ISSN 1032-3813
ISBN 0 7326 1027 3

MONASH UNIVERSITY

AUSTRALIA

PRIVATE AND PUBLIC CONSUMPTION EXPENDITURE SUBSTITUTABILITY: BAYESIAN ESTIMATES FOR THE G7 COUNTRIES

Gael M. Martin
Vance L. Martin

Working Paper 4/97
April 1997

DEPARTMENT OF ECONOMETRICS
AND BUSINESS STATISTICS

Private and Public Consumption Expenditure Substitutability: Bayesian Estimates for the G7 Countries*

Gael M. Martin[1] and Vance L. Martin[2]

[1]Corresponding Author. Department of Econometrics and Business Statistics
Monash University, Clayton 3168, Australia
Email: Gael.Martin@BusEco.monash.edu.au

[2]Department of Economics, University of Melbourne, Parkville 3052, Australia

April 1997

Abstract

The degree of substitution between private and public per capita consumption for the G7 countries is estimated over the period 1960 to 1996. Special attention is given to isolating both long-run and short-run substitution effects. Inferences are produced using a Bayesian methodology, with a Jeffreys prior being used to offset an identification problem in the likelihood function. The marginal posterior densities of interest are estimated via a hybrid of Markov chain Monte Carlo algorithms. The empirical results indicate that for the US, Germany, France and Italy, private and public consumption expenditure are substitutes in the short-run, but complements in the long-run. The opposite result occurs for the UK, whilst Japan and Canada exhibit no significant short-run or long-run relationships.

JEL classification: C11, C32, H31

*This paper is a revised version of Chapter 3 of the first author's Ph.D. thesis. The research has been supported by both an ARC Postgraduate Award and an ARC Research Grant. Both authors would like to thank Brett Inder and Keith McLaren for helpful comments on an earlier draft of the paper.

1 Introduction

This paper is concerned with estimating the degree of substitution between private and public consumption expenditure. In contrast to existing empirical work conducted in the area, special attention is given to estimating both long-run and short-run relationships. The long-run model is derived from the theoretical framework of Aschauer (1985). Combined with the assumption that private and public consumption expenditures are integrated processes of order one, it is shown that this framework leads to a cointegrating relationship between the two series. The short-run properties of the model are captured by the dynamic representations specified for the generating processes.

An important feature of this framework is that it allows for the possibility that the short-run impact of changes in government consumption can differ from the long-run impact in both magnitude and sign. In fact, this result occurs empirically for five of the G7 countries studied in the paper. Specifically, the empirical results suggest that in the case of the US, Germany, France and Italy, private and public consumption are substitutes in the short-run, but are negative substitutes, that is complements, in the long-run. The reverse result occurs for the UK. In the case of Japan and Canada, no significant short-run or long-run relationship between private and public consumption expenditure is found.

In the derivation of the model, the focus of attention is on the extent to which government consumption expenditure substitutes for, and hence leads to an *ex ante* crowding out of, private consumption expenditure. Issues of Ricardian equivalence are not dealt with. For examples of studies which consider both the degree of substitutability between private and public consumption expenditure and Ricardian equivalence, see Feldstein (1982), Kormendi (1983), Aschauer (1985), Leiderman and Razin (1988) and Graham and Himarios (1991).[1]

An outcome of the approach adopted here is that the long-run substitution properties of the model can be tested solely in terms of the variables, private and public consumption expenditure. Provided that cointegration is established, this rules out the need for the inclusion of other variables as suggested by Graham (1993), in testing for crowding out, at least in the long-run. The establishment of cointegration also rules out the need for decomposing government expenditure into its sub-components as suggested by Graham and Darby

[1]Literature which focuses on crowding out with respect to investment rather than consumption, includes Aschauer (1989a and b) and Cebula, Killingsworth and Belton (1994), amongst others.

and Malley (1996). Moreover, the cointegration approach highlights the fact that the Campbell and Mankiw (1990) modelling framework, as adopted by Graham, focuses only on the short-run relationship and, in failing to take into account the error correction term, is misspecified when a cointegrating relationship between private and public consumption expenditure is established.

In producing inferences about substitutability, or crowding out, a Bayesian methodology is used, with the marginal posterior densities of interest being estimated via a hybrid of Markov chain Monte Carlo (MCMC) sampling algorithms; see Geweke (1994) and Chib and Greenburg (1996). A Jeffreys prior is employed to circumvent an identification problem in the likelihood function whilst, at the same time, adhering to the principles of noninformative, or objective, prior specification. This line of approach follows Kleibergen and van Dijk (1994) and Martin (1997), and contrasts with the work of Bauwens and Lubrano (1993) and Geweke (1996), in which identification problems in models of cointegration are solved, in part, via subjective prior information.[2] The paper also draws on the work of Phillips and Ploberger (1994 and 1995) and Phillips (1996), in which objective Bayesian criteria for both model selection and hypothesis testing are developed. For comparison with the Bayesian results, a range of classical inferences are also reported.

The advantages of the Bayesian method are as follows. First, it is a coherent way of producing simultaneous estimates of the long-memory component in the cointegrating error and the cointegrating parameters. The simultaneous aspect of the procedure means that the accuracy with which the cointegrating parameters are estimated is appropriately affected by the probability that the relationship is indeed a cointegrating one. This contrasts with the classical approach, whereby all inference regarding the cointegrating parameters proceeds conditionally on the assumption that cointegration exists. Second, given that the joint posterior distribution for the assumed cointegration model induces manageable conditional posteriors, a fully parametric Bayesian approach can proceed via a Gibbs-based MCMC sampling scheme. This enables estimates of all long-run and short-run parameters to be produced, which contrasts with a nonparametric classical procedure such as the Phillips and Hansen (1990) fully modified ordinary least squares (FMOLS) method, which yields estimates of the long-run cointegrating vector and standard errors only. Alternatively, the parametric classical procedure of Phillips and Loretan (1991), whilst yielding both long-run and short-run parameter estimates, is adversely affected by the above-mentioned identification problem, which the Bayesian method counteracts via the use of the Jeffreys prior; see Martin (1996). Finally, conditional on the data, the results are exact with no recourse required to asymptotic

[2] The use of Jeffreys priors in overcoming identification problems in models other than those for cointegration occurs in Schotman and van Dijk (1991a), Chao and Phillips (1996) and Kleibergen and van Dijk (1996).

approximations.[3]

The rest of the paper proceeds as follows. A model of crowding out is derived in Section 2. The details of the Bayesian inferential method are given in Section 3. This section includes the specification of the requisite Jeffreys prior, plus an outline of the Gibbs and Metropolis-Hastings MCMC algorithms which are applied in the estimation of the relevant marginal posteriors. A more detailed derivation of the Jeffreys prior is given in an appendix to the paper. Section 4 provides the empirical results of testing for and estimating the nature of crowding out in both the long-run and the short-run, over the period 1960 to 1996 for the G7 countries. Section 5 contains some concluding remarks.

2 A Model of Crowding Out

2.1 Derivation of the Model

Following Aschauer (1985), consider a representative individual who chooses to maximize the following intertemporal utility function with respect to current and future "effective consumption" C^*_{t+j}, $j \geq 0$,

$$E_t(U) = \sum_{j=0}^{\infty} \left(\frac{1}{1+\delta}\right)^j E_t u(C^*_{t+j}), \tag{1}$$

where δ is the subjective rate of time preference, $u(.)$ is a time-invariant, concave, instantaneous utility function and E_t denotes conditional expectations formed at time t. To keep the analysis general, it is assumed that C^*_t is some unspecified increasing function of real private consumption (C_t) and real government consumption (G_t)

$$C^*_t = f(C_t, G_t), \tag{2}$$

in which case, (1) is re-expressed as

$$E_t(U) = \sum_{j=0}^{\infty} \left(\frac{1}{1+\delta}\right)^j E_t v(C_{t+j}, G_{t+j}), \tag{3}$$

with $v(.)$ now specified as a time-invariant, concave, instantaneous utility function. On the assumption that the individual takes the flow of government expenditure as given, optimizing (1) with respect to C^*_t amounts to optimizing (3) with respect to C_t.

[3] As a qualification to this last point Phillips and Ploberger (1994 and 1995) and Phillips (1996) argue that the data conditioning involved in the Bayesian paradigm and the simplifications which it produces over the classical approach, is not innocuous in a time series context.

The budget constraint for the individual is given by

$$\sum_{j=0}^{\infty}\left(\frac{1}{1+r}\right)^{j} E_tC_{t+j} = \sum_{j=0}^{\infty}\left(\frac{1}{1+r}\right)^{j} E_t(L_{t+j} - T_{t+j}) + W_t, \qquad (4)$$

where

$$\begin{aligned} L_t &= \text{period } t \text{ labour income,} \\ T_t &= \text{period } t \text{ tax payments,} \\ W_t &= \text{beginning of period holding of one-period bonds and} \\ r &= \text{the real interest rate.} \end{aligned}$$

Defining B_t as government debt of one-period maturity, the government's budget constraint is

$$\sum_{j=0}^{\infty}(\frac{1}{1+r})^j E_tT_{t+j} = B_t + \sum_{j=0}^{\infty}(\frac{1}{1+r})^j E_tG_{t+j},$$

which can, in turn, be substituted into (4) to obtain

$$\sum_{j=0}^{\infty}\left(\frac{1}{1+r}\right)^{j} E_tC_{t+j} = \sum_{j=0}^{\infty}\left(\frac{1}{1+r}\right)^{j} E_t(L_{t+j} - G_{t+j}) + (W_t - B_t). \qquad (5)$$

Maximizing (3) with respect to C_{t+j}, $j \geq 0$, subject to (5), yields the following first-order necessary conditions

$$E_tv'\left(C_{t+j}(G_{t+j})\right) = \lambda\left(\frac{1+\delta}{1+r}\right)^{j}, \quad j \geq 0, \qquad (6)$$

in addition to the intertemporal budget constraint, with λ denoting the Lagrangian multiplier and $v'(.)$ denoting the partial derivative of v with respect to C_{t+j}, $j \geq 0$. The marginal utility of private consumption is explicitly expressed as a function of government expenditure in order to highlight the fact that the latter is treated as given in the consumer's decision making process.

