APPENDIX OF "TIME-VARYING INTERCEPTS AND EQUILIBRIUM ANALYSIS: AN EXTENSION OF THE DYNAMIC ALMOST IDEAL DEMAND MODEL"

Philippe J. Deschamps

Université de Fribourg

Séminaire d'économétrie, Avenue de Beauregard 13, CH-1700 Fribourg, Switzerland. Telephone: +41-26-300-8252. Telefax: +41-26-481-3845.

E-mail address: philippe.deschamps@unifr.ch

The algorithm for simulating (4.3) requires the knowledge of the full conditional posterior densities of γ, (β_0, β_1), Φ, Ω, μ, and V. Before describing the algorithm proper, we will derive the forms of these six densities.

Since only the first and third terms of the right-hand side of (4.3) depend on γ, we have:

$$p^*(\gamma \mid \beta_0, \beta_1, \Phi, \Omega, \mu, V) \propto \mathcal{L}(\beta_1, \gamma, \Phi, \Omega \mid \text{data}) f_{NO}[\beta_0; w_0 - (I_r \otimes \Pi)x_0, \Sigma]. \tag{A1}$$

The first term of this product is the likelihood corresponding to equation (3.4):

$$\Phi(L)w_t - \Phi(1)\alpha_t = [(\, x_t' \quad d_t' \,) \otimes \Phi(1)]S\,\gamma + \epsilon_t \quad \text{for } t = r+1, \dots, r+T. \tag{A2}$$

Similarly, the second term may be viewed as a likelihood for the first r observations, corresponding to the following regression equation:

$$\begin{aligned}(w_t - \alpha_t) &= (\, \Pi \quad B \,) \begin{pmatrix} x_t \\ O_{kq\times 1} \end{pmatrix} + u_t \\ &= [(\, x_t' \quad O_{1\times kq} \,) \otimes I_n] \operatorname{vec}(\, \Pi \quad B \,) + u_t \\ &= [(\, x_t' \quad O_{1\times kq} \,) \otimes I_n] S\,\gamma + u_t. \end{aligned} \tag{A3}$$

In this case, the nr disturbances u_{it} (for $t = 1, \dots, r$) are jointly distributed as $N(0, \Sigma)$ and are independent of the ϵ_{it}, for $t = r+1, \dots, r+T$.

Since (A2) and (A3) involve the same coefficient vector γ, we may combine them into:

$$y_* = X_* \gamma + \psi \tag{A4}$$

where $\psi \sim N(0, \Psi)$, with:

$$y_* = \begin{pmatrix} w_1 - \alpha_1 \\ \vdots \\ w_r - \alpha_r \\ \Phi(L)w_{r+1} - \Phi(1)\alpha_{r+1} \\ \vdots \\ \Phi(L)w_{r+T} - \Phi(1)\alpha_{r+T} \end{pmatrix} \tag{A5}$$

$$X_* = \begin{pmatrix} [\, x_1' \quad O_{1\times kq} \,] \otimes I_n \\ \vdots \\ [\, x_r' \quad O_{1\times kq} \,] \otimes I_n \\ [\, x_{r+1}' \quad d_{r+1}' \,] \otimes \Phi(1) \\ \vdots \\ [\, x_{r+T}' \quad d_{r+T}' \,] \otimes \Phi(1) \end{pmatrix} S \tag{A6}$$

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

$$\Psi = \begin{pmatrix} \Sigma & O_{nr\times nT} \\ O_{nT\times nr} & I_T \otimes \Omega \end{pmatrix}. \tag{A7}$$

The form of the conditional posterior of γ follows immediately, as:

$$p^*(\gamma \mid \beta_0, \beta_1, \Phi, \Omega, \mu, V) = f_{NO}\left[\gamma; (X_*'\Psi^{-1}X_*)^{-1}X_*'\Psi^{-1}y_*, (X_*'\Psi^{-1}X_*)^{-1}\right]. \tag{A8}$$

