Modeling oil consumption in Baumeister and Hamilton's (2019) model of the global oil market

Michał RUBASZEK, Karol SZAFRANEK

Financial Markets Modelling Unit
SGH Warsaw School of Economics

SGH
Warsaw School of Economics

Energy Finance Christmas Workshop
Wrocław, December 2025

The project was supported by the National Science Centre Poland, grant No. 2020/39/B/HS4/00366.

Crude oil market

- Crude oil represent almost 30% of global total energy supply.
- It is a global market.
- Large swings in oil prices draw a lot of attention among policymakers, academics and practitioners.
- Important questions arise:
    1. Are price changes driven by demand or supply shocks?
    2. What are the elasticities on the crude oil market?

Structural VAR models for the crude oil market

- Large literature on using SVAR models for the global crude oil market
  e.g. Kilian and Zhou, 2023
- Key parameter is the short-run price elasticity of oil supply, which determines the importance of oil demand and oil supply shocks for the real price of oil
  Kilian and Murphy, 2012; Herrera and Rangaraju, 2020
- Baumeister and Hamilton (2019) is a recent alternative approach to the workhorse model by Kilian and Murphy (2014).
- Recently, it has been extensively criticized by Kilian (2022a,b).

Debate on how to model crude oil market

LUTZ KILIAN

CHRISTIANE BAUMAISTER

JAMES HAMILTON

About this article

- We add to the debate on modelling global crude oil market with SVAR.
- We correct the demand equation in the BH approach and run the model.
- We use both the original and updated sample.
- We document how these changes affect key oil market elasticities.
- The price elasticity of oil supply drops markedly and is closer to zero.
- Demand shocks are the key driver of the real price of oil.

Baumeister and Hamilton (2019) model for the global crude oil market

The SVAR model

Structural VAR model:

$$\mathbf{A}\mathbf{y}_t = \mathbf{B}\mathbf{x}_{t-1} + \mathbf{u}_t,\, \mathbf{u}_t \sim N(\mathbf{0}, \mathbf{D})$$

$\mathbf{y}_t = (y_{1t}, y_{2t}, \ldots, y_{nt})'$ is an $n \times 1$ vector of endogenous variables

$\mathbf{A}$ is an $n \times n$ matrix describing contemporaneous structural relations

$\mathbf{x}_{t-1}^{'} = (\mathbf{y}_{t-1}^{'}, \ldots, \mathbf{y}_{t-m}^{'}, 1)'$ is a $k \times 1$ vector, with $k = mn + 1$

$\mathbf{B}$ is an $n \times k$ matrix of parameters at lagged variables

$\mathbf{u}_t$ is an $n \times 1$ vector of uncorrelated structural shocks

$\mathbf{D} = \mathrm{diag}(d_{11}, \ldots, d_{nn})$ is a diagonal matrix of size $n \times n$.

The specification of the SVAR

The structure of contemporaneous relations in terms of observables:

$$\Delta q_t = \alpha_{qp}\Delta p_t + \mathbf{b}_1^{'}\mathbf{x}_{t-1} + u_t^S$$
$$\Delta y_t = \alpha_{yp}\Delta p_t + \mathbf{b}_2^{'}\mathbf{x}_{t-1} + u_t^Y$$
$$\Delta q_t = \beta_{qy}\Delta y_t + \beta_{qp}\Delta p_t + \chi^{-1}\Delta i_t + \mathbf{b}_3^{'}\mathbf{x}_{t-1} + u_t^D - \chi^{-1}e_t$$
$$\chi^{-1}\Delta i_t = \psi_1\Delta q_t + \psi_3\Delta p_t + \mathbf{b}_4^{'}\mathbf{x}_{t-1} + u_t^I + \chi^{-1}e_t$$

where: $\Delta q_t = \log(Q_t/Q_{t-1}), \Delta y_t = \log(Y_t/Y_{t-1}), \Delta p_t = \log(P_t/P_{t-1})$ and $\Delta i_t = \Delta I_t/Q_{t-1}$.