Consideration of the choice of consumption in the two adjacent periods, t and $t+1$, leads to the Euler equation

$$E_tv'\left(C_{t+1}(G_{t+1})\right) = bv'\left(C_t(G_t)\right), \qquad (7)$$

where $b = (1+\delta)/(1+r)$. Since the functional form of neither the initial utility function $u(.)$ nor the effective consumption function $f(.)$ is specified, some approximation for the marginal utility function $v'(.)$ is required. Motivated partly by the empirical considerations discussed by Campbell and Mankiw

(1990) and partly by the theoretical class of models used by Lucas and Rapping (1969), amongst others, $v'(.)$ is approximated by the following log-linear function

$$v'(C_t(G_t)) = \omega_0 + \omega_1 \ln(C_t) + \omega_2 \ln(G_t), \tag{8}$$

where the concavity of $v(.)$ requires $\omega_1 < 0$.[4] Substituting (8) into (7) and rearranging yields

$$E_t c_{t+1} = \frac{\omega_0}{\omega_1}(b-1) + bc_t + \frac{\omega_2}{\omega_1}(bg_t - E_t g_{t+1}), \tag{9}$$

where

$$\begin{aligned} c_\tau &= \ln(C_\tau) \text{ and} \\ g_\tau &= \ln(G_\tau) \text{ for } \tau = t, t+1. \end{aligned} \tag{10}$$

In evaluating the expectations in (9), it is necessary to make an assumption about both the nature of the generating processes for c_t and g_t and the nature of expectation formation. It is assumed that it is appropriate to represent both c_t and g_t as $I(1)$ processes,

$$c_t = c_{t-1} + w_t \text{ and} \tag{11}$$

$$g_t = g_{t-1} + v_t, \tag{12}$$

with w_t and v_t being stationary error processes. This assumption is tested empirically in Section 4. In order to simplify the proposed numerical procedure, v_t is specified as an autoregressive (AR) process of finite order q,

$$\Psi(L)v_t = \eta_t, \tag{13}$$

with $\Psi(L) = 1 - \psi_1 L - \psi_2 L^2 - \cdots - \psi_q L^q$ and L denoting the lag operator. The error process w_t is specified as a finite order AR process of order k,

$$\Lambda(L)w_t = \xi_t, \tag{14}$$

with $\Lambda(L) = 1 - \gamma_1 L - \gamma_2 L^2 - \cdots - \gamma_k L^k$. In keeping with the assumption of stationarity for v_t and w_t, the roots of both $\Psi(L)$ and $\Lambda(L)$ are specified as lying outside the unit circle. Both η_t and ξ_t are assumed to be white noise processes, with respective error variances σ_η^2 and σ_ξ^2 and covariance $\sigma_{\eta\xi}$.

Expectation formation is taken to be rational and, thus, consistent with the specifications in (11) to (14). Moreover, given the assumption invoked earlier regarding the status of g_t in the individual's decision making process, it is appropriate to construct expectations of c_{t+1} and g_{t+1} respectively as

$$\begin{aligned} E_t c_{t+1} = {} & c_t + \gamma_1 \Delta c_t + \gamma_2 \Delta c_{t-1} + \cdots + \gamma_k \Delta c_{t-[k-1]} \\ & + \frac{\sigma_{\xi\eta}}{\sigma_\eta^2}(\psi_1 \Delta g_t + \psi_2 \Delta g_{t-1} + \cdots + \psi_k \Delta g_{t-[q-1]}) \end{aligned} \tag{15}$$

[4]In contrast, Aschauer (1985) adopts a linear specification for marginal utility; see Section 2.3 below.

and

$$E_t g_{t+1} = g_t + \psi_1 \Delta g_t + \psi_2 \Delta g_{t-1} + \cdots + \psi_k \Delta g_{t-[q-1]}, \tag{16}$$

where $\Delta c_{t-j} = c_{t-j} - c_{t-(j+1)}$, $j = 0, 1, \ldots, k-1$ and $\Delta g_{t-j} = g_{t-j} - g_{t-(j+1)}$, $j = 0, 1, \ldots, q-1$. The representation of $E_t g_{t+1}$ in (16) implies that g_t is *strongly* exogenous, as it is both *weakly* exogenous for β and is not Granger caused by c_t. The allowance for $\sigma_{\eta\xi} \neq 0$ means that g_t is not *strictly* exogenous.

Substitution of (15) and (16) into (9) produces

$$\begin{aligned} c_t &= \frac{\omega_0}{\omega_1}(b-1) + bc_t + \frac{\omega_2}{\omega_1}(b-1)g_t \\ &\quad -(\frac{\omega_2}{\omega_1} + \frac{\sigma_{\xi\eta}}{\sigma_\eta^2}) \sum_{j=1}^{q} \psi_j \Delta g_{t-[j-1]} - \sum_{j=1}^{k} \gamma_j \Delta c_{t-[j-1]} \\ &= \alpha + \beta g_t + h(\Delta g, \Delta c), \end{aligned} \tag{17}$$

where

$$\begin{aligned} \alpha &= -\frac{\omega_0}{\omega_1}, \\ \beta &= -\frac{\omega_2}{\omega_1} \end{aligned}$$

and $h(\Delta g, \Delta c)$ denotes a function of present and past changes in c_t and g_t. Given the assumption that c_t and g_t are $I(1)$ processes, (17) constitutes a cointegrating relationship with cointegrating vector $(1, -\beta)$. This suggests the following stochastic representation,

$$c_t = \alpha + \beta g_t + h(\Delta g, \Delta c) + u_t^{\#},$$

where $u_t^{\#}$ is a general stationary error term. Alternatively, by grouping all stationary components together, the long-run relationship between c_t and g_t is expressed as

$$c_t = \alpha + \beta g_t + u_t,$$

where

$$u_t = u_t^{\#} + h(\Delta g, \Delta c)$$

is a stationary error term. Once again with a view to simplifying the numerical analysis, u_t is parameterized as the following finite AR process of order p

$$\Phi(L) u_t = \varepsilon_t. \tag{18}$$

The underlying error terms in (13) and (18) are assumed to have a bivariate Normal distribution of the form

$$e_t = (\varepsilon_t, \eta_t)' \sim i.i.d.\, N\left(0, \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}\right) \quad \forall t, \tag{19}$$

with Σ a (2×2) symmetric positive definite matrix. The allowance for $\sigma_{12} \neq 0$ accommodates the fact that g_t is not assumed to be strictly exogenous. Furthermore, it is sufficient, in combination with the specification of serial correlation in u_t, to accommodate the fact that u_t is a function of both the present and past values of Δg_t through $h(,.,)$.

2.2 Discussion of the Empirical Model

For convenience, all components of the derived empirical model are gathered together as follows:

$$\begin{aligned} c_t &= \alpha + \beta g_t + u_t, & (20) \\ g_t &= g_{t-1} + v_t, & (21) \\ \Phi(L)u_t &= \varepsilon_t, & (22) \\ \Psi(L)v_t &= \eta_t \text{ and} & (23) \\ e_t &= (\varepsilon_t, \eta_t)' \sim N(0, \Sigma). & (24) \end{aligned}$$

The parameter β is the long-run elasticity linking private and public expenditure and, therefore, represents a measure of substitutability between the two. If $\beta < 0$ government expenditure can be viewed as a substitute for and thereby as crowding out private expenditure. If $\beta > 0$ government expenditure is a complement for and thereby crowds in private expenditure. If $\beta = -1$ $(+1)$, private expenditure is crowded out (in) by the same proportionate amount as the increase in public expenditure. If $\beta = 0$, an increase in government expenditure has no impact on private expenditure.

The short-run dynamics of the model are captured by the coefficients associated with the lag polynomials $\Phi(L)$ and $\Psi(L)$ in (22) and (23) respectively. An important feature of this model is its flexibility, as it is possible for one type of crowding out to occur in the short-run but be reversed in the long-run. As noted previously, this is what is found to occur empirically for certain of the G7 countries considered in Section 4.

The cointegration properties of the model are governed by the polynomial $\Phi(L)$ in (22). In particular, applying the arguments outlined in Phillips (1991a and b), the long-term memory of the model is captured by the sum of the coefficients of this polynomial. If this sum is less than one, c_t and g_t are cointegrated. In order to produce, as the basis for inference, the marginal density of this parameter, it is convenient to reparameterize $\Phi(L)u_t$ as

$$\Phi(L)u_t = u_t - \rho' u^{\tau}_{t-1}, \tag{25}$$

where $u_{t-1}^{r} = (u_{t-1}, u_{t-1} - u_{t-2}, \ldots, u_{t-[p-1]} - u_{t-p})'$ and $\rho = (\rho_1, \rho_2, \ldots \rho_p)'$. As highlighted by Phillips, the parameter ρ_1, defined by

$$\rho_1 = \phi_1 + \phi_2 + \cdots \phi_p, \tag{26}$$

characterizes the behavior of u_t at the zero frequency. As such, it is the appropriate parameter for capturing the nature of the long-run relationship between c_t and g_t.[5] The remaining elements in ρ are defined by $\rho_k = -\sum_{i=k}^{p} \phi_i$.

2.3 Relationship with Aschauer's Model

The crowding out parameter β in (20), is directly related to the parameter representing the marginal rate of substitution in utility between private and public consumption in Aschauer's (1985) specification. Aschauer assumes a quadratic utility function,

$$u(C_t^*) = -0.5(\bar{C}^* - C_t^*)^2 \tag{27}$$

and a linear effective consumption function,

$$C_t^* = f(C_t, G_t) = C_t + \theta G_t, \tag{28}$$

where $\bar{C}^*$ is the bliss level of effective consumption and θ is the consumer's marginal rate of substitution between private and public consumption. The parameter θ is the focus of Aschauer's work, with estimates of it forming the basis of his conclusions regarding crowding out. For this model, the marginal utility function in (8) becomes

$$v'(C_t(G_t)) = \bar{C}^* - C_t - \theta G_t. \tag{29}$$

Invoking the distributional assumptions outlined in Sections 2.1 and 2.2, as well as taking expectations to be fully rational, substitution of (29) into (7) yields a linear relationship between the levels of private and public consumption, with long-run parameter

$$\beta = -\theta. \tag{30}$$

This relationship highlights the long-run interpretation of θ and the fact that inferences regarding it are best produced within the context of a cointegration model.[6]

[5] Discussion on the relative merits of the use of this parameter vis-à-vis the largest eigenvalue of the characteristic polynomial associated with the $\Phi(L)$ is to be found in DeJong and Whiteman (1991) and Phillips (1991b).

