EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Institute for the Protection and the Security of the Citizen
Econometrics and Statistical Support to Antifraud
I-21020 Ispra (VA) Italy

IPSC Exploratory Research 2003

Bayesian Cross-Country Estimation of Policy Assessment Models

Final Report

Prepared by:

Marco Ratto, Riccardo Girardi, Zoltan Palagyi
DG JRC - IPSC

Andries Brandsma, Constantin Ciupagea
DG JRC – IPTS

Werner Roeger
DG ECFIN

2004 **EUR 21194 EN**

LEGAL NOTICE

Neither the European Commission nor any person
acting on behalf of the Commission is responsible for
the use which might be made of the following information.

A great deal of additional information on the
European Union is available on the Internet.
It can be accessed through the Europa server
(http://europa.eu.int)

EUR 21194 EN
© European Communities, 2004
Reproduction is authorised provided the source is acknowledged
Printed in Italy

TABLE OF CONTENTS

1. Executive Summary

We explored and applied a Bayesian method for extending existing macro-economic models such as the Commission's QUEST II model. QUEST is a multi-country forward looking model, at present consisting of the EU Member States modules plus US, Japan and ten additional modules synthesizing the rest of the world. It is used for economic policy simulation. Accession of new countries requires an up-date of the present version of the model allowing the introduction of the modules of new countries.

Missing data and short time series over periods of structural change make it difficult to estimate additional country modules with the accuracy of existing modules. Bayesian methods allow considering a priori weights on model parameters on the basis of similarities with other countries, for which longer time series are available. Moreover, acceptability ranges can also be established based on theoretical and empirical considerations.

We applied the Bayesian approach to a prototypic EU accession country: Romania (Turkey, Bulgaria and Romania are still in the accession process, after the enlargement of May 1st 2004).

The estimation method has been implemented as a set of MATLAB routines. The core of the routines is given by the DYNARE software (Juillard, 1996, 2003), freely available and totally open source platform for the simulation and estimation of macro-economic models. Its development is lead by CEPREMAP (CNRS, Paris, F) and is carried out by a world-wide co-operation of researchers from different institutions, including JRC (M. Ratto). The DYNARE platform allowed a flexible handling and implementation of the specific estimation issues identified and developed within the present project.

The model specifications have been defined according to inputs from ECFIN (W. Roeger). Prior assumptions for Romania have been based on acceptability ranges other than on borrowing values from already estimate modules: Romanian economy revealed to be too far from actual EU25 countries.
IPTS contributed to the project in providing raw quarterly data starting in 1994. The data have been subsequently analysed and processed for seasonal adjustment, missing values, transformation of nominal series into real terms (constant prices).
Possible instability problems by adding the new module in QUEST II have been ruled out by tested specifications for current countries and by the priors for the new module.

The model specifications have been implemented into DYNARE and the investment module has been estimated applying the full Bayesian approach.
Other modules (labour demand, consumption, import, export price equations) are merely single equations, and have been estimated via a simple ordinary least squares (OLS) approach. Parameter values have been then checked ex post for consistency.

The analysis has been finalised through a dynamical simulation: impulse response functions for the investment module have been computed through a full Bayesian approach, allowing us to obtain a full range of dynamical responses, which are the core of policy simulation. This dynamic behaviour has been also tested ex-post for consistency.

The method proved to be effective and the computational tool based on DYNARE was flexible enough to allow the implementation of the specific issues concerning the present project. Estimation results for Romania are acceptable and consistent with theoretical macro-economic considerations.

2. Objectives of project

This research project aimed to identify estimation methods, which allow using a priori information, linked to observable similarities between countries, and assessing of the accuracy and consistency of the available estimates. More specifically, aims were

- To establish a formal Bayesian approach for extending existing multi-country models with additional country modules on the basis of time series of limited length, a priori similarities between countries and theoretical and empirical constraints. This is of particular relevance for the extension of the Commission's instruments of econometric analysis and policy evaluation.
- To implement in a user-friendly way the Bayesian approach: this includes attaching a priori weights on the basis of measurable similarities between countries, acceptability ranges based on theoretical considerations, a full Bayesian estimation computational tool, as well as a full Bayesian dynamical simulation engine for the final policy assessment. This also aims to derive a model that has been extended by the most objective and formalised methods, so that the results are transparent and reproducible.
- To apply the method to Accession Countries (Romania, Bulgaria, Turkey): Romania was selected as the prototypic country.

3. State of the art

Accession of new countries requires an update of the present version of the QUEST model allowing the introduction of the modules of new countries.

Traditional practice of macro-economic modelling would be based on the calibration of the parameters for each new country and the re-calibration of original modules. Recent development in Bayesian macro-economics has enhanced the possibility of estimating rather then calibrating the new modules. Most of the applications of Bayesian estimation of macro-economic models in the literature concern US or aggregated EU data.

As far as accession countries are concerned, missing data and short time series over periods of structural change make it difficult to estimate additional country modules

with the accuracy of existing modules. An accepted practice has been to exchange parameters between models and to borrow estimated coefficients of existing parts of the model, but this often raises questions of consistency and the possible distortion of the long-term properties of model projections. This requires the identification of appropriate tools.

4. Innovative aspects of work undertaken

The present project constitutes a first attempt to extend Bayesian macro-economic tools usually applied for big aggregated country models, such as US or EU, to an accession country.

For EU accession countries only short series are available and, for most of the observation period, far from equilibrium conditions. This makes standard maximum-likelihood estimation approach problematic. Bayesian estimation tools, adapted for the peculiar issue of accession countries, are particularly relevant since prior information and assumptions can play a major role, by constraining the estimation results into an economically acceptable region, ruling out instabilities in the model or allowing us to use estimation results of other countries with similar economies but longer observed series. Acceptability ranges can also be defined based on theoretical considerations. Moreover, estimation results can be checked ex-post for consistency by analysing the dynamical responses to shocks and verifying that the model behaves in accordance with macro-economic theory.

In this way, the analysis of dynamical responses is embedded in the whole Bayesian estimation procedure, which will provide, in the end, both the parameter estimate and the assessment of accuracy and consistency of the estimated model.

Finally, the implementation of the methodological approach into the totally free and open-source DYNARE software provides an interesting example, together with the contributions coming from other research institutions, of an effective convergence in a single computational tool of different methodological advances developed in the scientific community.

5. Approach and methodologies employed

The implementation of the research project required the following steps.

- Identification of the Bayesian approach for estimation and assessment for accuracy and consistency.
- Coding of the methodological approach into DYNARE.
- Specification of the model, according to ECFIN inputs. Each country module in QUEST is defined by various sectors: investment, labour supply, consumption, import, export and wages. Among these sectors, investment is the only relevant for the full Bayesian estimation procedure, and is characterised by a relatively complex system of equations. The remaining sectors are expressed by single

equations and a simple ordinary least squares estimation is performed for them. Parameter values are then checked ex-post for consistency.

- Pre-processing of the raw data collected.

5.1. Bayesian estimation

5.1.1. Estimation procedure

The Bayesian estimation approach at the core of the present project has been discussed by many authors in the literature in the last few years (see e.g. Schorfheide, 2000; Smets and Wouters, 2003; Lubik and Schorfheide, 2003).

Let a model be defined and first order conditions identified. This can be expressed as:

$$\begin{aligned} &E_t\{f(y_{t+1}, y_t, y_{t-1}, \varepsilon_t; \theta\} = 0 \\ &E\{\varepsilon_t\} = 0 \\ &E\{\varepsilon_t \varepsilon_t'\} = \Sigma \end{aligned} \tag{1}$$

where y is vector of endogenous variables, ε is the vector of exogenous stochastic shocks, θ is the vector of parameters and E is the expectation operator.

The non-linear model is solved via a linear approximation around the deterministic steady state $\bar{y}$ such that $f(\bar{y}, \bar{y}, \bar{y}, 0; \theta) = 0$. A linear rational expectation (LRE) system is obtained, with forward looking components

$$A^+ E_t \hat{y}_{t+1} + A^0 \hat{y}_t + A^- \hat{y}_{t-1} + B\varepsilon_t = 0, \text{ where } \hat{y}_t = y_t - \bar{y} \tag{2}$$

The system is solved for the reduced form state equation in its predetermined variables (Blanchard and Kahn, 1980; generalised Schur form, Klein, 2000). An observation equation is also added to link the observed variables y_t^* to the predetermined ones, obtaining:

$$\begin{aligned} &y_t^* = M(\bar{y}(\theta) + \hat{y}_t) + \eta_t \\ &\hat{y}_t = G(\theta)\hat{y}_{t-1} + H(\theta)\varepsilon_t \\ &E(\eta_t \eta_t') = V(\theta) \\ &E(\varepsilon_t \varepsilon_t') = Q(\theta) \end{aligned} \tag{3}$$

where η_t is the measurement error, if any. The system matrices G, H, V and Q and the steady state vector $\bar{y}(\theta)$ are functions of the vector of structural parameters θ of the original model. Vector θ includes the noise parameters Σ. The state space representation (3) allows use of Kalman filtering for the computation of the log-likelihood, and a subsequent inference based on it (maximum likelihood estimation, etc.).

In the original model specification (1), well-defined (and relatively few) shocks are usually present. If the number of shocks is smaller than the number of observed variables, singularities in the Kalman filter will be present, i.e. the probability distribution of the observables $p(Y^T \mid \theta)$ (the likelihood) can be degenerate. This implies the introduction of additional shocks until the system becomes non-singular, including either measurement errors η_t (as e.g. in Ireland, 2004, who also models the measurement error as a VAR(1) process) or additional structural shocks in the state

equation (as e.g. in Smets and Wouters, 2003). Rigorously, in such cases, as clearly stated by Schorfheide (2000), the evaluation approach here applied will "lead to an assessment of the modified model rather than the original one". In such cases, the parameter vector θ will be augmented for the additional noise terms and, if any, also for the VAR coefficients in the measurement errors as in Ireland (2004).