Additive Gaussian measurement error for inventories:
$\Delta i_t = \chi\Delta i_t^* + e_t \iff \Delta i_t^* = \chi^{-1}\Delta i_t - \chi^{-1}e_t$

Accounting identity

- Following Kilian and Murphy (2014): $C_t \equiv Q_t - \Delta I_t^*$
- BH assumes that oil consumption equals to: $\Delta c_t \approx \Delta q_t - \Delta i_t^*$
- This is a mistake, as pointed out by Kilian (2022a,b).
- We derive that:

$$\Delta c_t = \frac{C_t - C_{t-1}}{C_{t-1}} = \frac{\Delta C_t}{C_{t-1}} = \frac{\Delta(Q_t - \Delta I_t^*)}{C_{t-1}} \approx \frac{\Delta(Q_t - \Delta I_t^*)}{Q_{t-1}} =$$

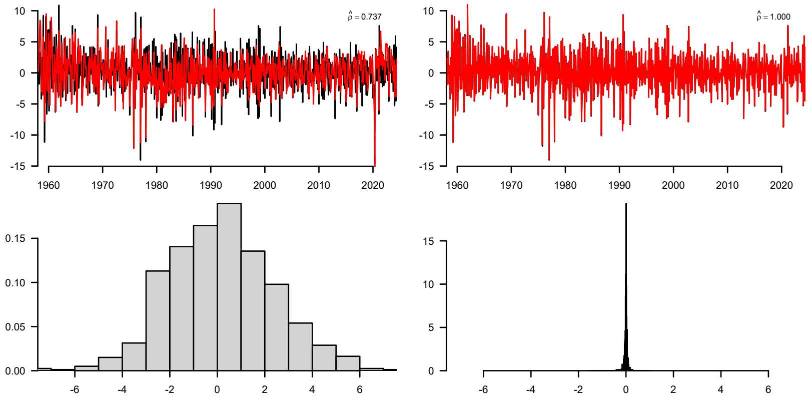
$$= \frac{\Delta Q_t}{Q_{t-1}} - \frac{\Delta^2 I_t^*}{Q_{t-1}} = \Delta q_t - \Delta^2 i_t^*.$$

While BH have used:

$$\Delta q_t - \Delta i_t^* = \frac{\Delta Q_t}{Q_{t-1}} - \frac{\Delta I_t^*}{Q_{t-1}} = \frac{Q_t - Q_{t-1} - \Delta I_t^*}{Q_{t-1}} =$$

$$= \frac{C_t - Q_{t-1}}{Q_{t-1}} \approx \frac{C_t - C_{t-1}}{C_{t-1}} = \Delta c_t$$

The approximation error



Notes: The figure presents the growth rate in oil consumption (Δc_t^*, black line) across the updated sample, where consumption is approximated as $C_t^* = Q_t - \chi^{-1}\Delta I_t$, and $\chi = 0.603$, as in BH. The series is compared to the BH proxy, $\Delta q_t - \chi^{-1}\Delta i_t$ (upper left corner, red line), and our proxy, $\Delta q_t - \chi^{-1}\Delta^2 i_t$ (upper right corner, red line), with $\hat{\rho}$ denoting the Pearson correlation coefficient. The distribution of differences between the series is presented in the bottom row.

Correcting the model

To correct this error, we substitute Δi_t in model with $\Delta^2 i_t$. New equations are:

$$\Delta q_t = \alpha_{qp}\Delta p_t + \mathbf{b}_1^{'}\mathbf{x}_{t-1} + u_t^S$$

$$\Delta y_t = \alpha_{yp}\Delta p_t + \mathbf{b}_2^{'}\mathbf{x}_{t-1} + u_t^Y$$

$$\Delta q_t = \beta_{qy}\Delta y_t + \beta_{qp}\Delta p_t + \chi^{-1}\Delta^2 i_t + \mathbf{b}_3^{'}\mathbf{x}_{t-1} + u_t^D - \chi^{-1}e_t$$

$$\chi^{-1}\Delta^2 i_t = \psi_1\Delta q_t + \psi_3\Delta p_t + \mathbf{b}_4^{'}\mathbf{x}_{t-1} + u_t^I + \chi^{-1}e_t$$

- Re-estimate the SVAR keeping all remaining settings unchanged.
- Use two samples: 01.1958-06.2024 and 01.1958-12.2016.

