsilon_t, \quad t = 1, \ldots, T. \tag{9.15}$$

The first requirement in the construction of the benchmark model is the selection of an *a priori* maximum lag order for the autoregressive and moving average processes, p and q, respectively. Here we choose 4, in light of the quarterly nature of the data, the number of available observations (140 observations for the sample period 1965q1–1999q4) and considering that the degree of serial correlation in the first difference of the macrovariables is not very high. We then examine the full set of model combinations that are spanned by all $p = 0, 1, \ldots, 4$ and $q = 0, 1, \ldots, 4$, providing 25 different combinations. Our preferred benchmark model is then selected on the basis of the Akaike Information Criterion (AIC).

The choice of AIC for model selection, compared with the Schwarz Bayesian Criterion (SBC) for example, relates to various practical and theoretical issues involved in the use of AIC and SBC. For example, choosing the SBC over AIC as a tool of model selection may be reasonable if we are confident that the true model lies in the set of models under consideration. Only in these circumstances (and assuming certain other regularity conditions are met) is SBC a consistent model selection criterion. In contrast, AIC is a more appropriate selection criterion if the aim is to select the best approximating model (in the information-theoretic sense), as we believe to be the case in our particular application. We certainly do not claim that the 'true model' lies in the set of models that we are considering (univariate or vector error correcting), so this suggests the use of AIC in model selection. Moreover, the theoretical grounds for the use of SBC in the case of models involving unit roots and cointegration has not been fully developed; there remains no clear practical guidance on how one would allocate degrees of freedom across the equations in a cointegrating system in calculating SBC; and there is evidence that SBC can seriously underestimate the lag order in these circumstances. Moreover, the AIC is designed for minimising the forecast error variance (see Lütkepohl (1991), Chapter 4). This is a feature that might be thought to be important since one of the key uses of our model will be in probability event forecasting (see Chapter 11).[11]

[10] Of course, these comparisons do not measure the usefulness of the more structural interpretation and understanding which the use of a long-run structural model, based on economic theory, can entail.

The results of the estimation and selection of the univariate ARMA models are summarised in Table 9.5, providing details of the AIC, SBC and $\bar{R}^2$ statistics calculated for different models estimated for each of the eight endogenous variables.

The first two columns of Table 9.5 relate to the unrestricted 'ARMA(4, 4)' specifications for each variable and to the error correction specification of the core model discussed above and reported in Table 9.4 (described as 'unrestricted' in the sense that the short-run dynamics are unconstrained). The third column relates to our preferred benchmark ARMA model chosen by AIC, and imposing restrictions on the short-run dynamics as discussed above. Comparison across these three columns show that the error correction specifications of our core model outperform the preferred ARMA(p, q) model for 7/8 of the variables, Δe_t being the exception, in terms of the AIC (and in all eight in terms of the estimated $\bar{R}^2$'s). For example, the preferred benchmark ARMA model selected for the relative price variable, $\Delta(p_t - p_t^*)$, in the third column is the ARMA(4, 3) process. This model explains as much as 31.6% of the total variation in $\Delta(p_t - p_t^*)$ but this compares unfavourably with the error correction specification for this variable in the core model, which explains 48.4% of the variation. The preferred benchmark ARMA model for the change in domestic inflation

[11] Note that the use of the same criterion for model selection and model evaluation can lead to misleading results. For model evaluation, we prefer to use out-of-sample forecast evaluation procedures as illustrated in Chapter 11.

Table 9.5 Model selection criteria for the core model and alternative time series specifications.