Likelihood-based inference presents a series of issues: specifically the lack of identification (global: multiple maxima; local: over-parameterisation, i.e. the maximum is given by a complex multidimensional combination/interaction structure rather then by a single point in the parameter space). This is even more critical for EU accession countries for which few data are available. Bayesian analysis is hence performed: prior distributions for model parameters have to be defined, representing the prior beliefs of the analyst on their plausible values, which, in combination with the likelihood function, allows us to obtain the posterior distribution.
Bayesian inference needs the use of stochastic simulations, specifically Markov Chain Monte Carlo (MCMC) techniques, allowing to obtain samples from the posterior joint pdf of the model parameters and subsequently to make an inference in which the parameter uncertainty and the shape of the likelihood are taken into account.
Finally, the dynamical properties of the model are analysed ex post, through an Impulse Response analysis, in order to check that the model respects some basic constraints about relationships between variables.
From the computational point of view, the linear approximation and the solution of the obtained LRE can be done automatically using the DYNARE program (Juillard, 1996, 2003), which applies the generalised Schur decomposition solution method. DYNARE is a software for the simulation of DSGE models, freely available and totally open source. Presently, an estimation module is being implemented on DYNARE by a world-wide co-operation of researchers comprising JRC (M. Ratto), to include the most recent developments in Bayesian estimation macro-economic models in an extremely efficient and easy way. This co-operative approach allows a considerable spare of coding and debugging time and a broader availability and applicability of advanced estimation techniques by the scientific community. DYNARE is also flexible, and allowed to easily incorporate the methodological issues concerning the present project.

5.2. *Model specification for private sector investment*

5.2.1. *Production function*

The output (GDP) is produced with a constant returns to scale Cobb Douglas production function:

$$Y_t = L_t^{P\alpha}(K_t)^{(1-\alpha)} \cdot TFP_t^{\alpha} \tag{4}$$

where *TFP* is an exogenous shock to technology, *L* is the labour supply, *K* is the capital stock and Y is output (GDP). The capital stock changes according to the rate of fixed capital formation J_t and the rate of geometric depreciation:

$$K_{t+1} = J_t + (1-\delta)K_t \tag{5}$$

To allow for a better description of the strong structural changes in a transition economy such as the accession country ones, a time variable depreciation rate is considered,

$\delta = \bar{\delta} + \delta_t$

where the deviation from the steady state level $\bar{\delta}$ follows a random walk process:

$\Delta\delta_t = e_{\delta,t}$

The level $\bar{\delta}$ is assumed equal to the depreciation rate used for current EU countries, i.e. $\bar{\delta}$ =0.01 on a quarterly basis.

Furthermore, in the actual calibrated version of QUEST, capital stock is usually initialized according to the empirical relationship $K_0 = cY_0$, with c~3 on an annual basis. For accession countries, a less stringent prior constraint is considered for this coefficient, with an acceptable range of [1, 5] on annual basis.

The estimation methodology applied is based on the linearization of the model, the solution of the obtained linear rational expectation model via a Schur decomposition, and the Bayesian estimation of the resulting backward looking state space model (see Section 5.1 below). The linearised state space representation of the original model also allows us to obtain smoothed estimates of the initial conditions of all state variables (either observed or unobserved). The acceptability of the estimate of K_0 can be then ex post checked with the empirical range [1, 5] (i.e. [4, 20] on a quarterly basis).

Both capital stock initialization and variable depreciation rate have an influence on the derivation of TFP, which will ultimately affect the estimation of investment dynamics.

5.2.2. *Investment equation*

Total investment expenditures are equal to investment purchases plus the cost of installation plus adjustment costs. The unit installation costs are assumed to depend on the investment to capital ratio and the relative change of investment with parameter ϕ_1, weighted by relative price of investment p_i^{IY} (ratio between investment and GDP deflators). Adjustment costs depend on the square of the relative change of investment with parameter ϕ_2. With this definition, neither installation costs nor adjustment costs affect the steady state, but they only have an effect in the dynamics. The investment specification assuming a stochastic trend for endogenous variables can be written as:

$$\frac{I_t}{J_t} = p_t^{IY}\left(1 + \frac{\phi_1}{2}\frac{J_t}{K_t}(J_t / J_{t-1} - \eta)^2\right) + \frac{\phi_2}{2}\left(\frac{J_t}{J_{t-1}} - \eta\right)^2 \qquad (6)$$

where η is the growth ratio in a stochastic trend definition of endogenous variables, i.e. the steady state value of the ratios J_t / J_{t-1}, Y_t / Y_{t-1}, K_t / K_{t-1}.

The corporate sector maximizes the net present value of its cash flow:

$$V_0 = E_t \sum_{t=0}^{\infty} \prod_{k=0}^{t} \left(1 + r_{0+k} + r_p\right)^{-1} \left\{(1 - t_c)[Y_t - w_t L_t] - I_t\right\} \qquad (7)$$

where r_t is the real interest rate, w_t are private sector wages, r_p is the risk premium and t_c is the corporate taxe rate.

5.2.3. Exogenous variables

TFP and *L*, together with the interest rate *r* and the relative price of investment p^{IY} are considered as exogenous in the investment estimation. After estimation of each single time series, the processes described below have been specified.

TFP:

$\Gamma_t = \eta_{TFP}\Gamma_{t-1}U_t$

$\ln(\Gamma_t) = \eta_{TFP} - 1 + \ln(\Gamma_{t-1}) + u_t$ where $u_t \sim N(0,\sigma_u)$

where η_{TFP} is the trend of technology.

Labour:

$L_t = \eta_{POP}L_{t-1}V_t$

$\ln(L_t) = \eta_{POP} - 1 + \ln(L_{t-1}) + v_t$ where $v_t \sim N(0,\sigma_v)$

where η_{POP} is the trend of employment (equal to the trend of population, i.e. employment is stationary in per capita terms). Hence $\eta = \eta_{POP}\eta_{TFP}$.

Relative price of investment

$p_t^{IY} = p_{t-1}^{IY} + \xi_t$ (RW) where $\xi_t \sim N(0,\sigma_p)$

Real interest rate

$r_t = (1-\rho_r)r + \rho_r r_{t-1} + \zeta_t$ (AR1)

where $\zeta_t \sim N(0,\sigma_r)$ and r is the steady state

5.2.4. The estimated model in per capita terms

The model equations can be re-written in per capita terms, normalising all variables with η_{POP}:

$$GY_t = (1-\alpha)GK_t + \alpha u_t + \alpha v_t + \alpha(\eta_{TFP} - 1) \tag{8}$$

$$\eta_{POP}(GK_{t+1} + 1) = JK_t + (1 - \bar{\delta} - \delta_t) \tag{9}$$

first order conditions from equation (7):

$$(1-t_c)(1-\alpha)YK_t + \frac{\phi_1}{2}p_t^{IY}[JK_t(DJ_t - \eta_{TFP})]^2 - \lambda_t + (1 - r_{t+1} - r_p - \bar{\delta} - \delta_{t+1})\lambda_{t+1} = 0 \tag{10}$$

$$\begin{aligned} &p_t^{IY} + \left(p_t^{IY}\phi_1 JK_t + \frac{\phi_2}{2}\right)(DJ_t - \eta_{TFP})^2 + \left(p_t^{IY}\phi_1 JK_t + \phi_2\right)DJ_t(DJ_t - \eta_{TFP}) \\ &- E\left\{(1 - r_{t+1} - r_p)\left(p_{t+1}^{IY}\phi_1 DJ_{t+1} + \phi_2\right)DJ_{t+1}^2(DJ_{t+1} - \eta_{TFP})\right\} = E\lambda_{t+1}(1 - r_{t+1} - r_p) \end{aligned} \tag{11}$$

exogenous processes

$$r_t = (1-\rho_r)r + \rho_r r_{t-1} + \zeta_t \tag{12}$$

$$p_t^{IY} = p_{t-1}^{IY} + \xi_t \tag{13}$$

$$\delta_t = \delta_{t-1} + e_{\delta,t} \tag{14}$$

identities

$$YK_t(GK_t + 1) = YK_{t-1}(GY_t + 1) \tag{15}$$

$$JK_{t-1}DJ_t = JK_t(GK_t + 1) \tag{16}$$

The equations (8-16) describe the dynamical evolution of the endogenous variables:

δ_t	time variable deviation from steady state depreciation rate
DJ_t	J_t / J_{t-1} ratio
GK_t	growth rate of capital
GY_t	growth rate of GDP (observed)
JK_t	J_t / K_t ratio
YK_t	Y_t / K_t ratio
λ_t	Lagrange multiplier (shadow price of capital)
r_t	real interest rate
p_t^{IY}	relative price of investment

with the following exogenous shocks

u_t	TFP shock
v_t	shock private sector employment
ζ_t	shock in r_t
ξ_t	shock in p_t^{IY}
$e_{\delta,t}$	shock in δ_t

For the estimation of the model, the following observed series are available

GY_t	Growth rate of GDP
DJ_t	Increment of investment of the private sector
p_t^{IY}	Relative price of investment
v_t	Shock to employment
r_t	Real interest rate

Measurement errors are also considered:

$e_{J,t}$: measurement error of DJ_t.

Taking the production function as fixed (α is computed as average of wage shares) and neglecting r_p (=0) and t_c (=0) at this stage, the list of parameters to be estimated for the investment equation is: ϕ_1, ϕ_2, η_{TFP}, i.e. 3 parameters plus the standard error of the shocks σ_δ, σ_u and σ_{DJ}. Due to the few data available, no joint estimation is performed for the other parameters for exogenous variables, i.e. ρ_r, σ_r, σ_p, σ_v. The latter are estimated separately from the observed time series. Moreover, η_{POP} is separately estimated using population data and all data are put in per capita terms prior to estimation.