The prior

Prior for matrix **A**:

$$\alpha_{qp} \sim t_3^+(0.1, 0.2)$$
$$\alpha_{yp} \sim t_3^-(-0.05, 0.10)$$
$$\beta_{qp} \sim t_3^-(-0.1, 0.2)$$
$$\beta_{qy} \sim t_3^+(0.7, 0.2)$$
$$\psi_1 \sim t_3(0, 0.5)$$
$$\psi_3 \sim t_3(0, 0.5)$$
$$\chi \sim Beta(0.6, 0.009)$$
$$h_1 \sim At_3(0.6, 1.6, 2)$$
$$h_2 \sim t_3(0.8, 0.2)$$

Prior for other parameters:

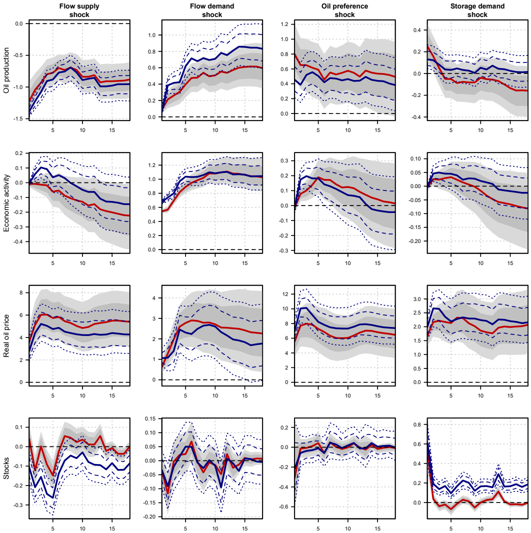
- τ_i and κ_i set in line with the standard Bayesian VAR literature.
- standard hyperparameters ($\lambda_0 = 0.5$, $\lambda_1 = 1$, $\lambda_3 = 100$).
- Bayesian estimation: M^* draws after M burn-in ($M = M^* = 1e6$).

Prior and posterior distribution

		Reported statistic	α_{qp}	α_{yp}	β_{yp}	β_{qp}	χ	ψ_1	ψ_3	ρ	h_1	h_2
Prior	All models	Type	t^+	t^-	t^+	t^-	*Beta*	*t*	*t*	*Beta*	*At*	*t*
		Location	0.100	-0.050	0.700	-0.100	0.600	0.000	0.000	0.25*χ	0.600	0.800
		Scale	0.200	0.100	0.200	0.200	0.009	0.500	0.500	0.12*χ	1.600	0.200
		D.o.f.	3.000	3.000	3.000	3.000	3.000	3.000	3.000	3.000	3.000	3.000
		Skew	—	—	—	—	—	—	—	—	2.000	—
Posterior	BH16	5%	0.072	-0.008	0.434	-0.622	0.438	-0.411	-0.090	0.061	0.491	0.988
		50%	**0.145**	**-0.002**	**0.724**	**-0.355**	**0.603**	**-0.138**	**-0.035**	**0.145**	**0.631**	**0.997**
		95%	0.303	0.000	1.092	-0.181	0.754	0.032	0.031	0.241	0.880	1.000
	RS16	5%	0.064	-0.008	0.429	-0.627	0.437	-0.446	-0.120	0.067	0.474	0.986
		50%	**0.132**	**-0.002**	**0.728**	**-0.353**	**0.604**	**-0.091**	**-0.047**	**0.161**	**0.629**	**0.997**
		95%	0.303	0.000	1.122	-0.166	0.756	0.092	0.029	0.269	0.925	1.000
	BH24	5%	0.032	-0.007	0.596	-0.715	0.446	-0.193	-0.071	0.060	0.460	0.987
		50%	**0.071**	**-0.002**	**0.876**	**-0.450**	**0.608**	**-0.047**	**-0.023**	**0.131**	**0.580**	**0.998**
		95%	0.126	0.000	1.421	-0.295	0.756	0.063	0.030	0.214	0.773	1.000
	RS24	5%	0.029	-0.007	0.595	-0.670	0.439	-0.152	-0.088	0.068	0.444	0.987
		50%	**0.065**	**-0.002**	**0.871**	**-0.435**	**0.604**	**0.011**	**-0.032**	**0.148**	**0.569**	**0.997**
		95%	0.119	0.000	1.376	-0.280	0.753	0.123	0.025	0.237	0.747	1.000