Next, we recall from (4.1) that:

$$\epsilon_t = \Phi(L)w_t - \Phi(1)[\alpha_t + z_t] \tag{A9}$$

with $z_t = \Pi x_t + Bd_t$, as defined in (4.2). This may be written as:

$$w_t - \alpha_t - z_t = \sum_{j=1}^{p} \Phi_j(w_{t-j} - \alpha_t - z_t) + \epsilon_t \quad \text{for } t = r+1, \ldots, r+T \tag{A10}$$

or, upon recalling the definition of Φ in (3.7), as:

$$Y = \Phi X + E \tag{A11}$$

where $\operatorname{vec} E \sim N(0, I_T \otimes \Omega)$, and where the $n \times T$ matrix Y and the $np \times T$ matrix X are defined as:

$$Y = \begin{pmatrix} w_{r+1} - z_{r+1} - \alpha_{r+1} & w_{r+2} - z_{r+2} - \alpha_{r+2} & \ldots & w_{r+T} - z_{r+T} - \alpha_{r+T} \end{pmatrix} \tag{A12}$$

$$X = \begin{pmatrix} w_{r-p+1} - z_{r+1} - \alpha_{r+1} & w_{r-p+2} - z_{r+2} - \alpha_{r+2} & \ldots & w_{r-p+T} - z_{r+T} - \alpha_{r+T} \\ \vdots & \vdots & \ldots & \vdots \\ w_r - z_{r+1} - \alpha_{r+1} & w_{r+1} - z_{r+2} - \alpha_{r+2} & \ldots & w_{r+T-1} - z_{r+T} - \alpha_{r+T} \end{pmatrix}. \tag{A13}$$

It follows from (4.3) and (A11) that:

$$p^*(\operatorname{vec}\Phi \mid \gamma, \Omega, \beta_0, \beta_1, \mu, V) \propto f_{NO}\left[\operatorname{vec}\Phi; \operatorname{vec} YX'(XX')^{-1}, (XX')^{-1} \otimes \Omega\right] I_S(\Phi) f(\Phi, \beta_0, \gamma, \Omega) \tag{A14}$$

$$p^*(\Omega \mid \Phi, \gamma, \beta_0, \beta_1, \mu, V) \propto f_{IW}[\Omega; T, (Y - \Phi X)(Y - \Phi X)'] f(\Phi, \beta_0, \gamma, \Omega) \tag{A15}$$

with:

$$f(\Phi, \beta_0, \gamma, \Omega) = (\det \Sigma)^{-1/2} \quad \times \\ \exp\left[-\frac{1}{2}[\beta_0 - w_0 + (I_r \otimes \Pi)x_0]'\Sigma^{-1}[\beta_0 - w_0 + (I_r \otimes \Pi)x_0]\right] (\det \Theta)^{-(n+1)/2}, \tag{A16}$$

and where Σ and Θ have been defined in Section 3.

Our derivation of the full conditional posterior of (β_0, β_1) is adapted from Min (1998). We write this density as:

$$p^*(\beta_0, \beta_1 \mid \gamma, \Phi, \Omega, \mu, V) = p^*(\beta_1 \mid \beta_0, \gamma, \Phi, \Omega, \mu, V) p^*(\beta_0 \mid \gamma, \Phi, \Omega, \mu, V). \tag{A17}$$

From (4.3), we see that:

$$\begin{aligned} p^*(\beta_0, \beta_1 \mid \gamma, \Phi, \Omega, \mu, V) \propto f_{NO}[\beta_0; w_0 - (I_r \otimes \Pi) x_0, \Sigma] \quad \times \\ \mathcal{L}(\beta_1, \gamma, \Phi, \Omega \mid \text{data}) f_{NO}[\beta_1; K^{-1}(\imath_T \otimes \mu + J\beta_0), K^{-1}(I_T \otimes V)(K')^{-1}]. \end{aligned} \tag{A18}$$

The information contained in the last two terms of (A18) may be summarized by:

$$W_1 = [I_T \otimes \Phi(1)]\beta_1 + \epsilon \tag{A19}$$

$$\beta_1 = K^{-1}(\imath_T \otimes \mu) + K^{-1} J \beta_0 + K^{-1} \eta \tag{A20}$$

where $\epsilon \sim N(0, I_T \otimes \Omega)$, $\eta \sim N(0, I_T \otimes V)$, and where:

$$W_1 = \begin{pmatrix} \Phi(L) w_{r+1} - \Phi(1) z_{r+1} \\ \vdots \\ \Phi(L) w_{r+T} - \Phi(1) z_{r+T} \end{pmatrix}. \tag{A21}$$