The use of exogenous data for the estimation of the investment equation is particularly useful for this kind of data, since it has to be expected that observations do not describe an equilibrium economy. For example, there has been a very strong crisis in 1997 in Romania, reflected for example by real interest rates of –30% in the first quarter of 1997. Using exogenous variables in the model, we feed some information on the crisis to the model and we can check if this information is used correctly by the model to estimate a reasonable dynamics for investment.

5.3. *Other modules: labour, import, export, consumption*

5.3.1. *Labour demand equation*

Private employment is expressed as:

$$\log(L_t^P) = a\log(L_{t-1}^P) + (1-a)[\log Y_t^P - \log(w_t / p_t^Y)] + c$$

where

w_t wage rate per employee

p_t^Y GDP deflator

a and c are parameters to be estimated. Usual values for EU countries is $a \approx 0.9$, This labour demand equation is consistent with the Cobb-Douglas production function used in QUEST.

5.3.2. *Import equations*

The equation reads:

$$\log(IM_t) = \sum_{i=0}^{n} \theta_i PCM_{t-i} + \log(C_t + G_t + J_t^{TOT} + EX_t) + c$$

where

$PCM = \log(p_t^C / p_t^M)$

p_t^C consumer price deflator

p_t^M import price deflator

G_t government purchases in goods and services

EX_t exports

θ_i and c are the parameters to be estimated.

5.3.3. Export price equations

Export prices respond to domestic prices and/or to world trade prices:

$$\log(p_t^X) = (1-\lambda)\log(p_t^Y) + \lambda\log(WP_t E_t) + c$$

where

p_t^X export deflator

p_t^Y domestic price index GDP deflator based on private sector data

WP_t competitors price index (world price)

E_t currency conversion

λ and c to be estimated.

5.3.4. Consumption equation

A simplified approach is here followed, with respect to the QUEST II specification for (private sector) consumption.

Consumption should be a random walk and current income should not add to the prediction of next periods consumption when lagged consumption is added. If lagged wage and transfer income add to the explanation of consumption, then this is a sign that a fraction of households follow some simpler consumption rules (i.e. consumption proportional to income)

Specifically

$$\log C_t = c + a\log C_{t-1} + b\log(INCOME_{t-1})$$

where

$$INCOME_t = (1-t_t)w_t / p_t^Y L_t + TR_t / p_t^Y$$

t_t labour tax rate

L_t total labour (private + gov.)

TR_t transfer income

p_t^Y total GDP deflator

w_t total wage rate per employee (from private and government sectors)

a and b are the parameters to estimate.

5.4. Data requirements and pre-processing

From the model specifications defined above, a number of quarterly series starting from 1994 were required to carry out the estimation. Raw data collected presented a series of problems which had to be dealt with before carrying on the estimation:

- missing values;
- seasonality;
- inconsistency among series in real terms, i.e. the reference price level was different between different series; this required a complete re-imputation of the

series in real terms.

Time series pre-processing was implemented in MATLAB, with an additional module attached at the top of the DYNARE engine. This module embeds the TRAMO-SEATS programs for missing values, outlier analysis and seasonal adjustment (TRAMO-SEATS, 1994, 1996, MATLAB version). All the procedures needed for the estimation have been hence coded into a unique environment, allowed to obtain a fully automatic engine, from the raw data, to pre-processing, to estimation, to policy simulation. This allows us to embed all the hypotheses done in the data pre-processing in the Bayesian procedure, for a complete assessment of its consistency and robustness. Even if this aspect fell outside the objectives of the project, we managed to design the whole engine comprising the 'data-processing external loop'. However, resources did not allow us to perform a complete analysis to test the pre-processing procedure. The full exploration of such issues could be the subject of future research in co-operation with ECFIN.

6. Scientific results

6.1. Investment equation

The model has been estimated for a quarterly dataset of Romanian data, running from the first quarter of 1994 till the second quarter of 2003.
The shocks in the observed exogenous variables are shown in Figure 1. The 1997 crisis is very evident in the series. Given these shocks, we will check if the dynamics of investment is described reasonably by the model.

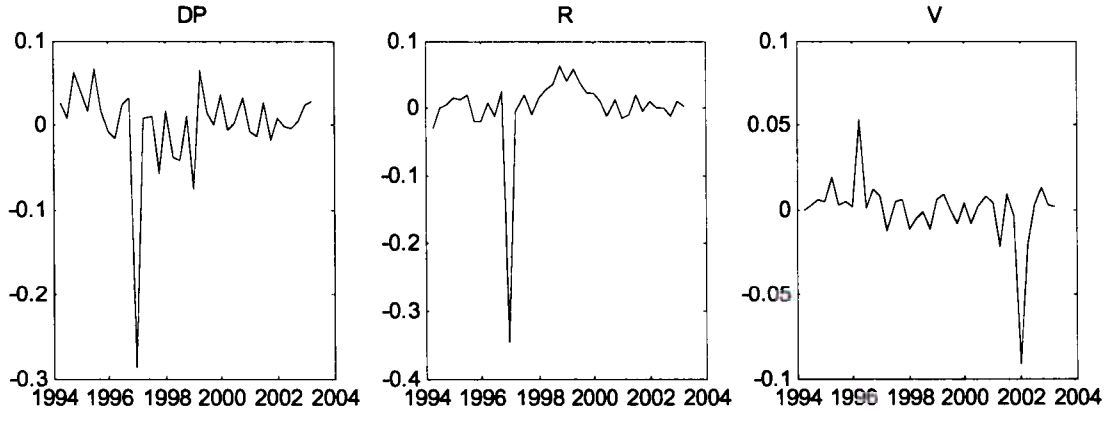


Figure 1: Exogenous observed variables

6.1.1. Setting up the estimation

Constant parameters (not estimated):

α = 0.5

$\bar{\delta}$ = 0.01
r_p = 0
t_c = 0
η_{POP} = 0.9995 (from population data)
ρ_r = 0.6013 (AR coefficient for interest rate)

Measured exogenous shocks (estimated separately from observed series):

σ_p = 0.0298
σ_r = 0.019
σ_v = 0.0091

Estimated parameters

Prior distributions for structural parameters have been set, by simply imposing plausible ranges. The possibility of using more informative priors based on values for actual EU countries was also planned. However, setting more informative priors is quite critical in this case, since we noted a notable sensitivity of posterior estimates to prior shapes. Moreover, Romanian economy was too far from any of the actual EU countries. Hence, we kept flat priors, based on theoretical constraints, on ruling out instabilities in the model and on empirical considerations, i.e. structural parameters cannot be too far from values used in any other country module:

ϕ_1 uniform prior [0 70]
ϕ_2 uniform prior [0 50]
η_{TFP} uniform prior [1, 1.01]

Prior distributions for shocks have been set to inverse gamma. For TFP and investment observation error, we assumed an infinite variance and a mean of the same order of the shocks of the exogenous variables. For variable depreciation rate, we imposed a much more constrained prior. This because we do not want depreciation rate wandering in a totally noisy manner, but, if any, we would like to estimate a smooth medium-long term history.

σ_δ inverse gamma prior, mean 0.0001, std. dev. 0.1
σ_u inverse gamma prior, mean 0.01, std. dev.= Inf
σ_{DJ} inverse gamma prior, mean 0.01, std. dev.= Inf

Steady state values of exogenous variables (from observed data):

p = 0.7695;
r = 0.038;

6.1.2. Bayesian estimation

The Bayesian estimation procedure starts with the maximisation of the posterior distribution, to estimate the mode. Results of the optimisation are shown in Table 1.

Table 1. Results from posterior maximisation.

parameters	prior mean	mode	s.d.	t-stat
PHI1	0.000	0.2946	18.4225	0.0160
PHI2	0.000	12.3482	1.1610	10.6356
ETA	1.000	1.0100	0.0050	203.6946
standard deviation of shocks	prior mean	mode	s.d.	t-stat
ED	0.000	0.0000	0.0000	3.1174
EU	0.010	0.0322	0.0038	8.3745
standard deviation of measurement errors	prior mean	mode	s.d.	t-stat
DJ	0.010	0.1219	0.0143	8.5399

In Figure 2 and Table 2 we show the posterior simulation of the estimated parameters with Metropolis Markov Chain Monte Carlo. Results are obtained after 120,000 runs of 4 parallel chains (480,000 runs in total). Posterior distributions are then analysed rejecting the first 65% of runs. Convergence was assessed based on the Multivariate Potential Scale Reduction Factor (MPSRF, Brooks and Gelman, 1998), as shown in Figure 3. This test aims at verifying that the samples obtained with the four parallel chains are drawn from the same distribution. The tuning of the jumping distribution has been quite critical. As usual, the jumping distribution between iteration (i-1) and i was a multivariate normal, with covariance matrix equal to the inverse of the Hessian at the posterior mode scaled by a factor c (usually c<1):

$$j(\theta^* \mid \theta^{(i-1)}) = N(\theta^{(i-1)} \mid c\Sigma_{\text{mod}\,e}),$$

where θ^* is the proposal value for iteration i.

Using the same scale factor for all parameters (c=0.9; acceptance rate of about 25%), yielded an extreme persistency of the chains for ϕ_2 and η_{TFP}, implying bad convergence tests. Then we allowed different scaling factors for the parameters, setting $c(\phi_2) = 5$ and $c(\eta_{TFP}) = 1.1$. The acceptance rate remained around 25%, indicating that the original jumping distribution did not allow a sufficient exploration of ϕ_2 and η_{TFP} values and that the Hessian at the posterior mode gave too narrow a standard error for these two parameters.

Comparing the shapes of prior and posterior distributions, we can see that the standard errors of shocks are shifted upwards by the likelihood while the shock to depreciation does not move so much from its narrow informative prior. As far as structural parameters are concerned, ϕ_1 has flat posterior and does not seem to be controlled by the data, while ϕ_2 has a much clearer peak in the lower part of its prior range (mode around 15) and η_{TFP} concentrates on the high values of its prior range.