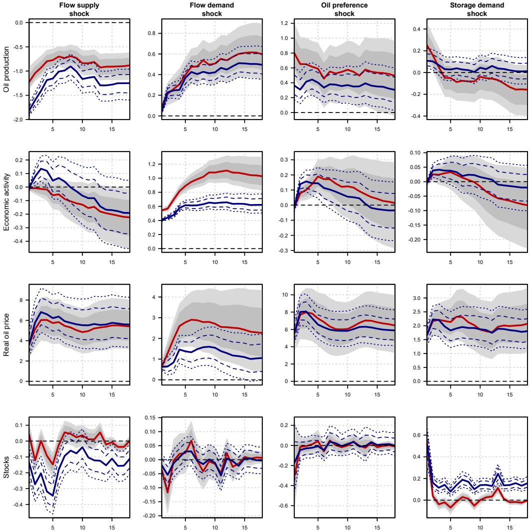
Notes: In the table *t* denotes a Student t distribution and *At* denotes an asymmetric Student t distribution. Signs + and − indicate that the distribution is truncated to be either positive or negative, respectively. D.o.f stands for degrees of freedom. BH16 and BH24 denote results for the original BH model estimated with data ending in December 2016 and in June 2024, respectively. RS16 and RS24 show the results after substituting equations 3 and 4 of the model on respective samples.

Impulse response functions: BH16 (red) vs RS24 (blue)



Note: The areas denotes the 68 and 95 percent posterior credible sets.

IRF – with similar response to oil prices: BH16 (red) vs RS24 (blue)



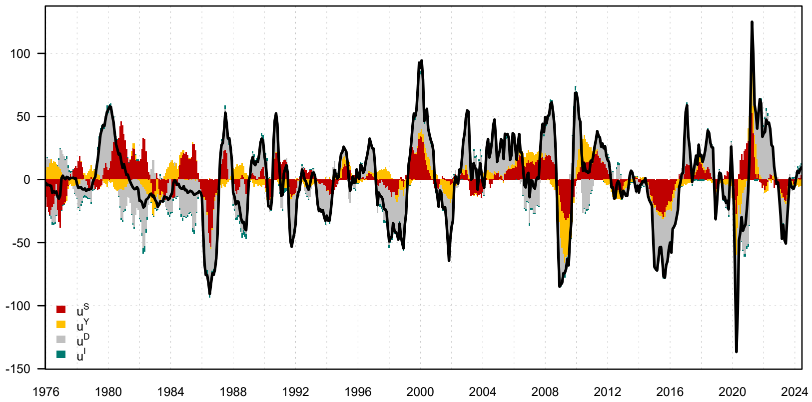
Notes: The areas denotes the 68 and 95 percent posterior credible sets. For BH16 responses to one standard deviation shocks. For RS24 the responses are normalized so that the initial response of the real price of oil to the shocks is the same as in BH16.