Variable		Unrestricted		Restricted		
		ARMA(4,4)	ECM	ARMA(p,q) order selected by AIC	ARMA(p,q) order selected by SBC	ECM with short-run restrictions
$\Delta(p-p^*)$						
	AIC	416.62	479.27	463.43	462.37	480.67
	SBC	448.38	455.74	453.43	457.96	460.08
	$\bar{R}^2$	0.308	0.484	0.316	0.277	0.487
	$(\hat{p},\hat{q})$	–	–	(4,3)	(1,1)	–
	$\chi^2(m)$	–	–	–	–	1.18 (2)
Δe						
	AIC	276.37	276.45	280.03	280.03	282.54
	SBC	263.13	253.21	277.09	277.09	270.00
	$\bar{R}^2$	0.028	0.070	0.026	0.026	0.098
	$(\hat{p},\hat{q})$	–	–	(0,1)	(0,1)	–
	$\chi^2(m)$	–	–	–	–	4.35 (8)
Δr^*						
	AIC	741.36	750.80	744.56	744.46	754.29
	SBC	728.12	727.27	741.52	741.52	738.11
	$\bar{R}^2$	0.218	0.345	0.213	0.213	0.356
	$(\hat{p},\hat{q})$	–	–	(1,0)	(1,0)	–
	$\chi^2(m)$	–	–	–	–	3.00 (5)
Δr						
	AIC	632.37	633.17	631.67	631.66	638.55
	SBC	619.14	609.63	628.73	630.19	625.31
	$\bar{R}^2$	0.090	0.115	0.007	0.000	0.142
	$(\hat{p},\hat{q})$	–	–	(1,0)	(0,0)	–
	$\chi^2(m)$	–	–	–	–	3.20 (7)
Δy						
	AIC	442.80	456.97	442.95	442.85	460.72
	SBC	429.54	433.44	437.07	441.38	444.54
	$\bar{R}^2$	−0.130	0.260	0.022	0.000	0.276
	$(\hat{p},\hat{q})$	–	–	(3,0)	(0,0)	–
	$\chi^2(m)$	–	–	–	–	2.47 (5)
Δy^*						
	AIC	540.46	550.45	539.70	538.71	555.17
	SBC	527.22	526.92	535.29	535.77	540.46
	$\bar{R}^2$	0.102	0.367	0.194	0.178	0.385
	$(\hat{p},\hat{q})$	–	–	(1,1)	(1,0)	–
	$\chi^2(m)$	–	–	–	–	0.38 (6)
$\Delta(h-y)$						
	AIC	379.53	388.49	374.68	374.68	393.73
	SBC	366.29	364.96	373.22	373.22	379.02
	$\bar{R}^2$	0.186	0.257	0.000	0.000	0.284
	$(\hat{p},\hat{q})$	–	–	(0,0)	(0,0)	–
	$\chi^2(m)$	–	–	–	–	2.51 (8)
$\Delta^2 p$						
	AIC	456.05	481.16	458.88	457.94	484.53
	SBC	422.81	457.63	448.59	454.99	468.35
	$\bar{R}^2$	0.199	0.445	0.191	0.152	0.454
	$(\hat{p},\hat{q})$	–	–	(3,3)	(0,1)	–
	$\chi^2(m)$	–	–	–	–	3.23 (5)

Note: The unrestricted ECM equations are those reported in Table 9.4. The restricted ARMA(p,q) models are obtained using AIC and SBC searching over all possible orders $p,q = 0,1,2,3,4$. The restricted ECM equations are obtained using a general-to-specific search procedure that begins with the unrestricted single equation ECM and takes the form of dropping, one at a time, the lagged change and cointegrating terms, starting with the variable with the largest p-value (assuming it is greater than 0.25). The search process ends when all the terms that remain in the equation have p-values of 0.25 or less. The $\chi^2(m)$ statistic is the Lagrange multiplier joint test of m zero restrictions on coefficients of the deleted variables.

is an ARMA(3, 3) process which explains 19.1% of the total variation in $\Delta^2\tilde{p}_t$. But this compares with 44.5% for the long-run structural error correction specification. The preferred ARMA benchmark model for domestic output growth is an ARMA(3, 0) process, whose explanatory power is low and accounts for 2.2% of the movement in Δy_t. This compares with 26% for the long-run structural error correction specification. This pattern is repeated for all variables as far as the $\bar{R}^2$ is concerned.

The conclusion to be drawn from these results is that the error correction model of the core model does indeed perform well in comparison to univariate time series models chosen according to our preferred AIC. For completion, though, the table also provides details of the SBC statistics, including in the fourth column of Table 9.5 details of the ARMA specification that would be chosen according to this criterion. The SBC statistic places greater weight on parsimony in model selection, and this is reflected by the fact that relatively simple models are chosen in the fourth column. Moreover, comparison of the SBC of the error correction specifications of the core model and that of the ARMA models of the fourth column suggests that the ARMA models outperform the core model (since the SBC shows the ARMA(p,q) model to be preferred for 7/8 variables by this criterion). However, this is not an even-handed comparison. The error correction specifications of the core model were obtained without imposing restrictions on the short-run dynamics and are bound to be disadvantaged relative to the ARMA models when parsimony is given more weight. For a more balanced comparison, therefore, we calculated the SBC statistics associated with a 'restricted ECM' in which a specification search was conducted, starting from the 'unrestricted' ECMs of the core model but dropping terms when the p-values of the estimates coefficient were greater than 0.25. As the results in the final column of Table 9.5 shows, the resultant 'restricted' ECM model outperforms the restricted univariate model in 8/8 cases according to AIC and in 5/8 of the cases according to SBC. So any criticism of our model for not adopting SBC as a selection mechanism for a benchmark comparison effectively disappears when the criteria are employed in a comparable manner.