Table 2. Posterior estimation results (Metropolis MCMC).

Log data density = 32.069309.

parameters

	prior mean	post. mean	conf. interval	
PHI1	0.000	35.0789	4.6723	67.5206
PHI2	0.000	19.6135	4.4028	38.8524
ETA	1.000	1.0062	1.0022	1.0100

standard deviation of shocks

	prior mean	post. mean	conf. interval	
ED	0.000	0.0001	0.0000	0.0002
EU	0.010	0.0341	0.0272	0.0405

standard deviation of measurement errors

	prior mean	post. mean	conf. interval	
DJ	0.010	0.1259	0.1018	0.1502

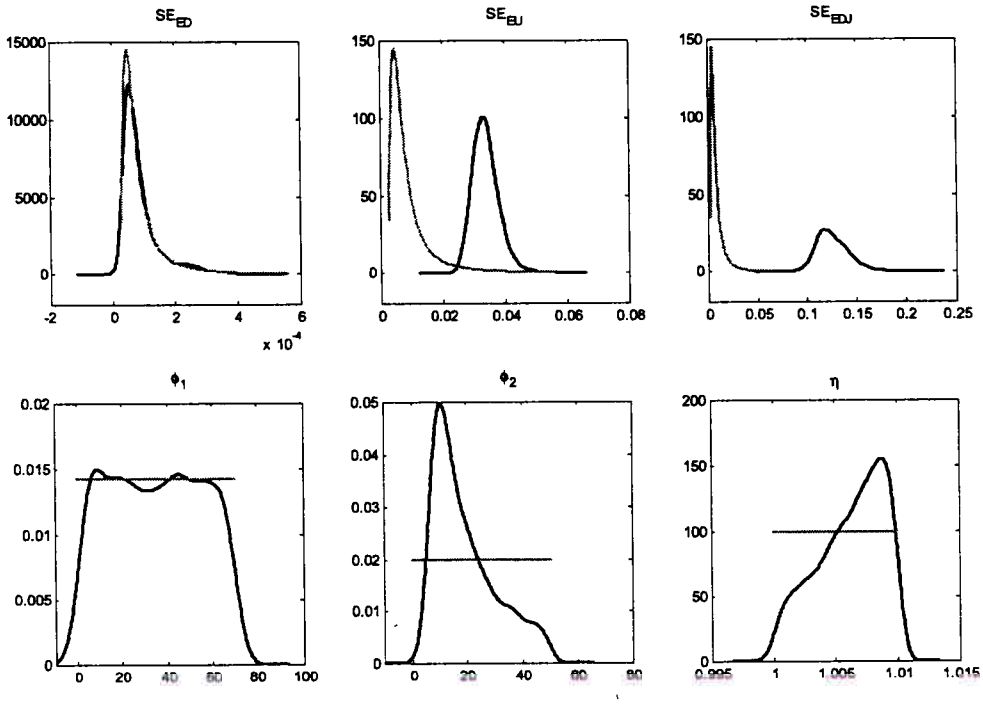


Figure 2. Posterior simulation for the estimated parameters with Metropolis algorithm. Gray lines are priors, black lines are posteriors.

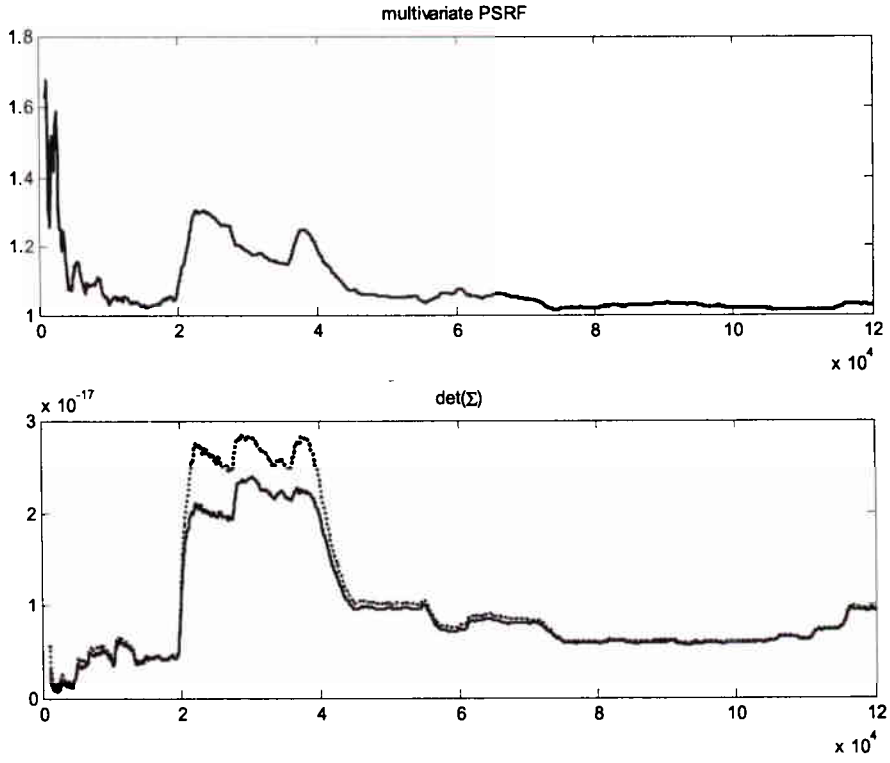


Figure 3. Convergence tests for Metropolis MCMC. Upper panel shows the multivariate potential scale reduction factor, which should be near to 1 at convergence. Lower panel shows the determinant of the 'between chains' and 'within chains' covariance matrices of the Monte Carlo sample.

6.1.3. Model behaviour

We present here the model behaviour using the values of the parameters at the posterior mean.

In Figure 4 we show the pure deterministic simulation of DJ for the whole estimation period, using the observed shocks in the exogenous variables (Figure 1) and setting to zero the measurement error of DJ. Some of the dynamics can be caught by this simple model. More specifically, the effect of the big shock in 1997 is reflected in the simulation. After that crisis, simulated DJ tends slowly to a steady state level. In the right panel of Figure 4 we show the reconstructed simulation in levels of *J*: the shock in 1997 is somehow smoothed as a medium term recession starting in 1994, but the simulated *J* pattern is fairly good, compared to observations.

In Figure 5 we show the 1-step-ahead prediction for DJ and GY. For DJ it does not differ significantly from the pure simulation, while for GY it gives a sort of smoothed trend. However, the fit is better appreciated in the lower panels of Figure 5, where the 1-step ahead predictions have been re-constructed in terms of levels of *J* and *Y*, applying the identities $J_{t|t-1} = J_{t-1} \cdot DJ_{t|t-1}$ and $Y_{t|t-1} = Y_{t-1} \cdot (GY_{t|t-1} + 1)$.

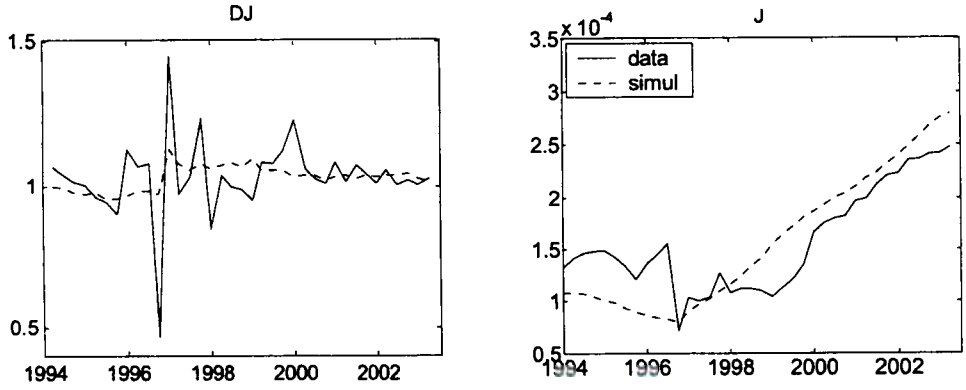


Figure 4.- Pure deterministic simulation of DJ at the posterior mean using the observed shocks for the exogenous variables and setting to zero the measurement error to DJ. Observations are dotted lines. The right panel shows the reconstructed simulation in *levels* for J.

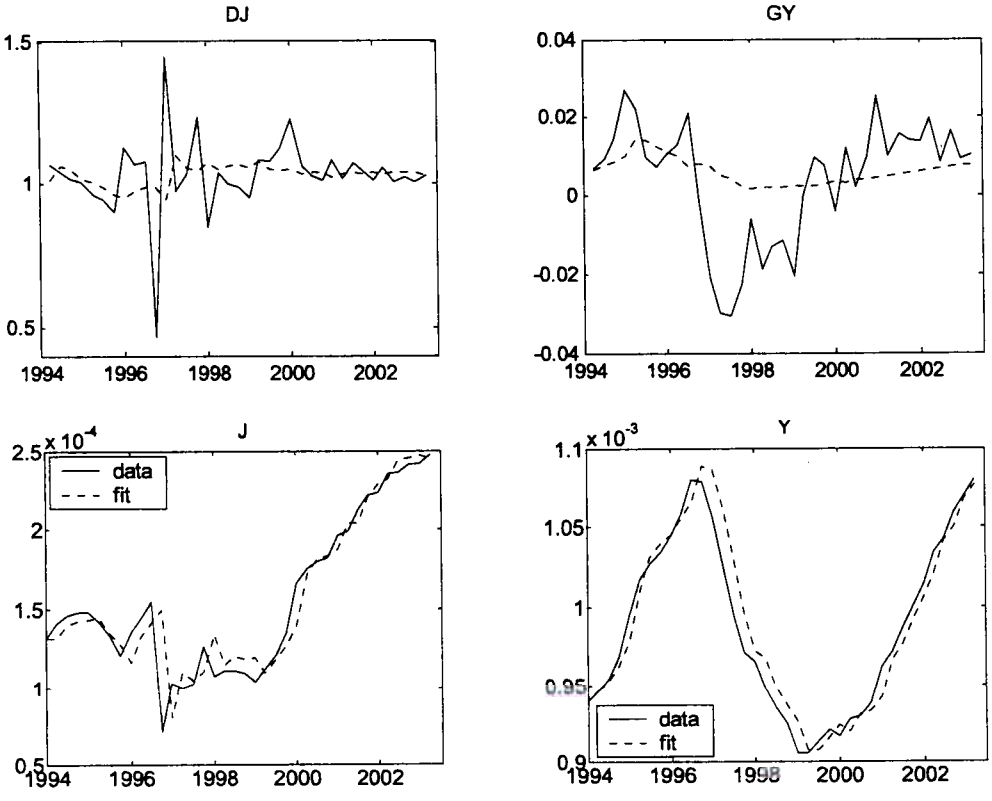


Figure 5. 1 step-ahead prediction (dotted) vs. actual observations. Lower panels show the reconstructed 1-step ahead prediction in levels for both *J* and *Y*.