Forecast Error Variance Decomposition

	u^S	u^Y	u^D	u^I	u^S	u^Y	u^D	u^I
	BH16				RS16			
h=1	27.3	0.8	65.2	5.8	23.4	0.8	67.6	7.2
	[10.0-63.4]	[0.2-1.9]	[30.3-85.4]	[3.1-8.4]	[7.7-64.1]	[0.2-1.9]	[28.2-87.2]	[3.6-10.9]
h=6	27.9	3.4	62.4	5.4	24.3	3.4	64.6	6.7
	[11.0-61.5]	[1.5-6.4]	[29.9-82.2]	[3.0-7.9]	[8.9-62.2]	[1.5-6.4]	[27.7-84.1]	[3.4-10.0]
h=12	28.0	3.8	61.8	5.5	24.6	3.8	63.9	6.6
	[11.4-60.6]	[1.8-6.8]	[30.1-81.4]	[3.1-8.0]	[9.4-61.6]	[1.8-6.9]	[27.9-83.2]	[3.4-9.9]
h=18	27.9	3.9	61.7	5.6	24.6	3.9	63.7	6.7
	[11.4-60.3]	[1.9-7.0]	[30.2-81.1]	[3.1-8.2]	[9.5-61.3]	[1.9-7.0]	[27.9-82.9]	[3.4-10.0]
	BH24				RS24			
h=1	14.3	1.8	78.3	5.3	12.2	1.8	79.2	6.4
	[6.2-29.8]	[0.8-3.7]	[60.9-88.9]	[2.8-8.1]	[5.2-27.2]	[0.9-3.7]	[61.7-89.4]	[3.4-10.1]
h=6	15.9	3.4	75.6	4.7	14.1	3.4	76.4	5.7
	[7.6-31.0]	[1.9-5.8]	[58.7-86.3]	[2.6-7.2]	[6.7-28.8]	[1.9-5.7]	[59.2-86.8]	[3.1-9.0]
h=12	16.1	4.1	74.6	4.8	14.4	4.0	75.4	5.8
	[7.9-30.9]	[2.4-6.5]	[58.1-85.3]	[2.7-7.4]	[7.1-28.9]	[2.4-6.4]	[58.5-85.8]	[3.2-9.0]
h=18	16.2	4.2	74.3	4.9	14.5	4.2	75.1	5.8
	[8.0-30.8]	[2.6-6.8]	[57.9-84.9]	[2.7-7.4]	[7.2-28.8]	[2.5-6.7]	[58.3-85.5]	[3.2-9.1]

Notes: In the table u^S, u^Y, u^D and u^I denote posterior median contributions of flow supply, flow demand, oil preference and storage demand shocks (in %) to the overall variability in the real price of oil along with the 5th and 95th percentile of the posterior distribution (in square brackets). BH16 and BH24 denote results for the original BH model estimated with data ending in December 2016 and in June 2024, respectively. RS16 and RS24 show the results after substituting equations 3 and 4 of the model on respective samples.

Historical decomposition of oil prices (RS24 model)



Notes: Black solid line represents the logarithmic annual rate of change in the real oil price. In the figure u^S, u^Y, u^D and u^I denote the posterior mean contribution of the flow supply, flow demand, oil preference and storage demand shocks, respectively.

Main takeaways

1. We studied how the estimates of the global oil market SVAR model proposed by Baumeister and Hamilton (2019) are affected by an error in the specification of their demand equation. We also checked whether the estimates are sensitive to the inflow of recent data.
2. We show that both changes reduce the estimate of the short-run price elasticity of oil supply.
3. It leads to a substantial decrease in the contribution of oil suppy shocks to real oil price.
4. Our results corroborate the substance of conclusions in Kilian and Murphy (2014), Zhou (2020) or Inoue and Kilian (2022).
5. Results from Baumeister and Hamilton (2019) are not robust to reasonable changes in specifications as shown also by Herrera and Rangaraju (2020) or Braun (2023).

Thank you for attention!