[6] In contrast with the Aschauer derivation, no attempt is made in this paper to link explicitly the parameter β in (20) with the marginal rate of substitution in utility between private and public expenditure. In principle this link could be made by combining a solution for (8) with a specific form for the utility function in (1).

An alternative way to understand Aschauer's model, which brings out more clearly the effect of the rational expectations restrictions as well as the interpretation to be placed on θ, is as follows. Consider the following restricted version of the bivariate vector autoregressive (VAR) model in the levels of private consumption (C_t) and public consumption (G_t) which Aschauer (1985, equations 15a and 15b) estimates[7]

$$\begin{aligned} C_t &= \delta_1 + \delta_2 C_{t-1} + \delta_3 G_{t-1} + \delta_4 G_{t-2} + \kappa_t \qquad (31) \\ &\text{and} \\ G_t &= \lambda_1 + \lambda_2 G_{t-1} + \lambda_3 G_{t-2} + \zeta_t. \qquad (32) \end{aligned}$$

Equation (31) can, in turn, be written as an error correction model (ECM) as follows,

$$\Delta C_t = \delta_1 - (1-\delta_2)\left[C_{t-1} - \frac{\delta_3+\delta_4}{1-\delta_2}G_{t-1}\right] - \delta_4 \Delta G_{t-1} + \kappa_t. \qquad (33)$$

The adoption of the cross-equation restrictions arising from the assumption of rational expectations; see Aschauer (1985, equation 16),[8]

$$\begin{aligned} \delta_3 &= \theta(\delta_2 - \lambda_2) \text{ and} \\ \delta_4 &= -\theta\lambda_3, \qquad (34) \end{aligned}$$

where θ is as defined in (29), then produces

$$\Delta C_t = \delta_1 - (1-\delta_2)\left[C_{t-1} - \frac{\theta(\delta_2-\lambda_2-\lambda_3)}{1-\delta_2}G_{t-1}\right] + \theta\lambda_3 \Delta G_{t-1} + \kappa_t. \qquad (35)$$

Inspection of the parameter estimates reported in Table 2 in Aschauer (1985) shows that $\lambda_2 + \lambda_3 \approx 1.0$, as based on either the constrained or unconstrained model.[9] Substitution of this estimated sum into (35) gives

$$\Delta C_t = \delta_1 - (1-\delta_2)\left[C_{t-1} + \theta G_{t-1}\right] + \theta\lambda_3 \Delta G_{t-1} + \kappa_t,$$

which shows $-\theta$ to be a cointegrating parameter and thus interpretable as a measure of long-run crowding out.

[7] Aschauer (1985) also includes the variable D_t, the real per-capita net government deficit. This variable is not important for the long-run specification of the model and is thus excluded in order to simplify the derivations that follow. In order to avoid confusion arising from an overlap between Aschauer's parameter notation and that used in this paper, reference to Aschauers' equations are made using notation which differs from his.

[8] Aschauer (1985, equation 16) also defines an equation for δ_1. As this equation is not needed to identify θ, it is excluded for simplicity.

[9] That is, Aschauer finds that G_t is $I(1)$.

3 A Bayesian Inferential Method

This section sets out the procedures for producing Bayesian inferences about the model in (20) to (24). After describing an identification problem in the likelihood function in Section 3.1, a Jeffreys prior which offsets these problems is presented in Section 3.2. Section 3.3 then outlines the proposed MCMC scheme used to estimate the marginal posterior densities.

3.1 An Identification Problem in the Likelihood

Given the model specified in (20) to (24), the likelihood function for the full parameter set $\Theta = \{\alpha, \beta, \Sigma, \rho, \psi\}$ is

$$L(\Theta|c,g) \propto |\Sigma|^{-n/2} \exp\{\frac{-1}{2}\mathrm{tr}(\Sigma^{-1}S)\}, \tag{36}$$

where c and g denote the n-dimensional observation vectors for c_t and g_t, and $S = \sum_{t=1}^{n} e_t e_t'$ is the error sums of squares matrix. Further define

$$\begin{aligned} c_t^* &= \Phi(L)c_t = c_t - \rho' c_{t-1}^r, \\ g_t^* &= \Phi(L)g_t = g_t - \rho' g_{t-1}^r \text{ and} \\ \sigma_{11.2} &= \sigma_{11} - \sigma_{12}^2/\sigma_{22}, \end{aligned}$$

where $c_{t-1}^r = (c_{t-1}, \Delta c_{t-1}, \ldots, \Delta c_{t-[p-1]})'$, $g_{t-1}^r = (g_{t-1}, \Delta g_{t-1}, \ldots, \Delta g_{t-[p-1]})'$ and ρ is as defined at the end of Section 2.2. Decomposing the bivariate Normal density for $e_t = (\varepsilon_t, \eta_t)'$ into the product of the conditional density for ε_t given η_t and the marginal density for η_t, and using this decomposition in the construction of the likelihood function, an alternative form for (36) is

$$\begin{aligned} L(\Theta|c,g) \propto \sigma_{11.2}^{-n/2} \exp\{\tfrac{-1}{2\sigma_{11.2}} \sum_t [c_t^* - \alpha(1-\rho_1) - \beta g_t^* - (\sigma_{12}/\sigma_{22})\Psi(L)\Delta g_t]^2\} \\ \times \sigma_{22}^{-n/2} \exp\{\tfrac{-1}{2\sigma_{22}} \sum_t [\Psi(L)\Delta g_t]^2\}. \end{aligned} \tag{37}$$

This form of the likelihood renders transparent two identification problems:

(i) When $\rho_1 = 1$, α is not identified;

(ii) When $\rho_1 = 1$; $\rho_i = \psi_{i-1}$, $i = 1, 2, \ldots, \min(p,q)$ and any excess elements in either ρ or ψ equal zero, $g_t^* = \Psi(L)\Delta g_t$ for all t. As a consequence, in this region of the parameter space, β enters (37) via $(\beta + \sigma_{12}/\sigma_{22})$ and is not individually identified.

The regression parameters are thus not identified when there is an exact unit root in the error term. Lack of identification of α has been documented elsewhere, within the context of similar models; see, for example, Zellner (1971), Schotman and van Dijk (1991b) and Lubrano (1995). The lack of identification of β specifically occurs in a part of the parameter space in which the quantities g_t^* and $\Psi(L)\Delta g_t$ coincide for all t and is equivalent to a problem of exact multicollinearity. The multicollinearity results, in turn, from the fact that, given both the assumption of Normality and the allowance for the endogeneity of g_t via the specification of a full covariance matrix Σ, β enters the likelihood via the regression of c_t^* on g_t^* *augmented* by $\Psi(L)\Delta g_t$.

The lack of identification of α and β when $\rho_1 = 1$ can be shown to manifest itself in terms of a singularity at $\rho_1 = 1$ in the marginal posterior for ρ_1, based on a flat prior. Most importantly, the flatness in the likelihood function in the subspace surrounding $\rho_1 = 1$ produces asymptoting behavior in the marginal posterior for ρ_1 in the region of the singularity. This, in turn, distorts inferences by strongly favoring a lack of cointegration in cases where the underlying generating processes are cointegrated. A more detailed account of this pathological posterior behavior is given in the Appendix. Similar results, for an alternative form of cointegration model, are to be found in Kleibergen and van Dijk (1994).

3.2 A Conditional Jeffreys Prior

Following the work of Schotman and van Dijk (1991a), Kleibergen and van Dijk (1994 and 1996), Chao and Phillips (1996) and Martin (1997), the identification problem described above can be offset using a Jeffreys prior. Being proportional to the determinant of the information matrix and, hence, inversely related to the curvature of the likelihood, this form of prior is a natural candidate for counteracting the flatness in the likelihood associated with the identification problem. The likelihood in (37) fails to identify α and β in a subspace defined by certain values of ρ and ψ. As such, it is a Jeffreys prior for β, *conditional* on ρ and ψ, which has the potential to eliminate both the exact lack of identification in this subspace and the near identification problem in the surrounding subspace.

Assuming a priori independence between Σ and the remaining parameters, a joint prior for Θ has the form

$$p(\Theta) \propto p(\Sigma)p(\alpha, \beta, \rho, \psi). \tag{38}$$

For the first component in (38), a Jeffreys prior is specified as

$$p(\Sigma) = |I_\Sigma|^{1/2} \propto |\Sigma|^{-3/2}, \tag{39}$$

where I_Σ denotes the submatrix of the information matrix which relates to the elements of Σ, that is $I_\Sigma = E(-\partial^2 \ln L/\partial\Sigma\partial\Sigma')$. The second component in (38) is decomposed as

$$p(\alpha, \beta, \rho, \psi) \propto p(\alpha, \beta|\rho, \psi)p(\rho, \psi), \tag{40}$$

with the conditional prior $p(\alpha,\beta|\rho,\psi)$ in (40) being assigned the form of a Jeffreys prior. The derivation of a Jeffreys prior for $p(\alpha,\beta|\rho,\psi)$ proceeds via the decomposed likelihood in (37), as it is in that form that the likelihood makes explicit the conditioning relationships underlying the lack of identification, relationships which need to be incorporated in the Jeffreys prior.

As detailed in the Appendix, the conditional Jeffreys prior is specified as

$$p(\alpha,\beta|\rho,\psi) = \left|I_{\alpha,\beta|\rho,\psi}\right|^{1/2} \propto |A|^{1/2}, \tag{41}$$

where

$$A = (g^{**\prime}g^{**})([1^*,g^*]^\prime P_{g^{**}}[1^*,g^*]) \tag{42}$$

and the notation 1^*, g^* and g^{**} is used to denote respectively, the n-dimensional vectors for $1_t^* = (1-\rho_1)$, $g_t^* = g_t - \rho^\prime g_{t-1}^r$ and $g_t^{**} = \Psi(L)\Delta g_t$. The notation $P_{g^{**}}$ denotes the idempotent matrix

$$P_{g^{**}} = I - g^{**}(g^{**\prime}g^{**})g^{**\prime}. \tag{43}$$

As $\rho_1 \to 1$ and the remaining elements of ρ approach those of ψ in the manner described in (ii) in Section 3.1, $1^* \to 0$ and $g^* \to g^{**}$, either of which, in turn, implies that $|A|^{1/2} \to 0$. The prior in (41) thus places zero weight on the region where the lack of identification problem occurs. More importantly, as shown in the Appendix, it serves to counteract the distortion to posterior inferences which occurs in the surrounding subspace. In an extension of the type of arguments presented in Phillips (1991a and b), (41) can be viewed as being characteristic of a noninformative, or objective, prior and is, thus, true to the conventional motivation for using a Jeffreys prior. In particular, (41) reflects the information content of the likelihood in the appropriate manner, indicating that, conditional on ρ and ψ approaching the values which define the region in which the identification problem arises, the likelihood becomes less and less informative about α and β.

Using a uniform prior for ρ and ψ in (40), the explicit form of the joint prior in (38) is

$$p(\alpha,\beta,\Sigma,\rho,\psi) \propto |\Sigma|^{-3/2}|A|^{1/2}. \tag{44}$$

The implications of defining a uniform prior for the dynamic parameters ρ and ψ are discussed briefly in Section 4.3. Combining the joint prior in (44) with the likelihood in (36) yields the joint posterior

$$p(\alpha,\beta,\Sigma,\rho,\psi|c,g) \propto |\Sigma|^{-(n+3)/2}\exp\{\frac{-1}{2}\mathrm{tr}(\Sigma^{-1}S)\}|A|^{1/2}, \tag{45}$$

with $(\alpha,\beta,\Sigma,\rho,\psi)$ defined on the support $D = \mathbb{R}^1 \times \mathbb{R}^1 \times \mathbb{S}^{\mathrm{PDS}} \times \mathbb{R}^p \times \mathbb{R}^q$, where $\mathbb{R}^j$ denotes the Euclidean space of dimension j and $\mathbb{S}^{\mathrm{PDS}}$ denotes the space of (2×2) symmetric positive definite matrices. This is the expression which forms the basis of the Bayesian inferences concerning the presence of cointegration, the value of the long-run crowding out parameter β and the short-run dynamics.

3.3 Estimation of Marginal Posteriors via MCMC Sampling

The use of a deterministic numerical integration method in producing marginal density estimates for the pertinent parameters from the joint posterior in (45) is impractical, due to the large number of parameters. The approach adopted in this paper is to produce estimates of the marginal densities using a combination of MCMC sampling strategies, namely the Gibbs and Metropolis-Hastings (MH) algorithms. For recent papers discussing both the theory and implementation of MCMC procedures, see Geweke (1994), Roberts and Smith (1994), Tierney (1994) and Chib and Greenburg (1996).

The Gibbs sampler involves generating a sample from the joint posterior in (45) indirectly, via an iterative generation of random drawings from all of the associated conditional posterior distributions. Given the satisfaction of certain convergence criteria, these drawings represent a realization of a Markov chain with equilibrium distribution equal to the joint posterior. Continued application of the algorithm after the so-called "burn-in" period produces a sample of values from the joint distribution, from which various functions of interest, including the relevant marginal densities, can be estimated. The explicit inclusion of the intercept term α in the model causes convergence problems for the Gibbs sampler; see Martin (1996). This problem is circumvented by expressing the data in demeaned form and applying the MCMC scheme to the reduced parameter set $(\beta, \Sigma, \rho, \psi)$. Using the notation p_i^d to denote the density function with associated distribution function p_i ; $i = 1,2,3,4$, the full set of conditional posterior densities induced by (45), and as applied to demeaned data c and g, are as follows:

$$p_1^d(\beta|\Sigma, \rho, \psi, c, g) \propto \exp\{\frac{-1}{2\sigma_\beta^2}(\beta - \tilde{\beta})^2\}, \tag{46}$$

with $\tilde{\beta} = A_2^{-1}A_1$ and $\sigma_\beta^2 = \sigma_{11.2}A_2^{-1}$,

$$p_2^d(\Sigma|\beta,\rho, \psi, c, g) \propto |\Sigma|^{-(n+3)/2} \exp\{\frac{-1}{2}tr(\Sigma^{-1}S)\}, \tag{47}$$

$$p_3^d(\rho|\beta, \Sigma, \psi, c, g) \propto \exp\{\frac{-1}{2\sigma_\rho^2}(\rho - \tilde{\rho})^2\} \times (g^{*\prime}P_{g^{**}}g^*)^{1/2}, \tag{48}$$

with $\tilde{\rho} = B_2^{-1}B_1$ and $\sigma_\rho^2 = \sigma_{11.2}B_2^{-1}$, and

$$p_4^d(\psi|\beta, \Sigma, \rho, c, g) \propto \exp\{\frac{-1}{2\sigma_\psi^2}(\psi - \tilde{\psi})^2\} \times (g^{**\prime}P_{g^*}g^{**})^{1/2}, \tag{49}$$

with $\tilde{\psi} = E_2^{-1}E_1$ and $\sigma_\psi^2 = \sigma_{22.1}E_2^{-1}$. The additional notation used is defined as follows:

$$
\begin{aligned}
A_1 &= g^{*\prime}[c^* - (\sigma_{12}/\sigma_{22})g^{**}],\\
A_2 &= g^{*\prime}g^*,\\
B_1 &= U_p'[u - (\sigma_{12}/\sigma_{22})g^{**}],\\
B_2 &= U_p'U_p,\\
E_1 &= V_q'[v - (\sigma_{12}/\sigma_{11})u^*],\\
E_2 &= V_q'V_q \text{ and}\\
\sigma_{22.1} &= \sigma_{22} - \sigma_{12}^2/\sigma_{11},
\end{aligned}
$$

where $U_p = (u^1, u^1 - u^2, \ldots, u^{p-1} - u^p)$ and $V_q = (v^1, v^2, \ldots, v^q)$, with $u^j = (u_{1-j}, u_{2-j}, \ldots, u_{n-j})'$, $j = 1, 2, \ldots, p$ and $v^j = (v_{1-j}, v_{2-j}, \ldots, v_{n-j})'$, $j = 1, 2, \ldots, q$. The notation u^* and v respectively denotes the n-dimensional vectors associated with $u_t^* = u_t - \rho' u_{t-1}^r$ and v_t.

The steps of the Gibbs algorithm for the parameter set $(\beta, \Sigma, \rho, \psi)$ are as follows:

Step 1 Specify initial values for Σ, ρ and ψ, $\Sigma^{(0)}$, $\rho^{(0)}$ and $\psi^{(0)}$.

Step 2 Cycle iteratively through the four conditional distributions, drawing respectively:

1. $\beta^{(i)}$ from $p_1(\beta^{(i)}|\Sigma^{(i-1)}, \rho^{(i-1)}, \psi^{(i-1)}, c, g)$
2. $\Sigma^{(i)}$ from $p_2(\Sigma^{(i)}|\beta^{(i)}, \rho^{(i-1)}, \psi^{(i-1)}, c, g)$
3. $\rho^{(i)}$ from $p_3(\rho^{(i)}|\beta^{(i)}, \Sigma^{(i)}, \psi^{(i-1)}, c, g)$
4. $\psi^{(i)}$ from $p_4(\psi^{(i)}|\beta^{(i)}, \Sigma^{(i)}, \rho^{(i)}, c, g)$

until $i = M$.

In practice, a finite value of M needs to be chosen, but one that is large enough to ensure that convergence to the joint posterior distribution has been achieved. Continued application of the algorithm for a further T iterations after the burn-in period M, produces a sample of size T values of $(\beta^{(i)}, \Sigma^{(i)}, \rho^{(i)}, \psi^{(i)})$ from the joint posterior distribution, from which the marginal densities are estimated.

The distributions for β and Σ associated with the densities in (46) and (47) are respectively univariate Normal and inverted Wishart. These distributions can be simulated from directly, using a Normal random number generator. However, the presence of the Jeffreys prior for β, conditional on ρ and ψ, renders the distributions of the latter two parameter vectors nonstandard in form. The approach adopted here is to simulate from these distributions via the insertion of MH subchains at the relevant points in the Gibbs sampler. With reference to ρ, the MH algorithm involves simulating a value for ρ at the ith iteration of the Gibbs chain via a so-called candidate distribution q_3. The Normal kernel factor, as forms part of the conditional density for p_3^d in (48), is used as the basis for a candidate distribution, producing a candidate value for ρ, ρ^* say. With $\rho^{(i-1)}$ being the value for ρ as produced in the $(i-1)th$ iteration of the Gibbs/MH algorithm, ρ^* is accepted as a simulated value from p_3 with probability δ, where

$$
\begin{aligned}
\delta(\rho^{(i-1)}, \rho^*) &= \min\{(p_3^d(\rho^*)/q_3^d(\rho^*))/(p_3^d(\rho^{(i-1)})/q_3^d(\rho^{(i-1)})), 1\} \\
& \qquad\qquad if \quad p_3^d(\rho^{(i-1)})q_3^d(\rho^*) > 0 \\
&= 1 \\
& \qquad\qquad if \quad p_3^d(\rho^{(i-1)})q_3^d(\rho^*) = 0.
\end{aligned}
$$

The MH algorithm is iterated prior to a simulated value being included in the outer Gibbs chain as $\rho^{(i)}$. As the parameterization of q_3^d remains fixed and the ratio $p_3^d(.)/q_3^d(.)$ can be shown to be bounded, by Tierney (1994) the iterated values of ρ form a Markov chain which are *uniformly ergodic* for p_3. The algorithm also satisfies the practical criterion of producing a high average proportion of acceptances across subchains. The same procedure is used in the generation of ψ via a candidate distribution based on the Normal kernel which forms part of the conditional density for ψ, given by p_4^d in equation (49). With reference to Roberts and Smith (1994) and Tierney (1994), the outer Gibbs chain can be shown to satisfy the sufficient conditions for convergence to the joint posterior and is *(simply) ergodic* for the joint distribution as a consequence.

With T being the number of simulated sets of parameter values produced after the burn-in period by the hybrid algorithm, the marginal densities of interest are estimated as finite mixtures. Demonstrated for the β marginal, the estimate is given by

$$\widehat{p(\beta|c, g)} = (1/T)\sum_{i=1}^{T} p_1^d(\beta|\Sigma^{(i)}, \rho^{(i)}, \psi^{(i)}, c, g). \tag{50}$$

This expression represents the sample estimate of the expectation implicit in the relationship between a conditional and marginal density. Computationally this involves evaluating each of the T conditional densities, p_1^d, over a grid of β values. For each value of β in the grid there are T values of the conditional densities which are simply averaged to form the ordinate value of the marginal density of β at that point.