A simple way to obtain better 1-step ahead predictions would be to introduce an auto-correlated measurement error, similarly to Ireland (2004). In Table 3 we show the posterior maximisation of this new version of the model. A quite high AR coefficient (RHOJ) is estimated for the measurement error while ϕ_2 is strongly reduced.

In Figures 6 and 7 we show the 1-step-ahead prediction and the pure deterministic simulation of the modified version of the model. While the 1-step ahead prediction of

DJ improves, no significant change can be noticed for GY. This improvement is rather 'artificial'. In fact, looking at the pure deterministic simulation in Figure 7, simulated DJ is strongly distorted and shifted to extremely low values at the beginning of the simulation, making the reconstructed *J* series totally wrong due to the deep fall in the first 10 periods. The patterns of Figure 4 have to be clearly preferred. In the latter case, the response of the model goes in the right direction, with a drop of DJ and a subsequent increase, followed by a slow recover to steady state levels. Moreover the reconstruction in terms of levels of *J* works correctly.

Table 3. Results from posterior maximisation of the modified model with auto-correlated measurement error of DJ.

parameters

	prior mean	mode	s.d.	t-stat
PHI1	0.000	0.1885	1.6550	0.1139
PHI2	0.000	1.9159	0.3316	5.7779
ETA	1.000	1.0077	0.0000	4063747.6368
RHOJ	0.950	0.9470	0.0221	42.7741

standard deviation of shocks

	prior mean	mode	s.d.	t-stat
ED	0.000	0.0000	0.0000	1873.4285
EU	0.010	0.0292	0.0034	8.5604
EJ	0.010	0.1408	0.0162	8.7124

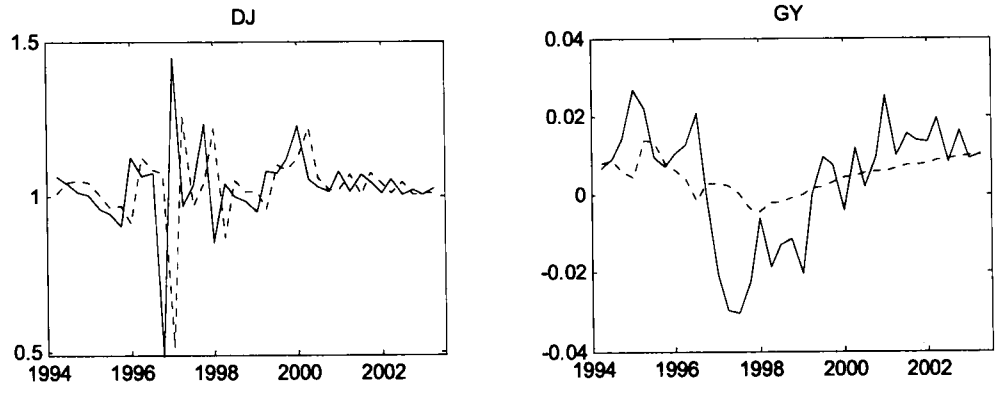


Figure 6. 1-step ahead prediction (dashed lines) compared to observations (continuous lines) for the modified model with auto-correlated measurement error of DJ.

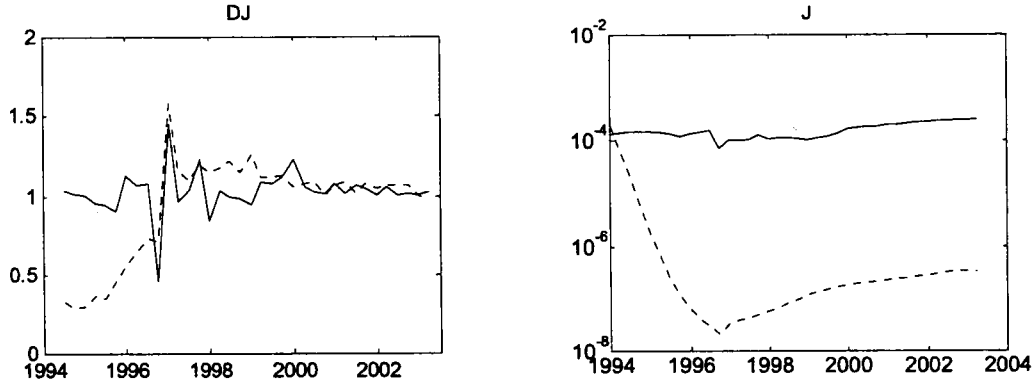


Figure 7. Pure deterministic simulation of the modified model with auto-correlated measurement error of DJ. The simulation is obtained using the observed shocks for the exogenous variables and setting to zero the measurement error to DJ. Dashed curve is model simulation; continuous line is the observed series. The right panel shows the reconstructed series in levels for *J*.

The state space approximation to the non-linear model also allows us to produce smoothed estimates of the unobserved variables (Figure 8). The crisis of 1997 is clearly displayed in the estimated series. For example, TFP growth rate has a strong negative shock; after that TFP growth tends to go back to a positive steady state level. It is also interesting to consider the estimated time variable deviation of from steady state of depreciation rate (DELT series in Figure 8). The smoother estimates a positive deviation of about 0.7% on a quarterly basis, making the effective depreciation rate increase from 0.01 to 0.017 during the estimation period. Further on, we can also check the smoothed estimate of the initial value of the K/Y ratio. The estimated value is around 3.4 on annual basis, which is perfectly in line with our prior acceptable range [1, 5].

Concerning steady state, we can see that the observed levels of GY and DJ (the sample means) are different from the theoretical values computed by the model. This can be due to different reasons:

1. the model parameterisation is too poor to allow a correct estimate of the steady state;
2. the current estimate of the steady state depends on the levels of observed interest rates, relative price of investment and labour; if the Romanian economy is still far from equilibrium, the level of observed time series might not be a good indicator of the equilibrium and thus alter a correct estimate of the steady state;
3. the computed steady state might be reasonable and the observed dynamics is still far from attaining a level comparable to the computed steady state.

Table 4. Steady state results at the posterior mean (sample means of observed seres are reposrted in parentheses).

Variable	Steady state	Observed sample mean
DJ	1.00618	(1.0283)
GK	0.00617886	(0.0039)
GY	0.00617886	-
JK	0.0156459	-
LAM	0.769482	-
YK	0.0738702	-

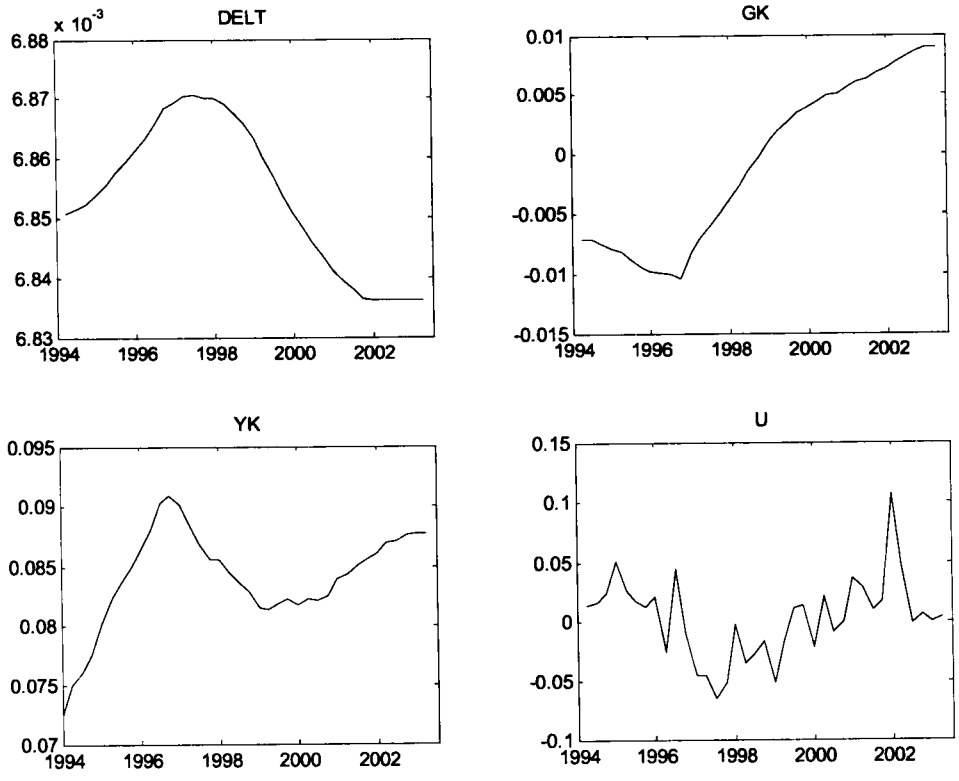


Figure 8: Smoothed estimates of deviation of depreciation rate from steady state (DELT), capital growth GK, YK ratio and TFP shock U.