9.5 An alternative model specification

The core model presented in the sections above fits the short-run dynamics well and embodies the economic theory's long-run relations in a transparent manner and in a way that is consistent with the data. Before moving on to discuss the use of the model, however, it is worth checking its robustness to alternative specifications. This also allows us to illustrate the types of choices typically encountered when performing empirical work of this sort. In what follows, we comment on one possible alternative model, which is similar in many respects to our preferred core model, but which is based on a different interpretation of the preliminary statistical analysis and one that places emphasis on the different aspects of the theoretical arguments.

Specifically, recall from the earlier discussion on the tests of unit roots in the variables that there is some ambiguity in the data regarding the order of integration of the price variables. The application of the ADF(s) tests to Δp_t and Δp_t^* yields mixed results. The hypothesis that there is a unit root in the domestic and foreign inflation rates is rejected for low orders of augmentation (namely, for $p = 0$ and 1), but not for higher orders. Overall the available data is not informative as to whether domestic and foreign prices are $I(1)$ or $I(2)$.

In our preferred model described in the previous section, we chose to follow Haldrup's (1998) advice on the analysis of $I(2)$ variables by working with the inflation series Δp_t and the relative price variable $p_t - p_t^*$ rather than the price levels p_t and p_t^* separately. The statistical evidence supports the view that $p_t - p_t^*$ is $I(1)$ and, on balance, the same is true for Δp_t. So we have some reassurance that our empirical work is statistically sound. However, this is not the only choice available. Investigation shows that the transformed series $p_t - p_t^o$ and $p_t^* - p_t^o$ are also unambiguously $I(1)$ according to the tests available. An alternative model might therefore be obtained employing exactly our modelling procedure, working instead with the vector of variables $\mathbf{z}_t^{ALT} = (p_t - p_t^o, e_t, r_t^*, r_t, y_t, p_t^* - p_t^o, h_t - y_t, y_t^*)'$ and in which the long-run relationships suggested by economic theory are captured by the vector

$$\boldsymbol{\beta}'_{ALT} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & -\beta_{44}^a & -\beta_{46}^a & 0 & 1 & 0 \end{pmatrix}.$$

The vector $\boldsymbol{\beta}_{ALT}$ incorporates the *PPP*, *IRP*, *OG*, and *MME* relationships of our preferred model (but not *FIP*) and, in terms of its treatment of the ambiguity on the order of integration of the price variables, the model is as justifiable as our preferred model. As shown in the empirical exercise of Garratt *et al.* (2000), this alternative model also performs well in terms of the fit of the data, with the associated test of the long-run relations accepted and with satisfactory diagnostics for the associated error correction equations.

In these circumstances, the judgement on which of the two models is preferred has to be based on economic as well as statistical analysis. The use of the variables $p_t - p_t^o$ and $p_t^* - p_t^o$ in the alternative model has at its base the view that, once the effects of oil price movements are taken into account, the price series are $I(1)$. This is appealing to those who point out that inflation rates are unlikely to grow without bounds and are therefore best modelled as being stationary. However, the statistical evidence indicates unambiguously that nominal interest rates are $I(1)$. If prices are treated as $I(1)$, then the modeller can only maintain the long-run *FIP* relationship if the interest rate is excluded from the cointegrating analysis (assuming nominal rates are $I(0)$ despite the statistical evidence). Or interest rates can be retained in the analysis as $I(1)$ variables, but then the long-run *FIP* relationship cannot be accommodated within the model (as is the case with $\boldsymbol{\beta}_{ALT}$ above). We preferred to work with the relative price variable, $p_t - p_t^*$ and the inflation rate Δp_t, since this allows us to accommodate the *FIP* relationship in the model in a straightforward way. While we recognise the difficulties in the view that price inflation and nominal interest rate series are $I(1)$, we also note the importance of capturing the statistical properties of the sample of data available and of accommodating the *FIP* relationship in the long-run model. We have, therefore, decided to continue with the model described in the earlier section, but we understand that others may make a different judgement. This highlights the importance of taking into account model uncertainty when making decisions on the basis of models. This is an issue that we explore in the work of Chapter 11 below.