The estimate of the marginal density of α is constructed from the output of the MCMC scheme as

$$\widehat{p(\alpha|c, g)} = (1/T)\sum_{i=1}^{T} p^d(\alpha|\beta^{(i)}, \Sigma^{(i)}, \rho^{(i)}, \psi^{(i)}, c, g), \tag{51}$$

where the conditioning values for the remaining parameters are those produced by the reduced scheme applied to the demeaned data. The conditional density to be averaged across in (51) is the univariate Normal conditional for α implied by the bivariate conditional for the vector $(\alpha, \beta)'$, which is, in turn, induced by the joint posterior in (45), with c and g entering the conditional density in *raw* form.

In order to improve the accuracy of the density estimates via a reduction of the Markovian dependence in the sample (see Lui, Wong and Kong, 1994), the output of every $10th$ iteration from the outer Gibbs chain is used in the density estimation. Thus, the number of iterations performed after the burnout period M is actually $(T \times 10)$.

4 Empirical Results for the G7 Countries

4.1 Data and Computational Details

The data consist of real per capita biannual private and public consumption expenditure for the G7 countries over the period 1960 to 1996. All data are obtained from the dX EconData database file EOL-C7S. Per capita data are used in order to render the empirical analysis of the crowding out model consistent with the individual optimization problem from which it has been derived. With the initial seven observations used for testing lag orders, the sample size on which the estimation is based is 67.

As explained in Section 2.1, the raw data is transformed into logarithmic form as dictated by the model specification. The data are then demeaned to circumvent the problems alluded to earlier regarding the convergence of the Gibbs sampler when applied to a model which includes an intercept term. Finally, the data are detrended to extract any deterministic trends by regressing the demeaned logarithmic data on a linear time trend.

The marginal posterior densities are estimated via the MCMC strategy outlined in the previous section, with $T = 2000$ simulated sets of parameter values produced after a burnout period of $M = 500$. In all instances in which an MH algorithm is inserted, namely when producing simulated values for ρ and ψ at each iteration of the Gibbs sampler, 20 iterations of the subchain are performed before a value is taken as a realization from the relevant conditional distribution. It has been found that any increase in these specifications produces essentially unchanged density estimates.

4.2 Univariate Unit Root Testing

The results of preliminary univariate unit root testing for the seven countries are presented in Table 1. The posterior information criterion (PIC) model selection procedure of Phillips and Ploberger (1994) is used to select the lag length in univariate AR specifications for all of the individual c_t and g_t series. The selected lag lengths are reported in column 2 of Table 1. Imposing the lag length as chosen by the PIC model selection criterion, each of the individual series is then tested for a unit root using the PIC odds ratio test. As explained by Phillips and Ploberger, the PIC odds ratio represents an objective Bayesian assessment of the unit root hypothesis, with no inherent tendency to conclude either in favor of or against the presence of a unit root. If prior odds of unity are specified, the criterion leads to rejection of the unit root hypothesis if PIC > 1.[10] For comparison, the results of augmented Dickey Fuller (ADF) and Phillips and Perron (1988) (PP) unit root tests are also reported in Table 1.

The results in Table 1 show that in all but three instances, the Bayesian and classical tests fail to reject the hypothesis that c_t and g_t are $I(1)$. In the case of the US, the rejection of a unit root in c_t is fairly clear cut according to both the PIC and ADF tests, but not according to the PP test. In the case of the UK c_t series, the PIC test only marginally rejects in favor of stationarity.

4.3 Inferences about Cointegration

With a qualification regarding the relevance of cointegration testing in the US case in particular, the next step is to test for cointegration between c_t and g_t for all seven countries. In Table 2, the results of the preliminary application of the PIC model selection procedure, as used to select the lag length for the AR error structures in (22) and (23), are presented. Given the imposition of a unit root in g_t, the lag length for the autoregression in v_t is chosen from alternative AR specifications for the series Δg_t. In choosing the lag length for the autoregression in u_t, the PIC procedure is applied to the residuals resulting from OLS estimation of (20). Included in Table 2 are the PIC odds ratios and the ADF and PP test statistics for the test of a unit root in the OLS residuals of (20), with the lag length for both the PIC and ADF procedures in the univariate specifications for the residuals set by the preliminary PIC selection procedure. Values of PIC exceeding unity constitute rejection of the null hypothesis of no cointegration.

The contrast between the Bayesian and classical results in Table 2 is very marked. The Bayesian odds ratios favor cointegration in six of the seven countries, with the seventh result, pertaining to Japan, going very close to rejection of the null of no cointegration. The PP test, on the other hand, fails to reject in favor of cointegration for all seven countries. The ADF test supports cointegration in the case of the US and Italy alone, with these two

[10] Obviously, if an investigator has a strong a priori belief that a series is indeed $I(1)$, then the evidence from the data in favour of stationarity, as measured by the PIC statistic, must be stronger for it to lead to the rejection of the unit root hypothesis.

Table 1:

Univariate unit root testing

(Bayesian and classical)

Variable	Lag[(a)]	PIC[(b)] odds ratio test	PP[(c)] unit root test	ADF[(c)] unit root test
US c_t	2	42.848*	-3.001	-4.171*
US g_t	3	0.673	-2.063	-2.855
Japan c_t	1	0.281	-2.160	-2.263
Japan g_t	1	0.043	-0.848	-0.653
Germany c_t	2	0.066	-1.907	−1.788
Germany g_t	1	0.044	-1.330	-1.025
France c_t	1	0.112	-1.940	-1.831
France g_t	1	0.050	-1.527	-1.305
UK c_t	2	1.137*	-2.375	-2.548
UK g_t	2	0.198	-1.726	-2.111
Italy c_t	3	0.050	-1.597	-1.794
Italy g_t	2	0.018	-1.208	-0.829
Canada c_t	2	0.101	-1.706	-1.618
Canada g_t	1	0.037	-1.444	-0.919

(a) Lag length as chosen by the preliminary PIC model selection procedure. Lag lengths of 1 to 4 are assessed against a maximum lag length of 5. The precise criterion applied is given in Phillips and Ploberger (1994, equation 25). These lag lengths are used in both the PIC and ADF unit root tests.

(b) The PIC odds ratio, as based on prior odds of unity, leads to the rejection of the hypothesis of a unit root if it exceeds one. * indicates values which lead to rejection.

(c) The ADF and PP test statistics are computed by the COINT module in GAUSS. The 5% critical value for the tests, as produced by the module, is −3.540. The PP results quoted are based on a lag length of 10 used in computing the nonparametric variance estimate. Results for lag lengths of 5 and 20 are not qualitatively different. * indicates values which lead to rejection of a unit root at the 5% level.

Table 2:

Residual-based cointegration testing

(Bayesian and classical)

Country	Polynomial	Lag[(a)]	PIC[(b)] odds ratio test	PP[(c)] coint. test	ADF[(c)] coint. test
US	$\Phi(L)$	2	1619.142*	-3.004	-4.665*
	$\Psi(L)$	2			
Japan	$\Phi(L)$	2	0.986	-2.886	-2.711
	$\Psi(L)$	1			
Germany	$\Phi(L)$	2	4.601*	-2.651	-3.050
	$\Psi(L)$	1			
France	$\Phi(L)$	1	1.788*	-2.882	-2.863
	$\Psi(L)$	1			
UK	$\Phi(L)$	2	1.276*	-2.421	-2.607
	$\Psi(L)$	1			
Italy	$\Phi(L)$	2	329.263*	-2.638	-4.199*
	$\Psi(L)$	3			
Canada	$\Phi(L)$	2	1.809*	-2.622	-2.641
	$\Psi(L)$	3			

(a) Lag lengths as chosen by the preliminary PIC criterion. Lag lengths of 1 to 4 are assessed against a maximum lag length of 5.

(b) The PIC odds ratio, as based on prior odds of unity, leads to the rejection of the hypothesis of a unit root in the residuals when it exceeds one. * indicates values which lead to rejection..

(c) The 5% critical value for the ADF and PP tests, as produced by the COINT module, is -3.938. As with the univariate results, the PP statistics quoted are based on a lag length of 10 used in the computation of the nonparametric variance estimate. The results for lag lengths of 5 and 20 are not qualitatively different. * indicates values which lead to rejection of no cointegration at the 5% level.

countries corresponding to those for which the Bayesian support is strongest.

— **Figure 1 about here** —

The Bayesian support for cointegration in Table 2 is, in general, substantiated by the results in Table 3, in which the posterior probability of cointegration, calculated as the area to the left of $\rho_1 = 1$ in the marginal posterior density of ρ_1, is given. For all countries with the exception of Canada, the posterior probabilities in Table 3 are greater than 97%, thereby providing very strong evidence of cointegration. There is overwhelming support for cointegration in the case of the US, the UK and Italy, where the total mass of the ρ_1 posterior density is to the left of the point $\rho_1 = 1$, and strong support in the case of Japan, Germany and France. For Canada, the posterior probability of cointegration is about 90%, indicating that there is approximately a 10% chance that private and public expenditure are not cointegrated. Also included in Table 3 are the modal point estimates of ρ_1, all of which fall in the stationary region. Further features of the marginal posterior densities of ρ_1 are highlighted in 1. All of the ρ_1 posterior marginals are essentially symmetric around their modal values, with the large probabilities assigned to the hypothesis of cointegration being made clear by the position of the densities in the support of ρ_1.

It should be noted that, following the arguments presented in Phillips (1991a and b) and Zivot and Phillips (1991), the flat prior on ρ_1, in particular, may slightly bias the results in Table 3 and Figure 1 in the direction of cointegration. This is in contrast with the PIC results in Table 2, which represent more objective evidence in favor of cointegration.

4.4 Long-run Substitutability

Table 4 gives both Bayesian and classical estimates of the long-run substitution parameter β in (20). The Bayesian point and interval estimates are respectively the mode and highest probability density (HPD) intervals constructed from the marginal posterior density for β.[11] The classical point estimates are based on the FMOLS procedure of Phillips and Hansen (1990), with the confidence intervals based on the asymptotic Standard Normal approximation. Given that the Bayesian evidence in favor of the existence of cointegration is strong for five of the countries considered, namely the US, Germany, France, the UK

[11] An HPD interval is defined as an interval with the specified probability coverage, whose inner density ordinates are not exceeded by any density ordinates outside the interval. For the current problem, in which the marginal posterior densities are unimodal, specification of an HPD interval involves choosing two values of β which yield ordinates of equal height and encompass the required probability content.

Table 3:

The posterior probability
of cointegration

Country	$\Pr(\rho_1 < 1)$	ρ_1 mode
US	1.000	0.841
Japan	0.983	0.951
Germany	0.984	0.911
France	0.979	0.921
UK	1.000	0.851
Italy	1.000	0.761
Canada	0.903	0.961

and Italy, with weaker support in the case of Japan and Canada, β is to be interpreted as a cointegrating parameter, with its sign indicating the direction of the impact of government on private expenditure in the long-run, and its magnitude indicating the percentage long-run response of private expenditure to a one per cent increase in public expenditure.[12]

The results in Table 4 can be summarized as follows. For the US, Germany, France and Italy, most of the evidence, both Bayesian and classical, points towards long-run crowding in. The support for crowding in is strongest for Germany and Italy, with the Bayesian and classical point estimates for β being large and positive and all interval estimates excluding zero and negative values. The evidence of long-run crowding in for France and the US is also strong, with the classical confidence intervals excluding nonpositive values and the 90% HPD intervals just extending into the negative region in each case. Only in the case of the UK does the evidence tend to favor long-run crowding out. Both the Bayesian and classical point estimates are negative and all interval estimates feature the negative region.

For both Japan and Canada, the situation is not clear-cut, with a conflict arising between the Bayesian and classical results. The Bayesian point esti-

[12] It should be recalled that the result for the US needs to be qualified, given the rejection of a unit root in the c_t series by both the PIC odds ratio and ADF tests.

Table 4:

Estimates of the crowding out parameter β

(Bayesian and classical)

Country	β mode	90% HPD interval	95% HPD interval	FMOLS estimate	95% conf. interval
US	0.180	(-0.010, 0.340)	(-0.050, 0.370)	0.276	(0.087, 0.465)
Japan	0.180	(-0.840, 0.860)	(-1.160, 1.000)	0.764	(0.212, 1.316)
Germany	0.820	(0.180, 1.120)	(0.000, 1.200)	1.016	(0.813, 1.220)
France	1.040	(-0.020, 1.640)	(-0.740, 1.800)	1.500	(1.198, 1.801)
UK	-0.380	(-0.980, 0.040)	(-1.140, 0.140)	-0.164	(-0.542, 0.214)
Italy	1.400	(1.200, 1.580)	(1.140, 1.640)	1.421	(1.120, 1.644)
Canada	-0.060	(-0.300, 0.220)	(-0.360, 0.280)	0.546	(0.320, 0.773)

mate for Japan is positive. However, the 90% and 95% HPD Bayesian intervals extend well into both the positive and negative regions, leading to distinct uncertainty as regards the sign of β. This result is supported by the marginal posterior density of β for Japan in Figure 2, where the bulk of the probability mass is seen to lie in the positive region, but with negative skewness producing a long left-hand tail. Figure 2 shows the marginal posterior density of β for Canada to be virtually symmetric around a modal estimate close to zero, with this result being reflected in the point and interval estimates reported in Table 4. In contrast to the ambiguity in the Bayesian results, the classical results for both countries strongly suggest crowding in. This conflict between the Bayesian and classical results may well be a reflection of the slightly doubtful status attached to the cointegrating relationship between private and public consumption expenditure as identified in Table 2 in the case of Japan and in Table 3 in the case of Canada.

— Figure 2 about here —

A comparison of the Bayesian 95% HPD intervals and the FMOLS 95% confidence intervals in Table 4 reveal that the latter are more narrow. However, the classical intervals do have asymptotic justification only and may be misleading given a sample size of 67. Also, the dispersion in the marginal posteriors for β reflects, in part, the non-zero probability of certain of the relationships not being cointegrated. This is an appropriate feature of a method which produces simultaneous, rather than sequential inferences regarding cointegration.

The evidence in favor of crowding in for the US in particular, is in contrast with the findings of Aschauer (1985), whereby the hypothesis of crowding out in the US, as tested in varying specifications of his model, is rarely rejected. However, a closer examination of Aschauer's figures reveals that the imposition of the cross equation restrictions has in fact had a drastic influence on his results, with the unrestricted estimates reported implying long-run crowding in. From equation (31), the long-run multiplier linking C_t to G_t in the Aschauer specification is $\beta = (\delta_3 + \delta_4)/(1 - \delta_2)$. From Table 2 in Aschauer, the unconstrained estimates produce $\widehat{\beta} = (-0.024 + 0.035)/(1 - 0.990) = 1.10$. Comparison with the estimate obtained from the constrained model, namely $\widehat{\beta} = -\widehat{\theta} = -0.229$, shows that the imposition of the cross equation restrictions has caused a reversal in the conclusions to be drawn regarding substitutability. Since the estimation procedure used in this paper accommodates rational expectations but does not impose severe parameter restrictions in the process of doing so, the validity or otherwise of the rational expectations hypothesis is not able to bear so heavily on the empirical results.[13]

[13]Note that the magnitude of Aschauer's estimates of β cannot be compared directly with

4.5 Short-run Substitutability

The discussion so far has concentrated on estimates of the long-run substitution between private and public expenditure. It is also of interest to investigate the dynamic time path of the response of private to public expenditure, so as to determine the speed of convergence to the long-run as well as the behavior of the dynamic path over the short-run and intermediate-run.

The dynamic multipliers are computed by estimating the impulse response functions associated with the model in (20) to (24), where the choice of the lag structures for each country is as given in Table 2. The ordering of the variables is g_t to c_t. This is appropriate given the exogeneity status of g_t, as discussed in Section 2.2. By computing the impulse response functions for $\partial c_{t+h}/\partial \varepsilon_t$ and $\partial g_{t+h}/\partial \varepsilon_t$, $h > 0$, that is the response of c_t and g_t to an unanticipated shock to g_t, the dynamic multiplier, $\partial c_{t+h}/\partial g_t$, is given by the ratio of these two expressions.

The dynamic multipliers and associated 90% probability intervals are given in Figure 3 for the G7 countries. The point estimates of the dynamic multipliers are based on the Bayesian modal estimates of $(\beta, \Sigma, \phi, \psi)$, produced via the MCMC scheme outlined in Section 3.3, with $M = 500$ and $T = 2000$. The point estimates of $\rho_1 = \sum_{j=1}^{p} \phi_j$ and β underlying the dynamic multipliers are thus equivalent to those reported in Sections 4.3 and 4.4 respectively.[14] The 90% probability intervals are computed using the following procedure. First, the $T = 2000$ simulated sets of parameter values of $(\beta, \Sigma, \phi, \psi)$ produced by the hybrid MCMC algorithm are stored. For each set of parameter values the impulse response functions are computed and the dynamic multipliers formed. The upper and lower bounds of the probability interval for each multiplier are found by ordering the $T = 2000$ impulse response values at each time point and choosing the upper (lower) bound as that value which has 5% of values greater (less) than it.[15]

— Figure 3 about here —

The point estimates of the dynamic multipliers reproduced in Figure 3 show the G7 countries as falling into three groups. The first group consists of the US, Germany, France and Italy. For this group, private consumption is crowded out in the short-run by public expenditure whilst in the long-run

the estimates given in Table 4, as his model is estimated in terms of levels, whilst the model estimated in this paper is in terms of logarithms of the variables.

[14] For computational convenience, the point estimates of the elements of Σ are taken as the modes of the histograms of simulated values produced by the MCMC algorithm for each element and not as the modes of the mixture density estimates of the relevant marginals.

[15] An alternative, more computationally burdensome, approach would be to estimate the marginal density associated with each dynamic multiplier by a nonparametric kernel method and then use the estimated densities to construct HPD intervals.

there is crowding in. The size of (the point estimate of) long-run crowding in given in Figure 3 is, by construction, equal to the corresponding estimate of β given in Table 4.[16] Based on the point estimates, short-run crowding out lasts for at least 1 year (France and Italy) and no more than 2 years (Germany). The second group comprises the UK, where the reverse results are found: there is short-run crowding in the first 2.5 years, but long-run crowding out. The third group consists of Japan and Canada. For this group the point estimate of the crowding out parameter is close to zero in both the short-run and the long-run.

The probability intervals reproduced in Figure 3 corroborate qualitatively the above interpretations of the short-run/long-run dichotomy in the case of the US, the UK and Italy. In the case of Germany and France, the probability intervals indicate some uncertainty as to the sign of the crowding out parameter in the short-run. For Japan and Canada, as is consistent with all results pertaining to these countries discussed previously in the paper, there is distinct uncertainty as regards both the sign and magnitude of the crowding out effect in both the short-run and the long-run.[17]

5 Conclusions

This paper has provided an empirical analysis of the substitution between private and public real, per capita consumption expenditure for the G7 countries over the period 1960 to 1996. Using Bayesian inferential techniques, the US, Germany, France and Italy have been found, overall, to exhibit short-run substitutability, but long-run complementarity. The opposite result occurs for the UK.

For Japan and Canada, no significant short-run or long-run relationship is found. This result may be interpreted as either reflecting that private consumption expenditure does not respond to public consumption expenditure, or that the empirical model is misspecified. This latter interpretation may be appropriate for the long-run model specification in particular, as there is some doubt over the cointegration status of private and public consumption expenditure in the case of these two countries.

[16] Whilst the long-run estimates of β in Figure 3 are, as noted, the same as the estimates given in Table 4, the associated 90% confidence intervals are computed differently. The confidence intervals in Table 4 are computed from a mixture estimate of the marginal posterior of β as given by (50), whereas the estimates given in Figure 3 are obtained from simulating the impulse response function, as described in the text. However, a comparison of the two different 90% confidence intervals for all countries shows that similar estimates are obtained.

[17] The finding of short-run crowding out for the US specifically tallies with the results in Graham (1993), based on regressing the change in C_t on the change in G_t. This result serves to highlight that regression equations based on first differences at best provide information on short-run substitution between private and public consumption expenditure.

References

[1] Aschauer, D.A., 1985, Fiscal policy and aggregate demand, *The American Economic Review* 75, 117-127.

[2] Aschauer, D.A., 1989a, Is public expenditure productive?, *Journal of Monetary Economics* 23, 177-200.

[3] Aschauer, D.A., 1989b, Does public capital crowd out private capital?, *Journal of Monetary Economics* 24, 171-188..

[4] Bauwens, L. and M. Lubrano, 1993, Identification restrictions and posterior densities in cointegrated Gaussian VAR systems, in *Advances in Econometrics,* Vol. 11b, ed. T.M. Fomby and R. Carter Hill, (JAI Press, Greenwich, CT).

[5] Campbell, J.Y. and N.G. Mankiw, 1990, Permanent income, current income and consumption, *Journal of Business and Economic Statistics* 8, No. 3, 265-279.

[6] Cebula, R.J., Killingsworth, J. and W.J. Belton, Jr, 1994, Federal government budget deficits and the crowding out of private investment in the United States, *Public Finance* 49, No. 2, 169-178.

[7] Chao, J.C. and P.C.B. Phillips, 1996, Bayesian posterior distributions in limited information analysis of the simultaneous equations model, *Mimeo, Yale University.*

[8] Chib, S. and E. Greenburg, 1996, Markov chain Monte Carlo simulation methods in Econometrics, *Econometric Theory* 12, 409-431.

[9] Darby, J and J. Malley, 1996, Fiscal policy and aggregate consumption: new evidence from the United States, *Scottish Journal of Political Economy* 43, No. 2, 129-145.

[10] DeJong, D.N. and C.H. Whiteman, 1991, The case for trend-stationarity is stronger than we thought, *Journal of Applied Econometrics* 6, 413-421.

[11] Drèze, J.H., 1977, Bayesian regression analysis using poly-t densities, *Journal of Econometrics* 6, 329-354.

[12] Feldstein, M., 1982, Government deficits and aggregate demand, *Journal of Monetary Economics* 9, 1-20.

[13] Geweke, J., 1994, Monte Carlo simulation and numerical integration, *Working Paper, University of Minnesota and Federal Reserve Bank of Minneapolis.*

[14] Geweke, J., 1996, Bayesian reduced rank regression in econometrics, *Journal of Econometrics* 75, No. 1, 121-146.

[15] Graham, F.C., 1993, Fiscal policy and aggregate demand: comment, *The American Economic Review* 83, 659-666.

[16] Graham, F.C. and D. Himarios, 1991, Fiscal policy and private consumption: instrumental variables tests of the "consolidated approach", *Journal of Money, Credit and Banking* 23, No. 1, 53-67.

[17] Kleibergen, F., and H.K. van Dijk, 1994, On the shape of the likelihood/posterior in cointegration models, *Econometric Theory* 10, 514-551.

[18] Kleibergen, F., and H.K. van Dijk, 1996, Bayesian simultaneous equation analysis using reduced rank structures, *Working paper, Econometric Institute and Tinbergen Institute, Rotterdam, The Netherlands.*

[19] Kormendi, R.C., 1983, Government debt, government spending and private sector behaviour, *American Economic Review* 73, 994-1010.

[20] Leiderman, L. and A. Razin, 1988, Testing Ricardian neutrality with an intertemporal stochastic model, *Journal of Money, Credit and Banking* 20, No. 1, 1-21.

[21] Lubrano, M., 1995, Testing for unit roots in a Bayesian framework, *Journal of Econometrics* 69, 81-109.

[22] Lui, J.S., Wong,W.H. and A. Kong, 1994, Covariance structure of the Gibbs sampler with applications to the comparisons of estimators and augmentation schemes, *Biometrika* 81, 27-40.

[23] Lucas, R.E. Jr. and L.A. Rapping, 1969, Real wages, employment and inflation, *Journal of Political Economy* 77, 721-754.

[24] Martin, G.M., 1996, Bayesian inference in models of cointegration: methods and applications, *Unpublished Ph.D. Thesis, Monash University.*

[25] Martin, G.M., 1997, Fractional cointegration: Bayesian inferences using a Jeffreys prior, *Manuscript, Monash University.*

[26] Phillips, P.C.B., 1988, Reflections on econometric methodology, *Economic Record* 64, 544-559.

[27] Phillips, P.C.B., 1991a, To criticize the critics: an objective Bayesian analysis of stochastic trends, *Journal of Applied Econometrics* 6, 333-364.

[28] Phillips, P.C.B., 1991b, Bayesian routes and unit roots: de rebus prioribus semper est disputandum, *Journal of Applied Econometrics* 6, 435-474.

[29] Phillips, P.C.B., 1996, Econometric model determination, *Econometrica* 64, 763-812.

[30] Phillips, P.C.B. and B.E. Hansen, 1990, Statistical inference in instrumental variables regression with I(1) processes, *Review of Economic Studies* 57, 99-125.

[31] Phillips, P.C.B. and M. Loretan, 1991, Estimating long-run economic equilibria, *Review of Economic Studies* 58, 407-436.

[32] Phillips, P.C.B. and P. Perron, 1988, Testing for a unit root in time series regression, *Biometrika* 75, No. 2, 335-346.

[33] Phillips, P.C.B. and W. Ploberger, 1994, Posterior odds testing for a unit root with data-based model selection, *Econometric Theory* 10, 774-808.

[34] Phillips, P.C.B. and W. Ploberger, 1995, An asymptotic theory of Bayesian inference for time series, *Econometrica* 64, 381-412.

[35] Roberts, G.O. and A.F.M. Smith, 1994, Simple conditions for the convergence of the Gibbs sampler and Metropolis-Hastings algorithms, *Stochastic Processes and their Applications* 49, 207-216.

[36] Schotman, P. and H.K. van Dijk, 1991a, On Bayesian routes to unit roots, *Journal of Applied Econometrics* 6, 387-401.

[37] Schotman, P. and H.K. van Dijk, 1991b, A Bayesian analysis of the unit root in real exchange rates, *Journal of Econometrics* 49, 195-238.

[38] Tierney, L., 1994, Markov chains for exploring posterior distributions, *The Annals of Statistics* 22, No.4, 1701-1762.

[39] Zellner, A., 1971, *An Introduction to Bayesian Inference in Econometrics,* (Wiley, New York).

[40] Zivot, E. and P.C.B. Phillips, 1991, A Bayesian analysis of trend determination in economic time series, *Cowles Foundation Discussion Paper No. 1002, Yale University.*

Appendix: Derivation of the Conditional Jeffreys Prior

The derivation of the conditional Jeffreys prior for (α, β) given ρ and ψ, as given in (41) proceeds as follows. Given the assumption of independence between Σ and the remaining parameters, the element σ_{12}/σ_{22} in the first exponent term in (37) can be replaced by an artificial parameter, γ say. Letting $Z = (1^*, g^*, g^{**})$, where 1^*, g^* and g^{**} are as defined in Section 3.2, the Jeffreys prior is given by

$$\left|I_{\delta|\rho,\psi}\right|^{1/2} \propto |E(Z'Z)|^{1/2}, \tag{52}$$

where E denotes expectations with respect to the assumed generating process. With the artificial parameter γ serving only to reflect the manner in which the regression parameters α and β enter the likelihood, that is, via the augmented regression, and serving no inferential purpose in itself, (52) can be viewed as a conditional prior for α and β. Given the nature of the assumed model, evaluation of the expectations in (52) produces a prior which is a function of σ_{22}, n, ρ, ψ and the fixed initial values of g_t. As $\rho_1 \to 1$ and the remaining elements of ρ approach those of ψ in the manner described in (ii) in Section 3.1, $1^* \to 0$ and $g^* \to g^{**}$, either of which, in turn, implies that $\left|I_{\delta|\rho,\psi}\right|^{1/2} \to 0$. Use of this prior would thus serve to counteract the singularity in the density of ρ and ψ in this region. However, only if the expectations in (52) are *approximated* by the realized values, does the conditional Jeffreys prior serve to offset exactly the distortion to inferences which would occur with a flat prior analysis.[18]

This can be most easily demonstrated within the context of a simplified version of the model, with $\alpha = 0$, $\Phi(L) = 1 - \rho_1 L$ and $\Psi(L) = 1$. Applying the standard diffuse prior for a bivariate regression model, $p(\beta, \Sigma, \rho_1) \propto |\Sigma|^{-3/2}$, the marginal density for ρ_1 can be derived analytically as

$$p(\rho_1|c, g) \propto A^{-1/2} B^{-(n-1)/2}, \tag{53}$$

where

$$\begin{aligned} A &= (\Delta g'\Delta g)(g^{*\prime} P_{\Delta g} g^*), \\ B &= (\Delta g'\Delta g)(c^{*\prime} P_{\Delta g} P_{P_{\Delta g} g^*} P_{\Delta g} c^*) \end{aligned} \tag{54}$$

and c^*, g^* and Δg denote respectively the n-dimensional vectors associated with $c_t^* = (1 - \rho_1 L)c_t$, $g_t^* = (1 - \rho_1 L) g_t$ and Δg_t. In (54), as well as in following expressions, the notation $P_y = I - y(y'y)^{-1}y'$ is used. Using the result that

$$P_{\Delta g}\Delta g = 0 \Rightarrow P_{\Delta g} g = P_{\Delta g} g_{-1} \Rightarrow P_{\Delta g} g^* = (1 - \rho_1) P_{\Delta g} g, \tag{55}$$

[18] Kleibergen and van Dijk (1994) provide extensive discussion on the impact of alternative forms of expectation evaluation in a Jeffreys prior for a cointegration model.

with $g_{-1} = (g_0, g_1, \ldots g_{n-1})'$, it follows that

$$A^{-1/2} \propto |1 - \rho_1|^{-1/2}. \tag{56}$$

The factor $A^{-1/2}$ is thus nonintegrable as a consequence of both its tail behavior and its behavior in the region of the singularity at $\rho_1 = 1$. This contrasts with the factor $B^{-(n-1)/2}$, which is given by

$$B^{-(n-1)/2} \propto (\frac{h}{\kappa})^{1/2}\{1 + \frac{h}{\kappa}(\rho_1 - \bar{\rho}_1)^2\}^{-(n-1)/2}, \tag{57}$$

where

$$\begin{aligned}
\bar{\rho}_1 &= (c_{-1}P'_{\Delta g}P_{P_{\Delta g}g^*}P_{\Delta g}c)/(c_{-1}P'_{\Delta g}P_{P_{\Delta g}g^*}P_{\Delta g}c_{-1}), \\
h/\kappa &= (c_{-1}P'_{\Delta g}P_{P_{\Delta g}g^*}P_{\Delta g}c_{-1})/d, \\
d &= (cP'_{\Delta g}P_{P_{\Delta g}g^*}P_{\Delta g}c) - (c_{-1}P'_{\Delta g}P_{P_{\Delta g}g^*}P_{\Delta g}c)^2/(c_{-1}P'_{\Delta g}P_{P_{\Delta g}g^*}P_{\Delta g}c_{-1})
\end{aligned}$$

and $c_{-1} = (c_0, c_1, \ldots c_{n-1})'$. Once again applying (55), it can be shown that both $\bar{\rho}_1$ and h are constant with respect to ρ_1, in which case (57) defines, for $\kappa > 2$, a Student t density for ρ_1 with mean $\bar{\rho}_1$, variance $(\frac{\kappa}{\kappa-2})(\frac{1}{h})$ and degrees of freedom $\kappa = n - 2$. Depending on the position of (53) in the support of ρ_1, one or other factor may dominate. In particular, if a large probability mass occurs in the region near $\rho_1 = 1$, the asymptoting behavior of $A^{-1/2}$ in the region near $\rho_1 = 1$ dominates the regular Student t form of $B^{-(n-1)/2}$, producing inferences which strongly favor a lack of cointegration even when the true value of ρ_1 falls well into the stationary region.

For this simple parameterization, approximation of the expectations in the Jeffreys prior as realized values produces

$$p(\beta|\rho_1) \propto A^{1/2}, \tag{58}$$

i.e. exactly the factor required to offset the distortion in the marginal ρ_1 density invoked by the factor $A^{-1/2}$.

In the case of the larger parameter set, Θ, it is the joint posterior of the vectors ρ and ψ which factorizes as in (53). Extending the notation used above, analytical integration yields

$$p(\rho, \psi|c, g) \propto |A|^{-1/2} B^{-(n-1)/2} \tag{59}$$

with

$$\begin{aligned}
A &= (g^{**\prime}g^{**})([1^*, g^*]'P_{g^{**}}[1^*, g^*]), \\
B &= (g^{**\prime}g^{**})(c^{*\prime}P_{g^{**}}P_{P_{g^{**}}[1^*,g^*]}P_{g^{**}}c^*)
\end{aligned} \tag{60}$$

and 1^*, c^*, g^* and g^{**} denoting respectively, the n-dimensional vectors for $1^*_t = (1 - \rho_1)$, $c^*_t = c_t - \rho' c^{\tau}_{t-1}$ $g^*_t = g_t - \rho' g^{\tau}_{t-1}$ and $g^{**}_t = \Psi(L)\Delta g_t$. Despite the fact that the two components in (59) do not reduce to simple parameter functions as in (56) and (57) above, the same qualitative result occurs. The first component possesses a singularity in and asymptoting behavior around the subspace described in (ii) in Section 3.1 in the text, whilst the second component possessing no such irregularities, but is dominated by the first factor when the bulk of the mass of the density is in the region of the singularity. The approximate Jeffreys prior for this full parameterization is given by

$$p(\alpha, \beta|\rho, \psi) \propto |A|^{1/2}. \tag{61}$$

Figure 1: Marginal ρ_1 posteriors for all G7 countries

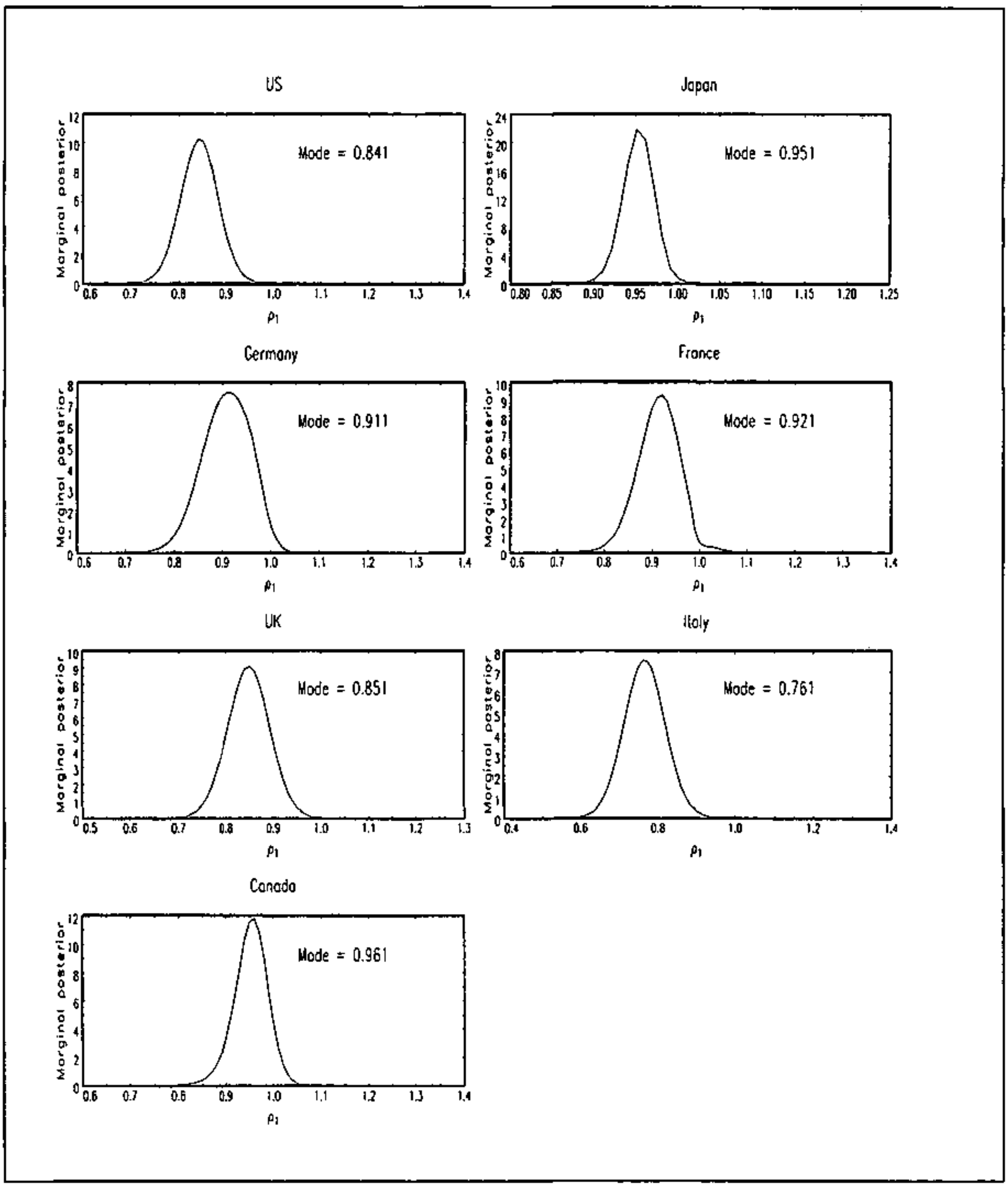

Figure 2: Marginal β posteriors for all G7 countries

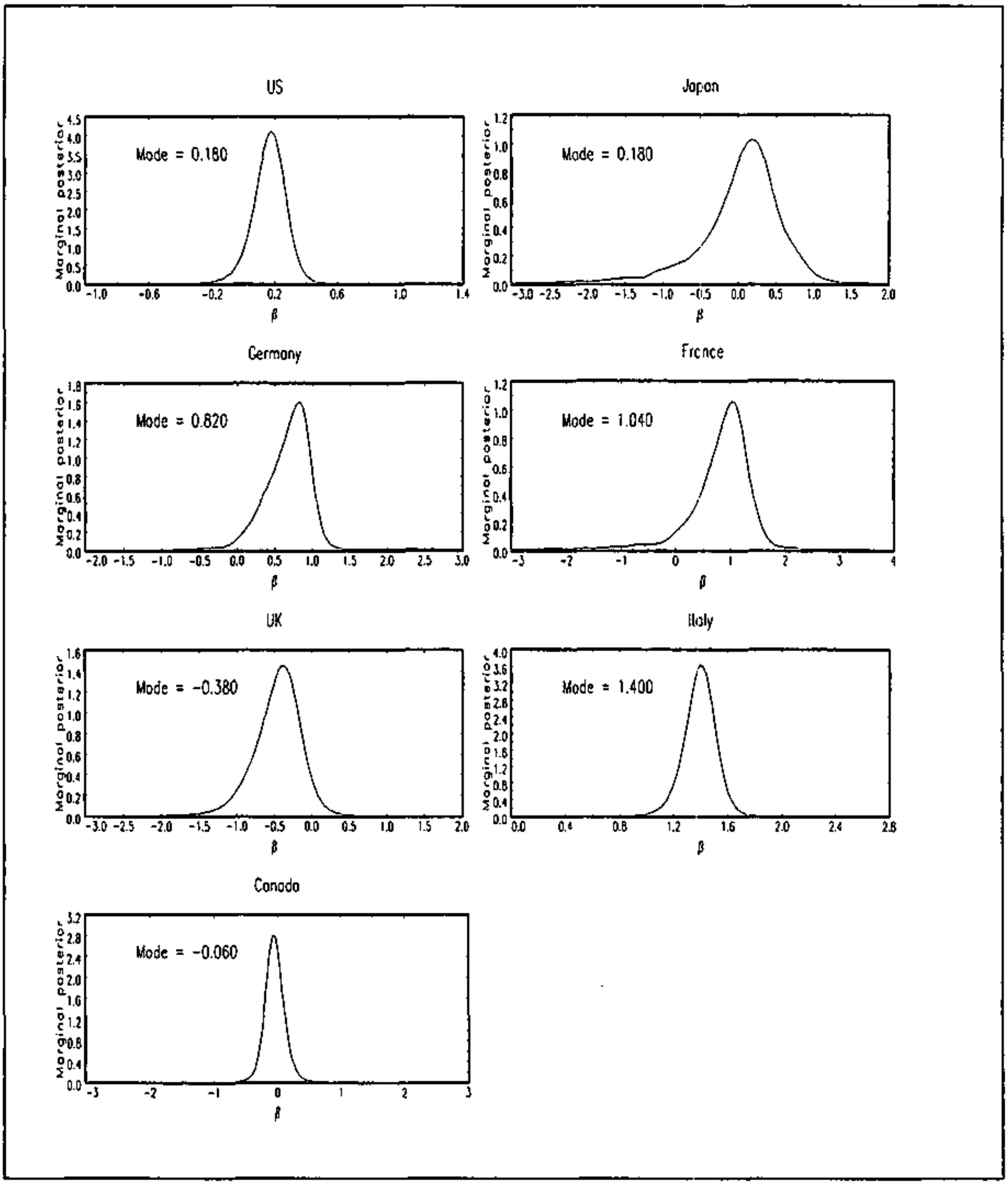

Figure 3: Impulse response functions (with 90% confidence bands given by dashed lines) for all G7 countries