6.1.4. Impulse response analysis

In Figures 9-13 we show the full Bayesian impulse response analysis of DJ, GY, GK, YK to shocks in the exogenous variables. All responses have a pattern complying theoretical issues, both in terms of mean path and error band. The linearised version of the model, already utilised for the estimation, has been applied. We selected randomly 1000 parameter sets out of the whole Monte Carlo sample and, for each parameter vector, IRF's were computed over a time span of T=50 periods imposing a unit shock in period t=1. This way we obtained a sample $f^{(j)}$, $j=1,...,1000$ from the posterior distribution of the IRF. One approach to represent the uncertainty in the posterior

sample of IRF's is to calculate the mean IRF ($\hat{f}$, dashed lines in the figures) and the one dimensional error band corresponding to the 95% confidence level (dotted lines in the figures). All such bands are asymmetrical with respect to the mean, and most of them present a sort of 'bottle-neck'.

Alternatively, as suggested by Sims and Zha (1998), one can represent each IRF, $f^{(j)}$, as a linear combination of the eigenvectors (v_i) of the T dimensional covariance matrix of the IRF's:

$$f^{(j)} = \hat{f} + \sum_{i=1}^{T} c_i^{(j)} v_i .$$

By construction, the mean of $c_i^{(j)}$ is zero for each i, and the T dimensional covariance matrix of the c_i's is diagonal, with the eigenvalues (λ_i) appearing as diagonal elements. In many applications (as in our case) most eigenvalues are negligible compared to the largest eigenvalue λ_M, thus most of the uncertainty in $f^{(j)}$ can be captured by the $c_M^{(j)} v_M$ term in the above expression. One can then calculate the 95% confidence interval $[c_M^{2.5}, c_M^{97.5}]$ of the $c_M^{(j)}$ coefficients, and plot the band $[\hat{f} + c_M^{2.5} v_M,\ \hat{f} + c_M^{97.5} v_M]$ corresponding to the main eigenvector (solid lines in the figures). It can be clearly seen that the latter band almost entirely covers the full 95% confidence band of the IRF's, except for 'bottle-neck' region. In this region, the main eigenvector highlights the presence of foldings, which could not be observed otherwise, adding valuable information to the IRF analysis.

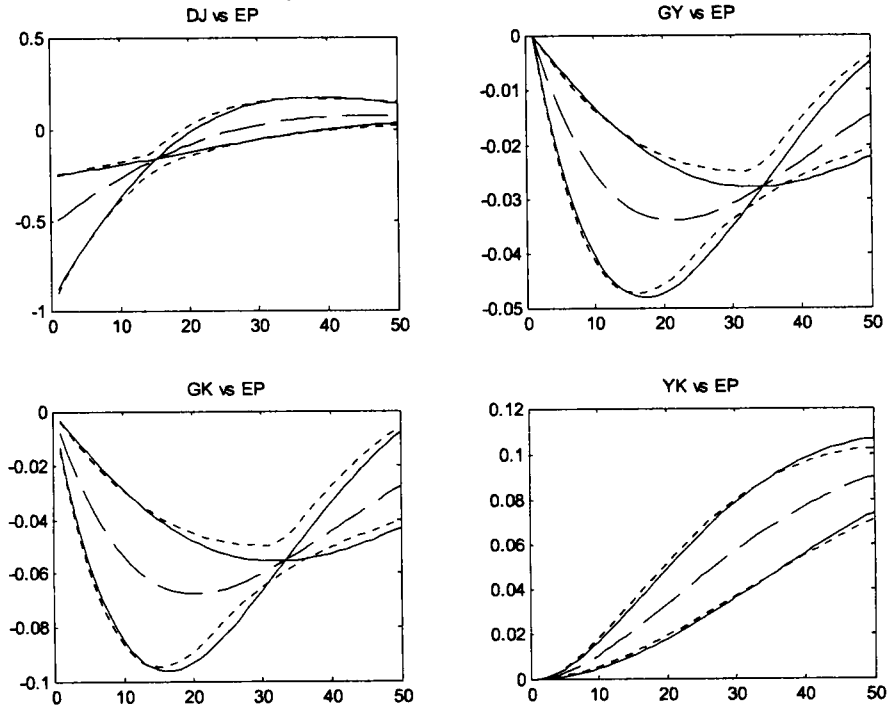


Figure 9: IRFs for a unit shock to relative price of investment (EP). Dashed line is the mean, dotted lines are the full 95% error band, solid lines are the 95% band for the eigenvector.

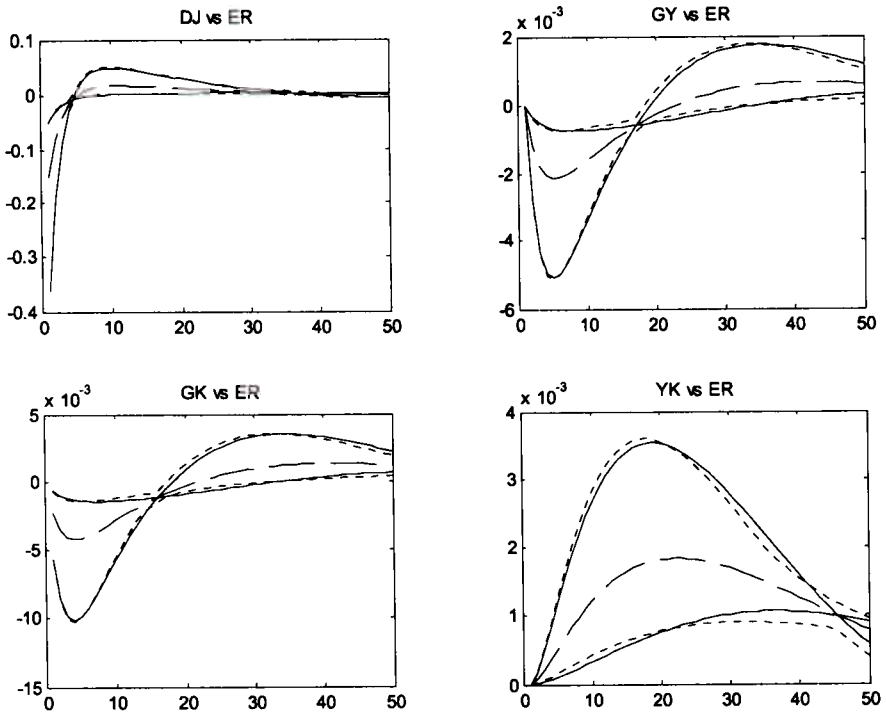


Figure 10: IRF's for a unit shock to interest rates (ER). Dashed line is the mean, dotted lines are the full 95% error band, solid lines are the 95% band for the eigenvector.

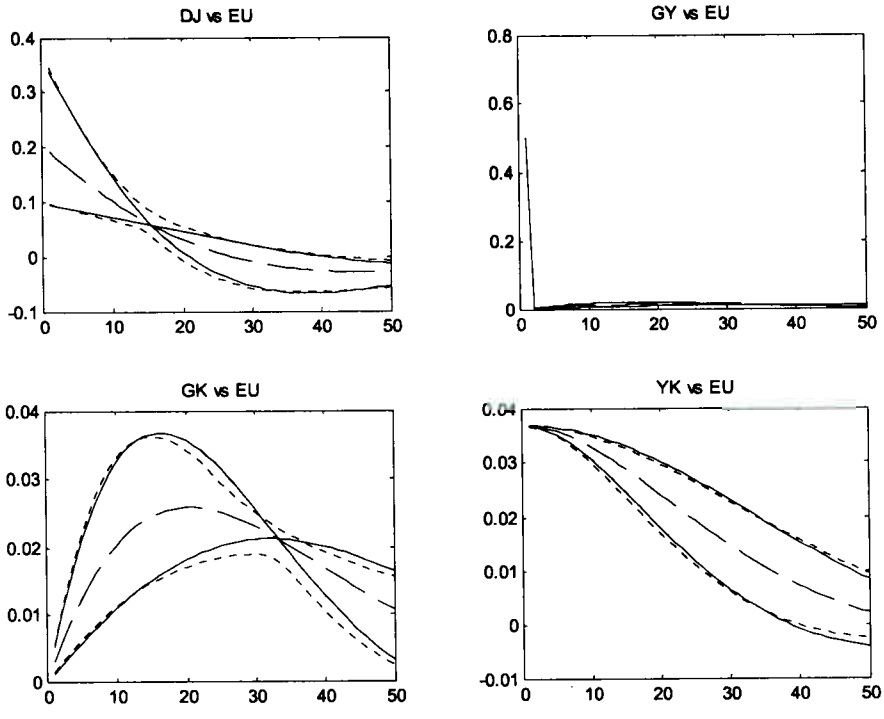


Figure 11. IRF's to a unit shock in TFP growth (EU). Dashed line is the mean, dotted lines are the full 95% error band, solid lines are the 95% band for the eigenvector.