Combining (A19) and (A20) yields:

$$W_2 = X_0 \beta_0 + B_0 \eta + \epsilon \tag{A22}$$

where $B_0 = [I_T \otimes \Phi(1)] K^{-1}$, $W_2 = W_1 - B_0(\imath_T \otimes \mu)$, and $X_0 = B_0 J$. Since $I_T \otimes \Phi(1)$ is nonsingular, the marginal in (A17) is obtained by updating the prior on β_0 with the information contained in (A22), as follows:

$$p^*(\beta_0 \mid \gamma, \Phi, \Omega, \mu, V) = f_{NO}(\beta_0; M_0, V_0), \quad \text{where:}$$

$$V_0 = [\Sigma^{-1} + X_0' \{B_0 (I_T \otimes V) B_0' + I_T \otimes \Omega\}^{-1} X_0]^{-1}, \tag{A23}$$

$$M_0 = V_0 [\Sigma^{-1}\{w_0 - (I_r \otimes \Pi) x_0\} + X_0' \{B_0 (I_T \otimes V) B_0' + I_T \otimes \Omega\}^{-1} W_2]. \tag{A24}$$

The derivation of $p^*(\beta_1 \mid \beta_0, \gamma, \Phi, \Omega, \mu, V)$ is straightforward. We simply combine the information in (A19) with the prior (3.12), and obtain:

$$p^*(\beta_1 \mid \beta_0, \gamma, \Phi, \Omega, \mu, V) = f_{NO}(\beta_1; M_1, V_1), \quad \text{where:}$$

$$V_1 = [K'(I_T \otimes V^{-1}) K + I_T \otimes \Phi'(1) \Omega^{-1} \Phi(1)]^{-1}, \tag{A25}$$

$$M_1 = V_1 [K'(I_T \otimes V^{-1})(\imath_T \otimes \mu + J\beta_0) + \{I_T \otimes \Phi'(1)\Omega^{-1}\} W_1]. \tag{A26}$$

Finally, conditionally on β_0 and β_1, the second term on the right-hand side of (4.3) is the likelihood of (μ, V) corresponding to the last T instances of equation (2.6), which imply:

$$W_3 = \mu \, \imath_T' + H \tag{A27}$$

with $\operatorname{vec} H \sim N(0, I_T \otimes V)$ and with the $n \times T$ matrix W_3 defined as:

$$W_3 = [\, F(L)\alpha_{r+1} \quad \cdots \quad F(L)\alpha_{r+T} \,] \,. \tag{A28}$$

It follows that:

$$p^*(\mu \mid \Phi, \gamma, \Omega, \beta_0, \beta_1, V) = f_{NO}(\mu; T^{-1} W_3 \imath_T, T^{-1} V), \tag{A29}$$

$$p^*(V \mid \Phi, \gamma, \Omega, \beta_0, \beta_1, \mu) = f_{IW}[V; T + 2n + 1, (W_3 - \mu \imath_T')(W_3 - \mu \imath_T')' + \delta^2 I]. \tag{A30}$$

We now describe the strategy for simulating the joint posterior density (4.3). The author generated two groups of N independent Metropolis-Hastings chains, each one starting from an initial candidate generated from a thick-tailed, heuristically chosen density. The first group consisted of chains of length L_1 and the second of chains of length L_2, with $L_2 > L_1$. The last links of the $2N$ chains should be an i.i.d. sample from a good approximation to the posterior density. Convergence was checked by means of a Wald test of the equality of the expectation vectors of the two groups of generated variates, complemented by two-sample Kolmogorov-Smirnov tests on all the marginal univariate distributions, as implemented in the IMSL subroutine DKSTWO: the value of L_1 was judged satisfactory when the Wald test did not reject the null hypothesis at the 1 percent significance level, and none of the Kolmogorov-Smirnov tests rejected the null at the 1 per thousand significance level (these low values are chosen in order to control the overall level of significance).

The initial candidate hyperparameter vector, $(\Phi^0, \gamma^0, \Omega^0, \mu^0, V^0)$, was generated as follows. Φ^0, μ^0, and γ^0 were jointly drawn from a truncated multivariate Student density with three degrees of freedom, and location and scale parameters derived from the asymptotic distribution of the corresponding maximum likelihood estimates in (2.9), where the author neglected for simplicity the MA structure of the disturbances; the density was truncated on the invertibility region S. Ω^0 was drawn from an inverted Wishart approximate marginal posterior of the disturbance covariance matrix in (2.1), where γ_t was approximated by a linear trend. Finally, conditionally on Φ^0, γ^0, and Ω^0, V^0 was drawn from an inverted Wishart with $T + r - s + n$ degrees of freedom, and scale matrix given by:

$$\sum_{t=s+1}^{r+T} [F(L)\alpha_t - \hat{\mu}][F(L)\alpha_t - \hat{\mu}]'$$

where $\alpha_1, \ldots, \alpha_r$ were obtained by solving the deterministic part of (2.8), and $\alpha_{r+1}, \ldots, \alpha_{r+T}$ were drawn from multivariate Student densities with three degrees of freedom, and location and scale parameters derived from equation (2.5). The vector $\hat{\mu}$ is the time average of the $F(L)\alpha_t$.

Once these initial candidates have been generated, the algorithm proceeds in the usual fashion: we draw from the full conditional densities described at the beginning of this Appendix, always using the most recent draw of the conditioning variables. However, drawing β_0 and β_1 directly from (A17) is impractical, due to the typically very large dimensions of the matrices in (A23) to (A26). We therefore present an adaptation of the Kalman smoothing algorithm of Carter and Kohn (1994) to the present case.

We first express the model for the last T observations in state space form. Equation (2.6) may be written as:

$$\xi_t = \imath_r \otimes \mu + F\xi_{t-1} + v_t, \quad \text{for } t = r+1, \ldots, r+T \tag{A31}$$

where:

$$\xi_t = \begin{pmatrix} \alpha_{t-r+1} \\ \vdots \\ \alpha_{t-1} \\ \alpha_t \end{pmatrix} \tag{A32}$$

$$F = \begin{pmatrix} O & I & O & \cdots & O \\ O & O & I & \cdots & O \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ O & O & O & \cdots & I \\ F_r & F_{r-1} & F_{r-2} & \cdots & F_1 \end{pmatrix} \tag{A33}$$

with $F_j = O$ for $j > s$, and where $v_t \sim N(0, V_*)$, with:

$$V_* = \begin{pmatrix} O_{n(r-1)\times n(r-1)} & O_{n(r-1)\times n} \\ O_{n\times n(r-1)} & V \end{pmatrix}. \tag{A34}$$

In view of (4.2), equation (2.5) is equivalent to:

$$\Phi(L)w_t = H\xi_t + \Phi(1)z_t + \epsilon_t \quad \text{for } t = r+1, \ldots, r+T, \text{ with:}$$

$$H = [\, O_{n\times n(r-1)} \quad \Phi(1) \,]. \tag{A35}$$

Next, we recall from (3.13) that, conditionally on the first r sample observations, $\beta_0 = \xi_r$ is distributed as $N(w_0 - (I_r \otimes \Pi)x_0, \Sigma)$. The algorithm for generating $(\beta_0, \beta_1) = (\alpha_1, \ldots, \alpha_{r+T})$ from the conditional posterior is then a straightforward adaptation of Kim and Nelson (1999, chap. 8). An exposition is given here for the sake of completeness.

Step 1. Compute $\xi_{r|r} = w_0 - (I_r \otimes \Pi)x_0$, and $P_{r|r} = \Sigma$ from (3.14)–(3.16).

Step 2. For $t = r+1, \ldots, r+T$, compute $z_t = \Pi x_t + Bd_t$, and compute and save $\xi_{t|t}$ and $P_{t|t}$ from the following recurrence:

$$\xi_{t|t-1} = \imath_r \otimes \mu + F\xi_{t-1|t-1}$$

$$P_{t|t-1} = FP_{t-1|t-1}F' + V_*$$

$$\eta_{t|t-1} = \Phi(L)w_t - \Phi(1)z_t - H\xi_{t|t-1}$$

$$K_t = P_{t|t-1}H'(HP_{t|t-1}H' + \Omega)^{-1}$$

$$\xi_{t|t} = \xi_{t|t-1} + K_t\eta_{t|t-1}$$

$$P_{t|t} = P_{t|t-1} - K_tHP_{t|t-1}.$$

Step 3. Draw ξ_{r+T} from $f_{NO}(\xi_{r+T}; \xi_{r+T|r+T}, P_{r+T|r+T})$, and let α_{r+T} consist of the last n elements of ξ_{r+T}.

Step 4. Let F_* be the matrix consisting of the last n rows of F in (A33), and let α_{t+1} consist of the last n elements of ξ_{t+1}. For $t = r+T-1, \dots, r$, compute:

$$\xi_{t|t,\alpha_{t+1}} = \xi_{t|t} + P_{t|t}F_*'(F_*P_{t|t}F_*' + V)^{-1}(\alpha_{t+1} - \mu - F_*\xi_{t|t})$$

$$P_{t|t,\alpha_{t+1}} = P_{t|t} - P_{t|t}F_*'(F_*P_{t|t}F_*' + V)^{-1}F_*P_{t|t},$$

and draw ξ_t from $f_{NO}(\xi_t; \xi_{t|t,\alpha_{t+1}}, P_{t|t,\alpha_{t+1}})$. Let, for $t > r$, α_t consist of the last n elements of ξ_t, and let:

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{pmatrix} = \xi_r.$$

In the sequel, we will use the superscript (i) to denote the i-th link of the Markov chain, for $i = 1, \dots, L_1$ or $i = 1, \dots, L_2$. Once $\beta_0^{(i)}$ and $\beta_1^{(i)}$ have been generated by the Kalman smoothing algorithm just described, $\gamma^{(i)}$ is generated from the multinormal in (A8), with conditioning values given by:

$$(\beta_0^{(i)}, \beta_1^{(i)}, \Phi^{(i-1)}, \Omega^{(i-1)}, \mu^{(i-1)}, V^{(i-1)}).$$

Next, a candidate Φ is drawn by truncating on the invertibility domain S the multinormal in (A14), with conditioning values given by:

$$(\beta_0^{(i)}, \beta_1^{(i)}, \gamma^{(i)}, \Omega^{(i-1)}, \mu^{(i-1)}, V^{(i-1)}).$$

$\Phi^{(i)}$ is set equal to this candidate with probability:

$$p = \min\left[\frac{f(\Phi, \beta_0^{(i)}, \gamma^{(i)}, \Omega^{(i-1)})}{f(\Phi^{(i-1)}, \beta_0^{(i)}, \gamma^{(i)}, \Omega^{(i-1)})}, 1\right]$$

and is set equal to $\Phi^{(i-1)}$ with probability $1-p$. For a justification of this procedure, see Chib and Greenberg (1995) and Deschamps (2000).

Next, a candidate Ω is drawn from the inverted Wishart in (A15), with conditioning values given by:

$$(\beta_0^{(i)}, \beta_1^{(i)}, \gamma^{(i)}, \Phi^{(i)}, \mu^{(i-1)}, V^{(i-1)}).$$

$\Omega^{(i)}$ is set equal to this candidate with probability:

$$q = \min \left[\frac{f(\Phi^{(i)}, \beta_0^{(i)}, \gamma^{(i)}, \Omega)}{f(\Phi^{(i)}, \beta_0^{(i)}, \gamma^{(i)}, \Omega^{(i-1)})}, 1 \right]$$

and is set equal to $\Omega^{(i-1)}$ with probability $1 - q$.

Finally, $V^{(i)}$ is drawn from (A30), using the conditioning values:

$$(\beta_0^{(i)}, \beta_1^{(i)}, \gamma^{(i)}, \Phi^{(i)}, \Omega^{(i)}, \mu^{(i-1)})$$

and $\mu^{(i)}$ is drawn from (A29), using the conditioning values:

$$(\beta_0^{(i)}, \beta_1^{(i)}, \gamma^{(i)}, \Phi^{(i)}, \Omega^{(i)}, V^{(i)}).$$

The link index i is then updated to $i + 1$, and the next link is initiated by another run of the Kalman smoothing algorithm.

We conclude this Appendix with some practical remarks. The method just described was found to generate well behaved Markov chains with acceptable rejection rates (for the UK data with durables included, the average rejection rates on Φ and Ω were about 0.79 and 0.48, respectively; the corresponding figures for the US data were 0.13 and 0.33).

The author defined S as the set of matrices Φ for which the companion matrix $F(\Phi)$ in (3.15) has a spectral radius less than 0.99. The chosen lengths of the Markov chains varied between 500 and 800 links across models; 10000 replications were used in each case. For the quarterly UK model, where Φ has dimensions 5×20, it is cost-effective to generate Σ as a fixed point of:

$$\Sigma_{k+1} = F(\Phi)\Sigma_k[F(\Phi)]' + \Omega_*$$

where Ω_* is generated from Ω in a fashion similar to V_* in (A34). The differences with the inversion method were not significant.

The program was written in FORTRAN 77; the author used the IMSL random generator with a multiplier of 950706376 and shuffling. The code of the Kalman smoothing algorithm was tested against a direct simulation (using GAUSS) of (A17). Following a suggestion of John Geweke (2001), the full program was tested by replacing (3.17) and (3.18) by proper priors, and simulating the joint density $f(y, \theta)$ by generating y from the full conditional $f(y \mid \theta)$ (given by the likelihood) at each pass of the Metropolis-Hastings algorithm. The resulting replications of θ indeed reproduced the prior.

Additional References

Chib S, Greenberg E. 1995. Understanding the Metropolis-Hastings algorithm. *The American Statistician* **49**: 327–335.

Geweke J. 2001. Getting it right: checking for errors in Bayesian models and posterior simulators. University of Iowa: Mimeo.