The prior

We decompose the prior:

$$p(\mathbf{A},\mathbf{B},\mathbf{D}) = p(\mathbf{B}|\mathbf{A},\mathbf{D}) \times p(\mathbf{D}|\mathbf{A}) \times p(\mathbf{A}). \tag{1}$$

Prior for the covariance matrix $p(\mathbf{D}|\mathbf{A})$

$$\begin{aligned} p(\mathbf{D}|\mathbf{A}) &= \prod_{i=1}^{n} p(d_{ii}|\mathbf{A}) \\ d_{ii}^{-1}|\mathbf{A} &\sim \Gamma(\kappa_i, \tau_i(\mathbf{A})), \end{aligned} \tag{2}$$

Prior for the matrix of parameters at lagged variables $p(\mathbf{B}|\mathbf{A},\mathbf{D})$

$$\begin{aligned} p(\mathbf{B}|\mathbf{A},\mathbf{D}) &= \prod_{i=1}^{n} p(\mathbf{b}_i|\mathbf{D},\mathbf{A}) \\ \mathbf{b}_i|\mathbf{A},\mathbf{D} &\sim N(\mathbf{m}_i, d_{ii}\mathbf{M}_i), \end{aligned} \tag{3}$$

Prior for the contemporaneous relations matrix $p(\mathbf{A})$ freely chosen.

The posterior (1)

The decomposition of the posterior:

$$p(\mathbf{A}, \mathbf{B}, \mathbf{D}|\mathbf{Y}_T) = p(\mathbf{B}|\mathbf{A}, \mathbf{D}, \mathbf{Y}_T) \times p(\mathbf{D}|\mathbf{A}, \mathbf{Y}_T) \times p(\mathbf{A}|\mathbf{Y}_T). \tag{4}$$

The posterior for the covariance matrix $p(\mathbf{D}|\mathbf{A}, \mathbf{Y}_T)$

$$\begin{aligned} p(\mathbf{D}|\mathbf{A}, \mathbf{Y}_T) &= \prod_{i=1}^{n} p(d_{ii}|\mathbf{A}, \mathbf{Y}_T) \\ d_{ii}^{-1}|\mathbf{A}, \mathbf{Y}_T &\sim \Gamma(\kappa_i^*, \tau_i^*(\mathbf{A})), \end{aligned} \tag{5}$$

where:

$$\begin{aligned} \kappa_i^* &= \kappa_i + (\mu T_1 + T_2)/2 \\ \tau_i^*(\mathbf{A}) &= \tau_i(\mathbf{A}) + \zeta_i^*(\mathbf{A}) \end{aligned} \tag{6}$$

The posterior (2)

$\zeta_i^*(\mathbf{A}) = \left(\widetilde{\mathbf{Y}}_i'(\mathbf{A})\widetilde{\mathbf{Y}}_i(\mathbf{A})\right) - \left(\widetilde{\mathbf{Y}}_i'(\mathbf{A})\widetilde{\mathbf{X}}_i\right)\left(\widetilde{\mathbf{X}}_i'\widetilde{\mathbf{X}}_i\right)^{-1}\left(\widetilde{\mathbf{X}}_i'\widetilde{\mathbf{Y}}_i(\mathbf{A})\right)$ is the sum of squared residuals from regression of $\widetilde{\mathbf{Y}}_i(\mathbf{A})$ on $\widetilde{\mathbf{X}}_i$ defined as:

$$
\begin{aligned}
\underset{(T+k)\times 1}{\widetilde{\mathbf{Y}}_i(\mathbf{A})} &= \left[\sqrt{\mu}\mathbf{y}_1'\mathbf{a}_i \quad \dots \quad \sqrt{\mu}\mathbf{y}_{T_1}'\mathbf{a}_i \quad \mathbf{y}_{T_1+1}'\mathbf{a}_i \quad \dots \quad \mathbf{y}_T'\mathbf{a}_i \quad \mathbf{m}_i'\mathbf{P}_i\right]' \\
\underset{(T+k)\times k}{\widetilde{\mathbf{X}}_i} &= \left[\sqrt{\mu}\mathbf{x}_0 \quad \dots \quad \sqrt{\mu}\mathbf{x}_{T_1-1}' \quad \mathbf{x}_{T_1}' \quad \dots \quad \mathbf{x}_{T-1}' \quad \mathbf{P}_i\right]'
\end{aligned}
\tag{7}
$$

with $\mathbf{P}_i$ being the Cholesky factor of $\mathbf{M}_i^{-1} = \mathbf{P}_i\mathbf{P}_i'$.

The posterior (3)

The posterior for the matrix of parameters at lagged variables $p(\mathbf{B}|\mathbf{A},\mathbf{D},\mathbf{Y}_T)$

$$
\begin{aligned}
p(\mathbf{B}|\mathbf{A},\mathbf{D},\mathbf{Y}_T) &= \prod_{i=1}^{n} p(\mathbf{b}_i|\mathbf{D},\mathbf{A},\mathbf{Y}_T) \\
\mathbf{b}_i|\mathbf{A},\mathbf{D},\mathbf{Y}_T &\sim N(\mathbf{m}_i^*(\mathbf{A}), d_{ii}\mathbf{M}_i^*),
\end{aligned} \tag{8}
$$

where:

$$
\begin{aligned}
m_i^*(\mathbf{A}) &= \left(\widetilde{\mathbf{X}}_i'\widetilde{\mathbf{X}}_i\right)^{-1}\left(\widetilde{\mathbf{X}}_i'\widetilde{\mathbf{Y}}_i(\mathbf{A})\right) \\
M_i^* &= \left(\widetilde{\mathbf{X}}_i'\widetilde{\mathbf{X}}_i\right)^{-1}.
\end{aligned} \tag{9}
$$

The posterior (4)

The posterior for the contemporaneous relations matrix, $p(\mathbf{A}|\mathbf{Y}_T)$

$$\widetilde{\mathbf{\Omega}}_1 = (T_1)^{-1}\left(\sum_{t=1}^{T_1}\mathbf{y}_t\mathbf{y}_t' - \left(\sum_{t=1}^{T_1}\mathbf{y}_t\mathbf{x}_{t-1}'\right)\left(\sum_{t=1}^{T_1}\mathbf{x}_{t-1}\mathbf{x}_{t-1}'\right)^{-1}\left(\sum_{t=1}^{T_1}\mathbf{x}_{t-1}\mathbf{y}_t'\right)\right)$$
$$\widetilde{\mathbf{\Omega}}_2 = (T_2)^{-1}\left(\sum_{t=T_1+1}^{T}\mathbf{y}_t\mathbf{y}_t' - \left(\sum_{t=T_1+1}^{T}\mathbf{y}_t\mathbf{x}_{t-1}'\right)\left(\sum_{t=T_1+1}^{T}\mathbf{x}_{t-1}\mathbf{x}_{t-1}'\right)^{-1}\left(\sum_{t=T_1+1}^{T}\mathbf{x}_{t-1}\mathbf{y}_t'\right)\right)$$
$$\widetilde{\mathbf{\Omega}}_T = (\mu T_1 + T_2)^{-1}\left(\mu T_1\widetilde{\mathbf{\Omega}}_1 + T_2\widetilde{\mathbf{\Omega}}_2\right) \tag{10}$$

The posterior marginal distribution for **A**:

$$p(\mathbf{A}|\mathbf{Y}_T) = k_T p(\mathbf{A})\left[\det(\mathbf{A}\widetilde{\mathbf{\Omega}}_T\mathbf{A}')\right]^{T^*}\prod_{i=1}^{n}\frac{[\tau_i(\mathbf{A})]^{\kappa_i}}{\left[\tau_i^*(\mathbf{A})/T^*\right]^{\kappa_i^*}}, \tag{11}$$