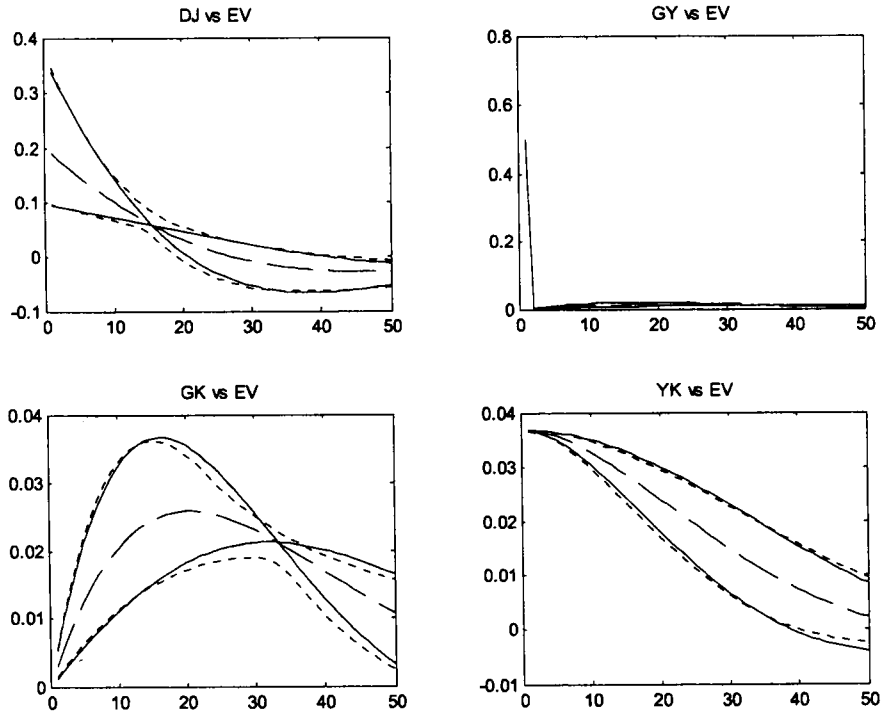


Figure 12: IRF's for a unit shock to growth rate of employment (EV). Dashed line is the mean, dotted lines are the full 95% error band, solid lines are the 95% band for the eigenvector.

In order to show the rapid decay of the spectrum of eigenvalues and the rapid covering of the error band by the few main eigenvectors, in Figure 13 we plotted the bands for the 2 largest eigenvalues for one type of IRF (DJ vs. EP). We can see that the band corresponding to the second eigenvector nicely accounts for the band-width in the 'bottle-neck' region, nicely complementing the band of the first eigenvector and so allowing for a complete description of the entire error band.

The Bayesian module for impulse response analysis has been entirely implemented within this project and attached to DYNARE.

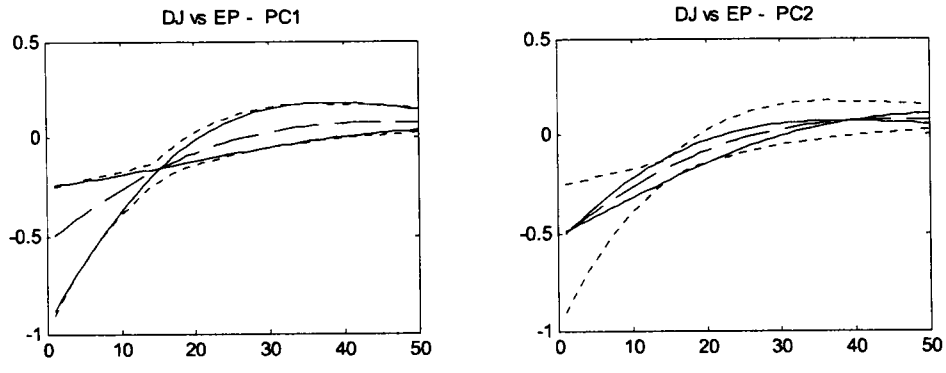


Figure 13. 95% error bands corresponding to the first two eigenvectors (principal components) for the IRF of DJ vs EP. Dashed line is the mean, dotted lines are the full 95% error band, solid lines are the 95% band for the principal component.

6.2. *Other modules: simplified estimation*

6.2.1. *Labour demand*

Private employment was expressed as (Section 5.3.1):

$$\log(L_t^P) = a\log(L_{t-1}^P) + (1-a)[\log Y_t^P - \log(w_t / p_t^Y)] + c$$

We estimated a and c by OLS after rearranging the equation as follows:

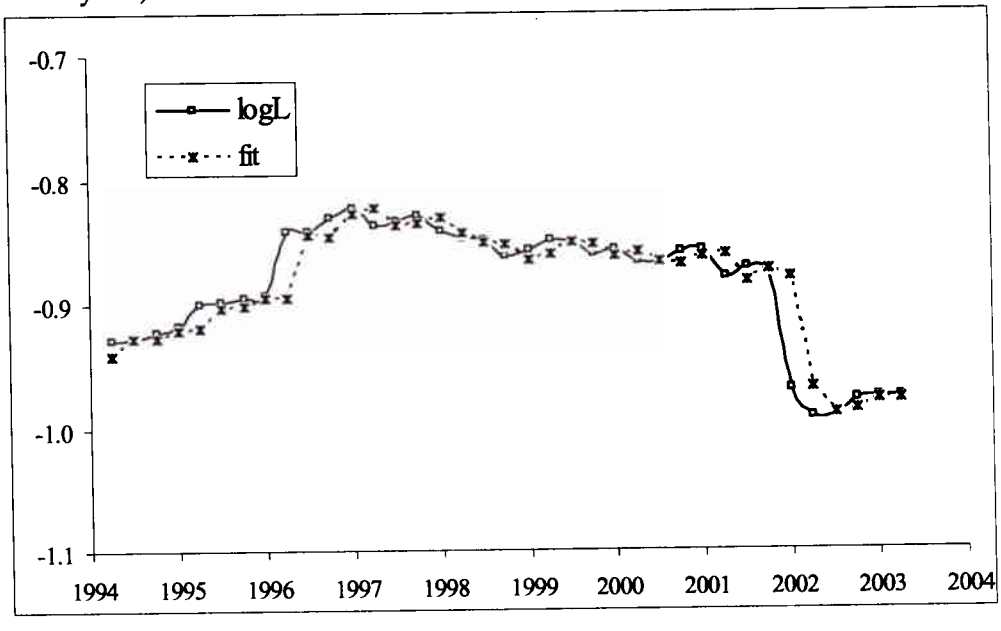
$$\log(L_t^P w_t / Y_t^P p_t^Y) = a\log(L_{t-1}^P w_t / Y_t^P p_t^Y) + c.$$

In the table below we report the results of the estimation:

parameter	estimated value	95% confidence interval
a	0.9690	[0.8716, 1.0665]
c	-0.0269	[-0.1094, 0.0555]

$R^2 = 0.921$ $F = 407.3$

The estimated a value is somewhat higher than the values from other countries (≈0.9). This seems to indicate a higher rigidity of the labour market in Romania. In the figure below we plotted $\log(L_t^P)$ ("logL") and the fit – the right hand side of the equation for $\log(L_t^P)$ evaluated with the estimated parameter values ("fit") as a function of time (quarters of a year).



We also estimated parameters in the extended version of the above labour equation, with distributed lags

$$\log(L_t^P) = \sum_{i=1}^{n} a_i \log(L_{t-i}^P) + (1 - \sum_{i=1}^{n} a_i)[\log Y_t^P - \log(w_t / p_t^Y)] + c$$

($n = 2,3$) and found that the effect of higher order lags L_{t-i}^P was insignificant.

6.2.2. Import

The import equation was (Section 5.3.2):

$$\log(IM_t) = \sum_{i=0}^{n} \theta_i PCM_{t-i} + \log(C_t + G_t + J_t^{TOT} + EX_t) + c$$

We estimated parameters θ_i and c by OLS after rearranging the equation as follows:

$$\log(IM_t / (C_t + G_t + J_t^{TOT} + EX_t)) = \sum_{i=0}^{n} \theta_i PCM_{t-i} + c.$$

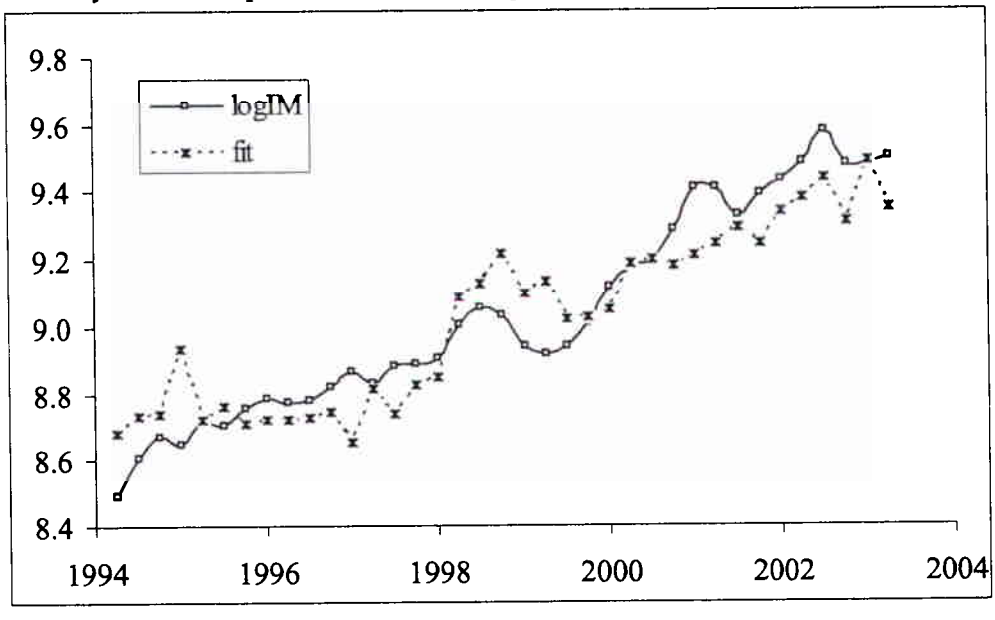
In the table below we report the results of the estimation for $n = 0$.

parameter	estimated value	95% confidence interval
θ_0	1.0473	[0.8584, 1.2363]
c	-1.3767	[-1.4191,-1.3343]

$R^2 = 0.778$ $F = 126.4$

For $n = 1$ confidence intervals for θ_0 and θ_1 become large, and $\theta_0 + \theta_1$ is approximately equal to θ_0 estimated for n=0, so the case of n=1 is over parameterised and the simplest expression with n=1 was the most appropriate.

In the figure below we plotted $\log(IM_t)$ ("logIM") and the fit – the right hand side of the equation for $\log(IM_t)$ evaluated with the estimated parameter values ("fit") as a function of time (quarters of a year). The increasing pattern is revealed by the simple equation, but cyclical components are not so good.



6.2.3. Export

Export prices were expressed as (Section 5.3.3):

$\log(p_t^X) = (1-\lambda)\log(p_t^Y) + \lambda\log(WP_tE_t) + c$

We estimated parameters λ and c by OLS after rearranging the equation as follows:

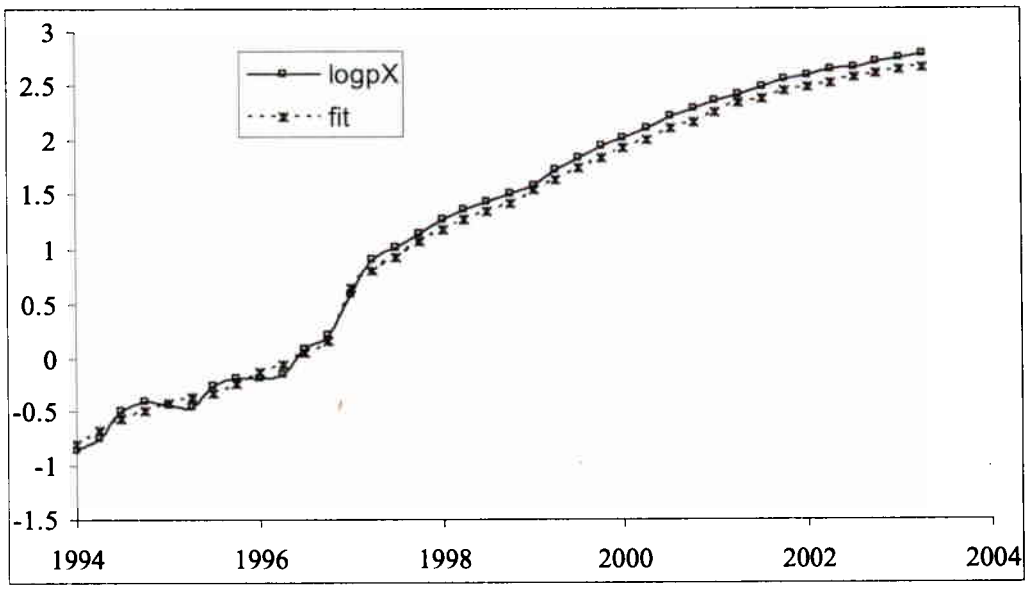
$\log(p_t^X / p_t^Y) = \lambda\log(WP_tE_t / p_t^Y) + c$

In the table below we report the results of the estimation:

parameter	estimated value	95% confidence interval
λ	0.0979	[-0.0420, 0.2378]
c	0.0071	[-0.0174, 0.0316]

$R^2 = 0.053 \quad F = 2.016$

In the figure below we plotted $\log(p_t^X)$ ("logpX") and the fit – the right hand side of the equation for $\log(p_t^X)$ evaluated with the estimated parameter values ("fit") as a function of time (quarters of a year). Although the R^2 value above is low, the fit is quite good, because the value of λ is low, thus most of $\log(p_t^X)$ is accounted for by the $(1-\lambda)\log(p_t^Y)$ term in the equation, indicating that export prices almost entirely depend on domestic prices.



6.2.4. Consumption equation

We could not complete the estimation of the consumption equation due to lacks in data of taxation and total income, that could not be solved in due time.

7. Comparison of results against objectives

The initial idea of using actual EU countries to put informative priors on the model parameters was not practicable. Romanian economy is too far from any of the actual EU countries. Using estimated models for developing countries could have been an alternative to what planned for the project, but this would have gone beyond the scopes

of the present project as well as require further unavailable resources to collect data and models from literature.
Prior assumptions have been then kept rather uninformative, by imposing a range of possible parameter values, based on both theoretical and empirical considerations. This did not impede a rather clear estimation of two out of the three parameters in the investment module.
On the other hand, data processing prior to estimation revealed to be a quite difficult job (see Section 5.4 above). This required a full data pre-processing routine to be attached at the top of the estimation procedure. The flexibility of DYNARE and the MATLAB environment also allowed us to design an 'external loop' of the whole estimation procedure, by which alternative hypotheses and time series approaches, useful to treat the data, could be tested to assess the robustness of the estimation to the data pre-processing. The full implementation and application of this external loop fell outside the scopes of the present project, and the resources available were not sufficient to complete this unexpected but interesting extension. We were able to design the numerical engine and we would be ready to apply it, if required by the institutional activities which will constitute the follow-up of this project.

8. Difficulties encountered and lessons learnt

Most of the difficulties encountered concerned the data set. When the project was proposed, neither the prototypic country was decided, nor any full data set was available for any of the accession countries. The availability of a full set of consistent time series of economic indicators revealed to be a problem. A lot of pre-processing was necessary, which took a significant part out of the resources allocated. This did not impede to complete satisfactorily the main core of the project: the coding of the Bayesian methodology and the estimation of the investment module, with a full Bayesian analysis of the dynamical responses for policy analysis. This main core had to be complemented by the estimation of single regressions for the labour, consumption, import and export modules. Among these, we could not complete the estimation of the consumption equation due to problems in data of taxation and total income, that could not be solved in due time. As a natural completion of the estimation work performed here, it would have been interesting to perform a full simulation of QUEST in a multi-country mode, including the new module for Romania. Due to the incomplete estimation of the consumption sector, the full simulation of QUEST had to be postponed.

9. Outlook and additional observations

The results of the present project are very likely to have a follow-up both on the customer DG's support side and on the purely scientific side. The methodological approach applied in this project, as well as the 'side-product' given by 'external loop' for assessment of consistency/robustness of data pre-processing, is suitable for a broad range of applications in framework of the upgrade of models used by the European Commission for policy simulation. The extension of the QUEST model for the ten new countries of the enlarged European Union is still in progress, with a clear relevance of

the methodological approach presented in this report. These extensions also have a very high scientific interest, due to the peculiarity of the countries analysed, with a short observation period characterised by strong structural changes. This poses a lot of methodological issues on the estimation procedures. So, not only the results obtained in the present project for the Romanian economy will be further disseminated in the open literature (in addition to the dissemination documented in the Appendix), but the present results are likely to produce a stream of scientific results in the framework of the extension of QUEST, to be disseminated in the open literature.

10. Conclusions

In the upcoming years the European Union will experience unprecedented political and socio-economic changes: ten countries entered the Union in May 2004 and thee more are negotiating accession for 2007. The new setting of the Union calls for a modification of the tools of socio-economic analysis of the European Commission.
One of the tools for economic policy simulation is the QUEST II model and accession of new countries requires an update of the present version of the model allowing the introduction of the modules of new countries. The present project constitutes a first attempt in this direction, applied to the Romanian economy. The methodological tools identified were appropriate to the goals of the present project and the MATLAB/DYNARE environment in which the methods have been implemented offered the flexibility necessary for the addition of the specific procedures required here. The results of the present project are very likely to have a follow-up both on the customer DG's support side (DG ECFIN) and on the scientific side, with publication of research articles on the application of the methodology to the extension of QUEST.

11. References

Blanchard O. J. and Kahn (1980), The solution of linear difference models under rational expectations, *Econometrica*, **48**, 5.

Brooks S.P., Gelman A. (1998), General methods for monitoring convergence of iterative simulations, *J. Computational and Graphical Statistics*, **7**, 434-455.

Collard F. and Juillard M. (2003), Stochastic simulations with DYNARE. A practical guide, Mimeo.

Ireland P.N. (2004) A method for taking model to the data, *J. Econ. Dynamics and Control*, **28**, 1205-1226.

Juillard M. (1996, 2003), DYNARE, A program for solving rational expectation models. Edition 2.6.1.1 for Dynare version 2.6.1, August 2003.

Klein P. (2000) Using the generalized Schur form to solve a multivariate linear rational expectations model, *Journal of Economic Dynamics & Control*, **24**, 1405-1423.

Lubik T. A., F. Schorfheide, Do central banks target exchange rates? A structural investigation, Working Paper, 2003, Mimeo.

Roeger W. and in't Veld Jan (1997) QUEST II A Multi Country Business Cycle and Growth Model, European Commission DG ECFIN Economic Papers, II/505/97-EN, Brussels.

Schorfheide, F. (2000), Loss function based evaluation of DSGE models, *Journal of Applied Econometrics*, **15**, 645-670.
Sims C. and Zha T. (1998), Error bands for impulse responses, Working Paper, Mimeo.
Smets F. and Wouters R. (2003), An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area, *Journal of the European Economic Association*, **1**, 1123-1175.
TRAMO-SEATS (TRAMO: Time Series Regression with ARIMA Noise Missing Observations and Outliers; SEATS: Signal Extraction in ARIMA Time Series), BETA VERSION (*), by V. Gomez and A. Maravall, with the programming assistance of G. Caporello. (*) Copyright : V. GOMEZ, A. MARAVALL (1994,1996).

Annex: List of publications, presentations given during exploratory research project

Workshop "Estimating stochastic dynamic general equilibrium (SDGE) models with Dynare", CEPREMAP (Paris), 9-14 February 2004. Invited Presentation of the estimation of the investment equation for Romania.

Joint JRC-ECFIN paper at the 10th International Conference on Computing in Economics and Finance (Amsterdam, July 8-10 2004): "Bayesian Estimation of Total Investment Expenditures For Romanian Economy using DYNARE", by M. Ratto, R. Girardi and W. Roeger.

ECFIN report on the estimation of model equations for the Romanian economy, by M. Ratto, R. Girardi, Z. Palagyi, W. Roeger, in preparation.

Mission of the JRC

The mission of the JRC is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of EU policies. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre