# Uncertainty shocks, monetary policy and long-term interest rates, technical appendix

Gianni Amisano*
Board of Governors
of the Federal Reserve System

Oreste Tristani†
European Central Bank

February 8, 2019

**Abstract**

This appendix describes some important details regarding aspects of the specification and the estimation of the model used in Amisano and Tristani (2019). The appendix is available at the following URL:
`https://sites.google.com/site/gianniamisanowebsite/`

# A Model details

## A.1 The household problem

The optimization problem is:

$$\max U\left[u_t, \mathrm{E}_t V_{t+1}\right] = \left\{(1-\beta)\, u_t^{1-\psi} + \beta \left(\mathrm{E}_t V_{t+1}^{1-\gamma}\right)^{\frac{1-\psi}{1-\gamma}}\right\}^{\frac{1}{1-\psi}},$$

where $u_t$ is shorthand for $u\left\{C_t(j) - h\Xi_t C_{t-1}, 1 - N_t(j)\right\}$, subject to

$$P_t C_t(j) + \mathrm{E}_t Q_{t,t+1} W_{t+1}(j) \leq W_t(j) + w_t(j)\, N_t(j) + \int_0^1 \Psi_t(i)\, di - T_t,$$

and

$$N_t(j) = L_t \left(\frac{w_t(j)}{w_t}\right)^{-\theta_{w,t}},$$

where the choice variables are $w_s$ and $c_s$.

The Bellman equation is

$$\begin{aligned} J_t = J(W_t) = \max &\left\{(1-\beta)\, u_t^{1-\psi} + \beta \left[\mathrm{E}_t J_{t+1}^{1-\gamma}\right]^{\frac{1-\psi}{1-\gamma}}\right\}^{\frac{1}{1-\psi}} + \\ &- \Lambda_t \left[P_t C_t + \mathrm{E}_t Q_{t,t+1} W_{t+1} - W_t - w_t N_t - \int_0^1 \Psi_t(i)\, di + T_t\right], \end{aligned}$$

*Email: gianni.amisano@frb.gov.
†Email: oreste.tristani@ecb.int.

where

$$N_t(j) = L_t \left(\frac{w_t(j)}{w_t}\right)^{-\theta_{w,t}}$$

and

$$\frac{\partial N_t(j)}{\partial w_t(j)} = -\theta_{w,t} \frac{N_t(j)}{w_t(j)}.$$

Using the aggregator function

$$\begin{aligned} U &= \left\{(1-\beta) u_t^{1-\psi} + \beta v_t^{1-\psi}\right\}^{\frac{1}{1-\psi}} \\ v_t &\equiv \left[\mathrm{E}_t J_{t+1}^{1-\gamma}\right]^{\frac{1}{1-\gamma}}, \end{aligned}$$

we define

$$\begin{aligned} U_{u,t} &= (1-\beta)\left\{(1-\beta) u_t^{1-\psi} + \beta v_t^{1-\psi}\right\}^{\frac{\psi}{1-\psi}} u_t^{-\psi}, \\ U_{v,t} &= \beta \left\{(1-\beta) u_t^{1-\psi} + \beta v_t^{1-\psi}\right\}^{\frac{\psi}{1-\psi}} v_t^{-\psi}. \end{aligned}$$

The FOCs include

$$U_{u,t} u_{c,t} = \Lambda_t P_t,$$

$$u_{N,t} U_{u,t} \frac{\partial N_t(j)}{\partial w_t(j)} = -\Lambda_t \left[N_t(j) + w_t(j) \frac{\partial N_t(j)}{\partial w_t(j)}\right],$$

and, state-by-state

$$U_{v,t} \left[\mathrm{E}_t J_{t+1}^{1-\gamma}\right]^{\frac{\gamma}{1-\gamma}} J_{t+1}^{-\gamma} J_{W,t+1} = \Lambda_t Q_{t,t+1},$$

$$J_{W,t+1} = \frac{\partial J_{t+1}}{\partial W_{t+1}}$$

plus envelope

$$J_{W,t} = \Lambda_t.$$

The FOCs can be rewritten as

$$\begin{aligned} \frac{\Lambda_t P_t}{u_{c,t}} &= U_{u,t}, \\ \frac{u_{N,t}}{u_{c,t}} &= \frac{1-\theta_{w,t}}{\theta_{w,t}} \frac{w_t(j)}{P_t}, \\ Q_{t,t+1} &= U_{v,t} \left[\mathrm{E}_t J_{t+1}^{1-\gamma}\right]^{\frac{\gamma}{1-\gamma}} J_{t+1}^{-\gamma} \frac{\Lambda_{t+1}}{\Lambda_t}, \end{aligned}$$

or

$$Q_{t,t+1} = \beta \left(\frac{\left[\mathrm{E}_t J_{t+1}^{1-\gamma}\right]^{\frac{1}{1-\gamma}}}{J_{t+1}}\right)^{\gamma-\psi} \frac{u_{t+1}^{-\psi}}{u_t^{-\psi}} \frac{u_{c,t+1}}{u_{c,t}} \frac{1}{\Pi_{t+1}}.$$

Using the definition of $\mu_{w,t}$, we obtain, as in the text,

$$-\frac{u_{N,t}}{u_{c,t}} = \mu_{w,t} \frac{w_t(j)}{P_t}$$

and

$$Q_{t,t+1} = \beta \left[\mathrm{E}_t \left(\frac{J_{t+1}}{J_t}\right)^{1-\gamma}\right]^{\frac{\gamma-\psi}{1-\gamma}} \left(\frac{J_{t+1}}{J_t}\right)^{-(\gamma-\psi)} \left(\frac{u_{t+1}}{u_t}\right)^{-\psi} \frac{u_{c,t+1}}{u_{c,t}} \frac{1}{\Pi_{t+1}}.$$

## A.2 Detrending

Given the stochastic trend $B_t$, define a detrended variable as $\widetilde{x}_t \equiv x_t/B_t$. It follows that we can rewrite the conditions above as

$$
\begin{aligned}
-\frac{\widetilde{u}_{N,t}}{u_{c,t}} &= \frac{\theta_{w,t}-1}{\theta_{w,t}}\frac{\widetilde{w}_t(j)}{P_t}, \\
\widetilde{J}_t^{1-\psi} &= (1-\beta)\,\widetilde{u}_t^{1-\psi} + \beta\left[\mathrm{E}_t \Xi_{t+1}^{1-\gamma}\widetilde{J}_{t+1}^{1-\gamma}\right]^{\frac{1-\psi}{1-\gamma}}, \\
\widetilde{u}_t &= u\left(\widetilde{C}_t(j) - h\widetilde{C}_{t-1}, 1-N_t(j)\right), \\
Q_{t,t+1} &= \beta\left(\frac{\left[\mathrm{E}_t \widetilde{J}_{t+1}^{1-\gamma}\Xi_{t+1}^{1-\gamma}\right]^{\frac{1}{1-\gamma}}}{\widetilde{J}_{t+1}\Xi_{t+1}}\right)^{\gamma-\psi}\left(\frac{\widetilde{u}_{t+1}}{\widetilde{u}_t}\right)^{-\psi}\frac{u_{c,t+1}}{u_{c,t}}\frac{1}{\Pi_{t+1}\Xi_{t+1}^{\psi}}.
\end{aligned}
$$

## A.3 Consumption

To second order, the Euler equation can be written as

$$
-\widehat{i}_t + \frac{1}{2}\widehat{i}_t^2 = \mathrm{E}_t\widehat{q}_{t,t+1} + \frac{1}{2}\mathrm{E}_t\widehat{q}_{t,t+1}^2
$$

where $\widehat{i}_t^2$ can be derived using only first order terms to obtain

$$
\widehat{i}_t = -\mathrm{E}_t\widehat{q}_{t,t+1} - \frac{1}{2}\mathrm{Var}_t\widehat{q}_{t,t+1}.
$$

We rewrite the stochastic discount factor $Q_{t,t+1}$ as

$$
\begin{aligned}
\widetilde{\Lambda}_t &\equiv \widetilde{u}_t^{-\psi}u_{c,t}, \\
D_t &\equiv \mathrm{E}_t\widetilde{J}_{t+1}^{1-\gamma}\Xi_{t+1}^{1-\gamma}, \\
Q_{t,t+1} &= \beta\frac{D_t^{\frac{\gamma-\psi}{1-\gamma}}}{\widetilde{J}_{t+1}^{\gamma-\psi}}\frac{\widetilde{\Lambda}_{t+1}}{\widetilde{\Lambda}_t}\frac{1}{\Pi_{t+1}\Xi_{t+1}^{\gamma}},
\end{aligned}
$$

so that

$$
\widehat{q}_{t,t+1} = \Delta\widehat{\widetilde{\lambda}}_{t+1} - \widehat{\pi}_{t+1} - \psi\widehat{\xi}_{t+1} + \frac{\gamma-\psi}{1-\gamma}\widehat{d}_t - (\gamma-\psi)\,\widehat{\widetilde{j}}_{t+1}.
$$

Now, we expand $\widehat{d}_t$ to second order (again using only first order terms to evaluate $\widehat{d}_t^2$) to find

$$
\begin{aligned}
\widehat{d}_t = {} & (1-\gamma)\,\mathrm{E}_t\widehat{\xi}_{t+1} + (1-\gamma)\,\mathrm{E}_t\widehat{\widetilde{j}}_{t+1} + \frac{1}{2}(1-\gamma)^2\,\mathrm{Var}_t\widehat{\xi}_{t+1} + \\
& + \frac{1}{2}(1-\gamma)^2\,\mathrm{Var}_t\widehat{\widetilde{j}}_{t+1} + (1-\gamma)^2\,\mathrm{Cov}_t\widehat{\xi}_{t+1}\widehat{\widetilde{j}}_{t+1}.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
\widehat{q}_{t,t+1} = {} & \Delta\widehat{\widetilde{\lambda}}_{t+1} - \psi\widehat{\xi}_{t+1} - \widehat{\pi}_{t+1} - (\gamma-\psi)\left(\widehat{\xi}_{t+1} + \widehat{\widetilde{j}}_{t+1} - \mathrm{E}_t\left[\widehat{\xi}_{t+1} + \widehat{\widetilde{j}}_{t+1}\right]\right) + \\
& + \frac{1}{2}(1-\gamma)(\gamma-\psi)\,\mathrm{Var}_t\left[\widehat{\xi}_{t+1} + \widehat{\widetilde{j}}_{t+1}\right].
\end{aligned}
$$

We now evaluate $\mathrm{E}_t\widehat{q}_{t,t+1}$ and $\mathrm{Var}_t\widehat{q}_{t,t+1}$ to obtain

$$\mathrm{E}_t\widehat{q}_{t,t+1} = \mathrm{E}_t\Delta\widehat{\widehat{\lambda}}_{t+1} - \psi\mathrm{E}_t\widehat{\xi}_{t+1} - \mathrm{E}_t\widehat{\pi}_{t+1} + \frac{1}{2}\left(1-\gamma\right)\left(\gamma-\psi\right)\mathrm{Var}_t\left[\widehat{\xi}_{t+1} + \widehat{\widehat{j}}_{t+1}\right],$$

and (using first order terms to evaluate $\widehat{q}^2_{t,t+1}$)

$$\begin{aligned}\mathrm{E}_t\widehat{q}^2_{t,t+1} = {} & \mathrm{Var}_t\left[\Delta\widehat{\widehat{\lambda}}_{t+1} - \psi\widehat{\xi}_{t+1} - \widehat{\pi}_{t+1}\right] + \left(\gamma-\psi\right)^2\mathrm{Var}_t\left[\widehat{\xi}_{t+1} + \widehat{\widehat{j}}_{t+1}\right] + \\ & - 2\left(\gamma-\psi\right)\mathrm{Cov}_t\left[\Delta\widehat{\widehat{\lambda}}_{t+1} - \psi\widehat{\xi}_{t+1} - \widehat{\pi}_{t+1}, \widehat{\xi}_{t+1} + \widehat{\widehat{j}}_{t+1}\right].\end{aligned}$$

Hence

$$\begin{aligned}\widehat{i}_t = {} & \mathrm{E}_t\left[-\Delta\widehat{\widehat{\lambda}}_{t+1} + \psi\widehat{\xi}_{t+1} + \widehat{\pi}_{t+1}\right] + \frac{1}{2}\left(\gamma-\psi\right)\left(\psi-1\right)\mathrm{Var}_t\left[\widehat{\xi}_{t+1} + \widehat{\widehat{j}}_{t+1}\right] + \\ & - \frac{1}{2}\mathrm{Var}_t\left[\Delta\widehat{\widehat{\lambda}}_{t+1} - \psi\widehat{\xi}_{t+1} - \widehat{\pi}_{t+1}\right] + \\ & + \left(\gamma-\psi\right)\mathrm{Cov}_t\left[\Delta\widehat{\widehat{\lambda}}_{t+1} - \psi\widehat{\xi}_{t+1} - \widehat{\pi}_{t+1}, \widehat{\xi}_{t+1} + \widehat{\widehat{j}}_{t+1}\right]\end{aligned}$$

We now expand $\widehat{\widehat{\lambda}}_{t+1}$ and $\widehat{\widehat{j}}_{t+1}$ for the specific case of the Trabandt and Uhlig (2011) form for temporary utility, which we use in the paper

$$\widetilde{u}_t = \left(\widetilde{C}_t - h\widetilde{C}_{t-1}\right)\left(1 - \eta\left(1-\psi\right)N_t^{1+\frac{1}{\phi}}\right)^{\frac{\psi}{1-\psi}}.$$

The expression above, after defining surplus consumption $\overleftrightarrow{c}_t = \widetilde{C}_t - h\widetilde{C}_{t-1}$, implies

$$\begin{aligned}\widetilde{\Lambda}_t &= \overleftrightarrow{c}_t^{-\psi}\left(1 - \eta\left(1-\psi\right)N_t^{1+\frac{1}{\phi}}\right)^{\psi}, \\ \widetilde{J}_t^{1-\psi} &= \left(1-\beta\right)\overleftrightarrow{c}_t\widetilde{\overleftrightarrow{\Lambda}}_t + \beta\left[\mathrm{E}_t\Xi_{t+1}^{1-\gamma}\widetilde{J}_{t+1}^{1-\gamma}\right]^{\frac{1-\psi}{1-\gamma}}.\end{aligned}$$

### A.3.1 Expanding $\Lambda_t$

To second order

$$\begin{aligned}\widehat{\widehat{\lambda}}_t + \frac{1}{2}\widehat{\widehat{\lambda}}_t^2 = {} & -\psi\widehat{\overleftrightarrow{c}}_t - \psi\left(1+\frac{1}{\phi}\right)\frac{\eta\left(1-\psi\right)N_t^{1+\frac{1}{\phi}}}{1-\eta\left(1-\psi\right)N_t^{1+\frac{1}{\phi}}}\widehat{l}_t + \frac{1}{2}\psi^2\widehat{\overleftrightarrow{c}}_t^2 + \\ & - \frac{1}{2}\psi\left(1+\frac{1}{\phi}\right)^2\frac{\eta\left(1-\psi\right)N_t^{1+\frac{1}{\phi}}\left(1-\eta\sigma\left(1-\psi\right)N_t^{1+\frac{1}{\phi}}\right)}{\left(1-\eta\left(1-\psi\right)N_t^{1+\frac{1}{\phi}}\right)^2}\widehat{l}_t^2 + \\ & + \psi^2\left(1+\frac{1}{\phi}\right)\frac{\eta\left(1-\psi\right)N_t^{1+\frac{1}{\phi}}}{1-\eta\left(1-\psi\right)N_t^{1+\frac{1}{\phi}}}\widehat{\overleftrightarrow{c}}_t\widehat{l}_t,\end{aligned}$$

and using first order terms to evaluate $\widehat{\widehat{\lambda}}_t^2$

$$\widehat{\widehat{\lambda}}_t = -\psi \widehat{\overleftrightarrow{c}}_t - \psi \left(1 + \frac{1}{\phi}\right) \frac{\eta (1-\psi) N_t^{1+\frac{1}{\phi}}}{1 - \eta (1-\psi) N_t^{1+\frac{1}{\phi}}} \widehat{l}_t - \frac{1}{2} \psi \left(1 + \frac{1}{\phi}\right)^2 \frac{\eta (1-\psi) N_t^{1+\frac{1}{\phi}}}{\left(1 - \eta (1-\psi) N_t^{1+\frac{1}{\phi}}\right)^2} \widehat{l}_t^2$$

It follows that

$$\Delta \widehat{\widehat{\lambda}}_{t+1} = -\sigma \Delta \widehat{\overleftrightarrow{c}}_{t+1} - \psi \left(1 + \frac{1}{\phi}\right) \frac{\overline{n}}{1-\overline{n}} \Delta \widehat{l}_{t+1} - \frac{1}{2} \psi \left(1 + \frac{1}{\phi}\right)^2 \frac{\overline{n}}{(1-\overline{n})^2} \left(\widehat{l}_{t+1}^2 - \widehat{l}_t^2\right)$$

for

$$\overline{n} \equiv \eta (1-\psi) N^{1+\frac{1}{\phi}}$$

Surplus consumption $\overleftrightarrow{c}_t$ can be expanded as

$$\widehat{\overleftrightarrow{c}}_t = \frac{1}{1-h} \left(\widehat{\widehat{c}}_t - h \widehat{\widehat{c}}_{t-1}\right) - \frac{1}{2} \frac{h}{(1-h)^2} \left(\widehat{\widehat{c}}_t - \widehat{\widehat{c}}_{t-1}\right)^2,$$

so that

$$\begin{aligned}
\Delta \widehat{\widehat{\lambda}}_{t+1} = & -\psi \frac{1}{1-h} \left(\Delta \widehat{\widehat{c}}_{t+1} - h \Delta \widehat{\widehat{c}}_t\right) - \psi \left(1 + \frac{1}{\phi}\right) \frac{\overline{n}}{1-\overline{n}} \Delta \widehat{l}_{t+1} + \\
& + \frac{1}{2} \psi \frac{h}{(1-h)^2} \left[\left(\Delta \widehat{\widehat{c}}_{t+1}\right)^2 - \left(\Delta \widehat{\widehat{c}}_t\right)^2\right] - \frac{1}{2} \psi \left(1 + \frac{1}{\phi}\right)^2 \frac{\overline{n}}{(1-\overline{n})^2} \left(\widehat{l}_{t+1}^2 - \widehat{l}_t^2\right).
\end{aligned}$$

### A.3.2 Expanding $\widetilde{J}$

Note that $\widehat{\widehat{J}}_{t+1}$ only enters the interest rate in terms of second order. It can therefore be evaluated to first order. We obtain

$$\widetilde{J}_t^{1-\psi} = (1-\beta) \overleftrightarrow{c}_t \widetilde{\Lambda}_t + \beta D_t^{\frac{1-\psi}{1-\gamma}}$$

so that

$$(1-\psi) \widehat{\widehat{j}}_t = \frac{(1-\beta) \overleftrightarrow{c} \widetilde{\Lambda}}{(1-\beta) \overleftrightarrow{c} \widetilde{\Lambda} + \beta D^{\frac{1-\psi}{1-\gamma}}} \widehat{\overleftrightarrow{c}}_t + \frac{(1-\beta) \overleftrightarrow{c} \widetilde{\Lambda}}{(1-\beta) \overleftrightarrow{c} \widetilde{\Lambda} + \beta D^{\frac{1-\psi}{1-\gamma}}} \widehat{\widehat{\lambda}}_t + \frac{1-\psi}{1-\gamma} \frac{\beta D^{\frac{1-\psi}{1-\gamma}}}{(1-\beta) \overleftrightarrow{c} \widetilde{\Lambda} + \beta D^{\frac{1-\psi}{1-\gamma}}} \widehat{\widehat{d}}_t$$

and, using $\widehat{d}_t \equiv (1-\gamma) \mathrm{E}_t \widehat{\xi}_{t+1} + (1-\gamma) \mathrm{E}_t \widehat{\widehat{j}}_{t+1}$,

$$\begin{aligned}
(1-\psi) \widehat{\widehat{j}}_t = & \frac{(1-\beta) \overleftrightarrow{c} \widetilde{\Lambda}}{(1-\beta) \overleftrightarrow{c} \widetilde{\Lambda} + \beta D^{\frac{1-\psi}{1-\gamma}}} \widehat{\overleftrightarrow{c}}_t + \frac{(1-\beta) \overleftrightarrow{c} \widetilde{\Lambda}}{(1-\beta) \overleftrightarrow{c} \widetilde{\Lambda} + \beta D^{\frac{1-\psi}{1-\gamma}}} \widehat{\widehat{\lambda}}_t + \\
& + (1-\psi) \frac{\beta D^{\frac{1-\psi}{1-\gamma}}}{(1-\beta) \overleftrightarrow{c} \widetilde{\Lambda} + \beta D^{\frac{1-\psi}{1-\gamma}}} \left(\mathrm{E}_t \widehat{\xi}_{t+1} + \mathrm{E}_t \widehat{\widehat{j}}_{t+1}\right).
\end{aligned}$$

Since to first order $\widehat{\widehat{\lambda}}_t = -\psi \widehat{\overleftrightarrow{c}}_t - \psi \left(1 + \frac{1}{\phi}\right) \frac{\eta(1-\psi) N^{1+\frac{1}{\phi}}}{1-\eta(1-\psi) N^{1+\frac{1}{\phi}}} \widehat{l}_t$, we further obtain

$$\begin{aligned}
\widehat{\widehat{j}}_t = & \frac{(1-\beta) \overleftrightarrow{c} \widetilde{\Lambda}}{(1-\beta) \overleftrightarrow{c} \widetilde{\Lambda} + \beta D^{\frac{1-\psi}{1-\gamma}}} \widehat{\overleftrightarrow{c}}_t - \frac{\psi}{1-\psi} \left(1 + \frac{1}{\phi}\right) \frac{\eta (1-\psi) N^{1+\frac{1}{\phi}}}{1 - \eta (1-\psi) N^{1+\frac{1}{\phi}}} \frac{(1-\beta) \overleftrightarrow{c} \widetilde{\Lambda}}{(1-\beta) \overleftrightarrow{c} \widetilde{\Lambda} + \beta D^{\frac{1-\psi}{1-\gamma}}} \widehat{l}_t \\
& + \frac{\beta D^{\frac{1-\psi}{1-\gamma}}}{(1-\beta) \overleftrightarrow{c} \widetilde{\Lambda} + \beta D^{\frac{1-\psi}{1-\gamma}}} \left(\mathrm{E}_t \widehat{\xi}_{t+1} + \mathrm{E}_t \widehat{\widehat{j}}_{t+1}\right).
\end{aligned}$$

Recall that in steady state

$$
\begin{aligned}
\widetilde{J}^{1-\psi} &= \frac{1-\beta}{1-\beta\Xi^{1-\psi}} \overleftrightarrow{c}\widetilde{\Lambda}, \\
\widetilde{\Lambda} &= \overleftrightarrow{c}^{-\psi}\left(1-\eta\left(1-\psi\right)N^{1+\frac{1}{\phi}}\right)^{\psi}, \\
D &\equiv \Xi^{1-\gamma}\widetilde{J}^{1-\gamma},
\end{aligned}
$$

to obtain

$$
\widehat{\widehat{j}}_t = \left(1-\beta\Xi^{1-\psi}\right)\left(\widehat{\overleftrightarrow{c}}_t - \frac{\psi}{1-\psi}\left(1+\frac{1}{\phi}\right)\frac{\overline{n}}{1-\overline{n}}\widehat{l}_t\right) + \beta\Xi^{1-\psi}\left(\mathrm{E}_t\widehat{\xi}_{t+1} + \mathrm{E}_t\widehat{\widehat{j}}_{t+1}\right).
$$

This can be solved forward to obtain

$$
\begin{aligned}
\widehat{\widehat{j}}_t + \widehat{\xi}_t = \lim_{n\to\infty}\sum_{i=0}^{n}\left[\beta\Xi^{1-\psi}\right]^i \mathrm{E}_t\left[\widehat{\xi}_{t+i} + \left(1-\beta\Xi^{1-\psi}\right)\left(\widehat{\overleftrightarrow{c}}_{t+i} - \frac{\psi}{1-\psi}\left(1+\frac{1}{\phi}\right)\frac{\overline{n}}{1-\overline{n}}\widehat{l}_{t+i}\right)\right] + \\
+ \lim_{n\to\infty}\left[\beta\Xi^{1-\psi}\right]^n\left(\mathrm{E}_t\widehat{\xi}_{t+n} + \mathrm{E}_t\widehat{\widehat{j}}_{t+n}\right).
\end{aligned}
$$

Assuming that $\lim_{n\to\infty}\left[\beta\Xi^{1-\psi}\right]^n\left(\mathrm{E}_t\widehat{\xi}_{t+n} + \mathrm{E}_t\widehat{\widehat{j}}_{t+n}\right) = 0$ and that the other sums converge, we obtain

$$
\widehat{\widehat{j}}_t + \widehat{\xi}_t = \sum_{i=0}^{\infty}\left(\beta\Xi^{1-\psi}\right)^i \mathrm{E}_t\left[\widehat{\xi}_{t+i} + \left(1-\beta\Xi^{1-\psi}\right)\left(\widehat{\overleftrightarrow{c}}_{t+i} - \frac{\psi}{1-\psi}\left(1+\frac{1}{\phi}\right)\frac{\overline{n}}{1-\overline{n}}\widehat{l}_{t+i}\right)\right],
$$

or using the first order expansion of $\widehat{\overleftrightarrow{c}}_t$,

$$
\widehat{\widehat{j}}_t+\widehat{\xi}_t = \sum_{i=0}^{\infty}\left(\beta\Xi^{1-\psi}\right)^i \mathrm{E}_t\left[\widehat{\xi}_{t+i} + \left(1-\beta\Xi^{1-\psi}\right)\left(\frac{1}{1-h}\left(\widehat{\widehat{c}}_{t+i} - h\widehat{\widehat{c}}_{t+i-1}\right) - \frac{\psi}{1-\psi}\left(1+\frac{1}{\phi}\right)\frac{\overline{n}}{1-\overline{n}}\widehat{l}_{t+i}\right)\right].
$$

Note that the first order approximation of temporary utility is

$$
\widehat{\widehat{u}}_t = \frac{1}{1-h}\left(\widehat{\widehat{c}}_t - h\widehat{\widehat{c}}_{t-1}\right) - \frac{\psi}{1-\psi}\left(1+\frac{1}{\phi}\right)\frac{\overline{n}}{1-\overline{n}}\widehat{l}_t.
$$

Hence

$$
\widehat{\widehat{j}}_t + \widehat{\xi}_t = \sum_{i=0}^{\infty}\left(\beta\Xi^{1-\psi}\right)^i \mathrm{E}_t\left[\widehat{\xi}_{t+i} + \left(1-\beta\Xi^{1-\psi}\right)\widehat{\widehat{u}}_{t+i}\right]
$$

### A.3.3 Second order approximation to the Euler equation

It follows that the second order approximation to the Euler equation, using also the assumption $\mathrm{E}_t\left[\widehat{\xi}_{t+1}\right] = 0$, can be written as

$$
\begin{aligned}
\widehat{\widehat{c}}_t = \frac{1}{1+h}\mathrm{E}_t\widehat{\widehat{c}}_{t+1} + \frac{h}{1+h}\widehat{\widehat{c}}_{t-1} - \frac{1}{\psi}\frac{1-h}{1+h}\left(\widehat{i}_t - \mathrm{E}_t\left[\widehat{\pi}_{t+1}\right]\right) + \frac{1-h}{1+h}\left(1+\frac{1}{\phi}\right)\frac{\overline{n}}{1-\overline{n}}\mathrm{E}_t\Delta\widehat{l}_{t+1} - \frac{1}{2}\mathrm{Var}_t\Omega_{t+1} \\
+ \frac{1}{2}\left(\gamma-\psi\right)\frac{\psi-1}{\psi}\frac{1-h}{1+h}\mathrm{Var}_t\left[\widehat{\xi}_{t+1} + \widehat{\widehat{j}}_{t+1}\right] - \left(\gamma-\psi\right)\frac{1}{\psi}\frac{1-h}{1+h}\mathrm{Cov}_t\left[\psi\widehat{\xi}_{t+1} + \widehat{\pi}_{t+1}, \widehat{\xi}_{t+1} + \widehat{\widehat{j}}_{t+1}\right] \\
- \left(\gamma-\psi\right)\frac{1-h}{1+h}\mathrm{Cov}_t\left[\frac{1}{1-h}\left(\Delta\widehat{\widehat{c}}_{t+1} - h\Delta\widehat{\widehat{c}}_t\right) + \left(1+\frac{1}{\phi}\right)\frac{\overline{n}}{1-\overline{n}}\Delta\widehat{l}_{t+1}, \widehat{\xi}_{t+1} + \widehat{\widehat{j}}_{t+1}\right]
\end{aligned}
$$

for

$$
\begin{aligned}
\mathrm{Var}_t\Omega_{t+1} = {} & \frac{h}{1-h^2}\left[\mathrm{E}_t\left(\Delta\widehat{\widehat{c}}_{t+1}\right)^2 - \left(\Delta\widehat{\widehat{c}}_t\right)^2\right] - \frac{1-h}{1+h}\left(1+\frac{1}{\phi}\right)^2 \frac{\overline{n}}{(1-\overline{n})^2}\left(\mathrm{E}_t\widehat{l}_{t+1}^2 - \widehat{l}_t^2\right) \\
& + \frac{1}{\psi}\frac{1-h}{1+h}\mathrm{Var}_t\left[\psi\frac{1}{1-h}\left(\Delta\widehat{\widehat{c}}_{t+1} - h\Delta\widehat{\widehat{c}}_t\right) + \psi\left(1+\frac{1}{\phi}\right)\frac{\overline{n}}{1-\overline{n}}\Delta\widehat{l}_{t+1} + \psi\widehat{\xi}_{t+1} + \widehat{\pi}_{t+1}\right],
\end{aligned}
$$

and

$$
\widehat{\widehat{\jmath}}_{t+1} + \widehat{\xi}_{t+1} = \sum_{i=0}^{\infty}\left(\beta\Xi^{1-\psi}\right)^i \mathrm{E}_t\left[\widehat{\xi}_{t+1+i} + \left(1-\beta\Xi^{1-\psi}\right)\widehat{\widehat{u}}_{t+1+i}\right]
$$

and

$$
\overline{n} = \eta\left(1-\psi\right)N^{1+\frac{1}{\phi}}.
$$

## A.4 Expected excess holding period returns

Recall that

$$
HPR_{n,t} = \frac{\mathrm{E}_t B_{n-1,t+1}}{B_{n,t}},
$$

and

$$
\widehat{\imath}_{t,t+n} = -\frac{1}{n}\widehat{b}_{t,t+n},
$$

so that

$$
HPR_{2,t} = \frac{\mathrm{E}_t B_{1,t+1}}{B_{2,t}},
$$

To second order, expected holding period return are

$$
\widehat{h}_{n,t} + \frac{1}{2}\widehat{h}_{n,t}^2 = \mathrm{E}_t\widehat{b}_{n-1,t+1} - \widehat{b}_{n,t} + \frac{1}{2}\mathrm{E}_t\widehat{b}_{n-1,t+1}^2 + \frac{1}{2}\widehat{b}_{n,t}^2 - \widehat{b}_{n,t}\mathrm{E}_t\widehat{b}_{n-1,t+1}
$$

with

$$
\frac{1}{2}\widehat{h}_{n,t}^2 = \frac{1}{2}\left(\mathrm{E}_t\widehat{b}_{n-1,t+1}\right)^2 + \frac{1}{2}\widehat{b}_{n,t}^2 - \widehat{b}_{n,t}\mathrm{E}_t\widehat{b}_{n-1,t+1},
$$

so that

$$
\widehat{h}_{n,t} = -\widehat{b}_{n,t} + \mathrm{E}_t\widehat{b}_{n-1,t+1} + \frac{1}{2}\mathrm{E}_t\widehat{b}_{n-1,t+1}^2 - \frac{1}{2}\left(\mathrm{E}_t\widehat{b}_{n-1,t+1}\right)^2,
$$

or

$$
\widehat{h}_{n,t} = -\widehat{b}_{n,t} + \mathrm{E}_t\widehat{b}_{n-1,t+1} + \frac{1}{2}\mathrm{Var}_t\widehat{b}_{n-1,t+1},
$$

and since bond prices are

$$
\widehat{b}_{t,n} = \widehat{b}_{t,1} + \mathrm{E}_t\widehat{b}_{t+1,n-1} + \frac{1}{2}\mathrm{Var}_t\widehat{b}_{t+1,n-1} + \mathrm{Cov}_t\left[\widehat{b}_{t+1,n-1}, \widehat{q}_{t,t+1}\right],
$$

we can in general rewrite expected holding period returns as

$$
\widehat{h}_{n,t} = \widehat{\imath}_t - \mathrm{Cov}_t\left[\widehat{b}_{t+1,n-1}, \widehat{q}_{t,t+1}\right].
$$

Excess holding period returns are therefore

$$
\widehat{h}_{n,t} - \widehat{\imath}_t = -\mathrm{Cov}_t\left[\widehat{b}_{t+1,n-1}, \widehat{q}_{t,t+1}\right].
$$

Using the approximated stochastic discount factor, we obtain

$$\widehat{h}_{n,t} - \widehat{i}_t = -\mathrm{Cov}_t \left[ \widehat{b}_{t+1,n-1}, \Delta\widehat{\widehat{\lambda}}_{t+1} - \psi\widehat{\xi}_{t+1} - \widehat{\pi}_{t+1} - (\gamma - \psi) \left( \widehat{\xi}_{t+1} + \widehat{\widehat{j}}_{t+1} \right) \right]$$

Now use the first order expansion of $\widehat{\widehat{\lambda}}_t$ to write

$$\widehat{h}_{n,t} - \widehat{i}_t = \mathrm{Cov}_t \left[ \widehat{b}_{t+1,n-1}, -\frac{\psi}{1-h}\widehat{\widehat{c}}_{t+1} - \psi \left(1 + \frac{1}{\phi}\right) \frac{\overline{n}}{1-\overline{n}}\widehat{l}_{t+1} - \psi\widehat{\xi}_{t+1} - \widehat{\pi}_{t+1} - (\gamma - \psi) \left( \widehat{\xi}_{t+1} + \widehat{\widehat{j}}_{t+1} \right) \right].$$

Define the first order approximation of variable $v$ as $F_v\widehat{\mathbf{x}}_t$. Then (note that we use $F_j$ to denote the first order approximation of the infinite sum $\widehat{\xi}_{t+1} + \widehat{\widehat{j}}_{t+1}$)

$$\widehat{h}_{n,t} - \widehat{i}_t = \mathrm{Cov}_t \left[ F_{B_{n-1}}\widehat{x}_{t+1}, \left( \psi\frac{1}{1-h}F_c + \psi \left(1 + \frac{1}{\phi}\right) \frac{\overline{n}}{1-\overline{n}}F_l + \psi F_\xi + F_\pi + (\gamma - \psi) F_j \right) \widehat{x}_{t+1} \right].$$

It follows that

$$\widehat{h}_{n,t} - \widehat{i}_t = \mathrm{E}_t \left[ F'_{B_{n-1}}\widehat{x}_{t+1}\widehat{x}'_{t+1} \left( \psi\frac{1}{1-h}F_c + \psi \left(1 + \frac{1}{\phi}\right) \frac{\overline{n}}{1-\overline{n}}F_l + \psi F_\xi + F_\pi + (\gamma - \psi) F_j \right)' \right] +$$
$$- \mathrm{E}_t F'_{B_{n-1}}\widehat{x}'_{t+1}\mathrm{E}_t \left[ \widehat{\mathbf{x}}'_{t+1} \left( \psi\frac{1}{1-h}F_c + \psi \left(1 + \frac{1}{\phi}\right) \frac{\overline{n}}{1-\overline{n}}F_l + \psi F_\xi + F_\pi + (\gamma - \psi) F_j \right)' \right],$$

and using the law of motion for $\widehat{x}_{t+1}$

$$\widehat{h}_{n,t} - \widehat{i}_t = \widetilde{\sigma}^2 F_{B_{n-1}}\mathrm{E}_t \left[ u_{t+1}u'_{t+1} \right] \left( \psi\frac{1}{1-h}F_c + \psi \left(1 + \frac{1}{\phi}\right) \frac{\overline{n}}{1-\overline{n}}F_l + \psi F_\xi + F_\pi + (\gamma - \psi) F_j \right)'$$

## A.5 Firms' optimization problem

Under Rotemberg prices, firm $j$ maximizes real profits

$$\max_{P_t^j} \mathrm{E}_t \sum_{s=t}^{\infty} Q_{t,s} \left[ \frac{P_s^j Y_s^j}{P_s} - \frac{w_s}{P_s} \left( \frac{Y_s^j}{A_s} \right)^{\frac{1}{\alpha}} - \frac{\zeta}{2} \left( \frac{P_s^j}{P_{s-1}^j} - (\Pi^*)^{1-\iota} \Pi_{s-1}^{\iota} \right)^2 Y_s \right],$$

subject to the total demand for its output

$$Y_t(j) = \left( \frac{P_t(j)}{P_t} \right)^{-\theta} Y_t,$$

and to the production function

$$Y_t(j) = A_t L_t^{\alpha}(j),$$

where $L_t$ is the labour index defined above.

The FOC is

$$0 = (1-\theta) \left( \frac{P_t^j}{P_t} \right)^{-\theta} Y_t \frac{1}{P_t} + \frac{\theta}{\alpha} \frac{w_t}{P_t} \left( \frac{Y_t}{A_t} \right)^{\frac{1}{\alpha}} \left( \frac{P_t^j}{P_t} \right)^{-\frac{\theta}{\alpha}-1} \frac{1}{P_t} - \zeta \left( \frac{P_t^j}{P_{t-1}^j} - (\Pi^*)^{1-\iota} \Pi_{t-1}^{\iota} \right) Y_t \frac{1}{P_{t-1}^j} +$$
$$+ \mathrm{E}_t Q_{t,t+1} \zeta \left( \frac{P_{t+1}^j}{P_t^j} - (\Pi^*)^{1-\iota} \Pi_t^{\iota} \right) Y_{t+1} \frac{P_{t+1}^j}{P_t^j} \frac{1}{P_t^j},$$

or, noting that all firms will set the same price and expressing variables in detrended form,

$$(\theta - 1)\widetilde{Y}_t + \zeta \left( \Pi_t - (\Pi^*)^{1-\iota} \Pi_{t-1}^{\iota} \right) \widetilde{Y}_t \Pi_t = \frac{\theta}{\alpha} \frac{\widetilde{w}_t}{P_t} \frac{1}{Z_t^{\frac{1}{\alpha}}} \widetilde{Y}_t^{\frac{1}{\alpha}} + \mathrm{E}_t Q_{t,t+1} \zeta \left( \Pi_{t+1} - (\Pi^*)^{1-\iota} \Pi_t^{\iota} \right) \widetilde{Y}_{t+1} \Xi_{t+1} \Pi_{t+1}.$$

### A.6 Equilibrium

Equilibrium is described by the following system:

- households

$$
\begin{aligned}
\frac{\Lambda_t P_t}{u_{c,t}} &= (1-\beta)\,\widetilde{u}_t^{-\psi}\widetilde{J}_t^{\psi},\\
-\frac{\widetilde{u}_{N,t}}{u_{c,t}} &= \frac{\theta_{w,t}-1}{\theta_{w,t}}\frac{\widetilde{w}_t}{P_t},\\
\widetilde{J}_t^{1-\psi} &= (1-\beta)\,\widetilde{u}_t^{1-\psi} + \beta\left[\mathrm{E}_t \Xi_{t+1}^{1-\gamma}\widetilde{J}_{t+1}^{1-\gamma}\right]^{\frac{1-\psi}{1-\gamma}},\\
\widetilde{u}_t &= u\left(\widetilde{C}_t - h\widetilde{C}_{t-1}, 1-N_t\right),\\
Q_{t,t+1} &= \beta\left[\mathrm{E}_t \widetilde{J}_{t+1}^{1-\gamma}\Xi_{t+1}^{1-\gamma}\right]^{\frac{\gamma-\psi}{1-\gamma}}\frac{\widetilde{J}_t^{\psi}}{\widetilde{J}_{t+1}^{\gamma}\Xi_{t+1}^{\gamma}}\frac{\Lambda_{t+1}}{\Lambda_t};
\end{aligned}
$$

- firms

$$
\begin{aligned}
(\theta-1)\,\widetilde{Y}_t = -\zeta\left(\Pi_t - (\Pi^*)^{1-\iota}\,\Pi_{t-1}^{\iota}\right)\widetilde{Y}_t\Pi_t + \frac{\theta}{\alpha}\frac{\widetilde{w}_t}{P_t}\frac{1}{Z_t^{\frac{1}{\alpha}}}\widetilde{Y}_t^{\frac{1}{\alpha}} + \\
+\,\mathrm{E}_t Q_{t,t+1}\zeta\left(\Pi_{t+1} - (\Pi^*)^{1-\iota}\,\Pi_t^{\iota}\right)\widetilde{Y}_{t+1}\Xi_{t+1}\Pi_{t+1};
\end{aligned}
$$

- market clearing

$$
\begin{aligned}
\widetilde{Y}_t &= \widetilde{C}_t + \widetilde{G}_t + \frac{\zeta}{2}\left(\Pi_t - (\Pi^*)^{1-\iota}\,\Pi_{t-1}^{\iota}\right)^2\widetilde{Y}_t,\\
N_t &= \widetilde{Y}_t^{\frac{1}{\alpha}} Z_t^{-\frac{1}{\alpha}};
\end{aligned}
$$

- policy rule

$$
I_t = \left(\frac{\Pi^*\Xi_{t+1}^{\psi}}{\beta}\right)^{1-\rho_I}\left(\frac{\Pi_t}{\Pi_t^*}\right)^{\psi_\Pi}\left(\frac{\widetilde{Y}_t}{\widetilde{Y}}\right)^{\psi_Y} I_{t-1}^{\rho_I} e^{\varepsilon_{t+1}^{I}};
$$

- shocks

$$
\begin{aligned}
\Xi_t &= \overline{\Xi}^{1-\rho_\xi}\Xi_{t-1}^{\rho_\xi}e^{\varepsilon_t^{\xi}}, \qquad \varepsilon_{t+1}^{\xi} \sim N\left(0,\sigma_\xi\right),\\
\widetilde{G}_t &= \left(g\widetilde{Y}\right)^{1-\rho_g}\widetilde{G}_{t-1}^{\rho_g}e^{\varepsilon_t^{g}}, \qquad \varepsilon_{t+1}^{g} \sim N\left(0,\sigma_g\right),\\
\mu_{w,t+1} &= \mu_w^{1-\rho_\mu}\left(\mu_{w,t}\right)^{\rho_\mu}e^{\varepsilon_{t+1}^{\mu}}, \qquad \varepsilon_{t+1}^{\mu} \sim N\left(0,\sigma_\mu\right),\\
Z_t &= Z_{t-1}^{\rho_z}e^{\varepsilon_t^{z}}, \qquad \varepsilon_{t+1}^{z} \sim N\left(0,\sigma_{z,s_{z,t}}\right),\\
\eta_{t+1} &= e^{\varepsilon_{t+1}^{\eta}}, \qquad \varepsilon_{t+1}^{\eta} \sim N\left(0,\sigma_{\eta,s_{\eta,t}}\right);
\end{aligned}
$$

- standard deviations

$$
\begin{aligned}
\sigma_{z,s_{z,t}} &= \sigma_{z,0}s_{z,t} + \sigma_{z,1}\left(1-s_{z,t}\right),\\
\sigma_{\eta,s_{\eta,t}} &= \sigma_{\eta,0}s_{\eta,t} + \sigma_{\eta,1}\left(1-s_{\eta,t}\right);
\end{aligned}
$$

- $C_{-1}$, $I_{-1}$, $\Pi_{-1}$ given.

### A.7 Numerical implementation

For the numerical implementation of the model, we scale the maximum value function by a constant $\kappa$ to increase accuracy. Define a dummy variable $\widetilde{D}_t = \mathrm{E}_t \Xi_{t+1}^{1-\gamma} \widetilde{J}_{t+1}^{1-\gamma} / \kappa^{1-\gamma}$. It follows that $\kappa^{1-\gamma} \widetilde{D}_t = \mathrm{E}_t \Xi_{t+1}^{1-\gamma} \widetilde{J}_{t+1}^{1-\gamma}$. This implies

$$\widetilde{D}_t = \frac{\mathrm{E}_t \Xi_{t+1}^{1-\gamma} \widetilde{J}_{t+1}^{1-\gamma}}{\kappa^{1-\gamma}},$$

$$\widetilde{J}_t^{1-\psi} = (1-\beta)\, \widetilde{u}_t^{1-\psi} + \beta \kappa^{1-\psi} \widetilde{D}_t^{\frac{1-\psi}{1-\gamma}},$$

$$Q_{t,t+1} = \beta \left( \frac{\kappa \widetilde{D}_t^{\frac{1}{1-\gamma}}}{\widetilde{J}_{t+1}} \right)^{\gamma-\psi} \left( \frac{\widetilde{u}_{t+1}}{\widetilde{u}_t} \right)^{-\psi} \frac{u_{c,t+1}}{u_{c,t}} \frac{1}{\Xi_{t+1}^{\gamma}} \frac{1}{\Pi_{t+1}}.$$

### A.8 Functional forms

We rely on the Trabandt and Uhlig (2011) form for temporary utility, i.e.

$$u_t = (C_t - h \Xi_t C_{t-1}) \left( 1 - \eta (1-\psi) N_t^{1+\frac{1}{\phi}} \right)^{\frac{\psi}{1-\psi}}.$$

As a result

$$\frac{\widetilde{w}_t}{P_t} = \frac{\eta \psi \left(1 + \frac{1}{\phi}\right) \left( \widetilde{C}_t - h \widetilde{C}_{t-1} \right) N_t^{\frac{1}{\phi}}}{1 - \eta (1-\psi) N_t^{1+\frac{1}{\phi}}} \frac{\theta_{w,t}}{\theta_{w,t} - 1},$$

$$\widetilde{J}_t^{1-\psi} = (1-\beta) \left( \widetilde{C}_t - h \widetilde{C}_{t-1} \right)^{1-\psi} \left( 1 - \eta (1-\psi) N_t^{1+\frac{1}{\phi}} \right)^{\psi} + \beta \kappa^{1-\psi} \widetilde{D}_t^{\frac{1-\psi}{1-\gamma}},$$

$$Q_{t,t+1} = \beta \left( \frac{\kappa \widetilde{D}_t^{\frac{1}{1-\gamma}}}{\widetilde{J}_{t+1}} \right)^{\gamma-\psi} \left( \frac{\widetilde{C}_{t+1} - h \widetilde{C}_t}{\widetilde{C}_t - h \widetilde{C}_{t-1}} \right)^{-\psi} \left( \frac{1 - \eta (1-\psi) N_{t+1}^{1+\frac{1}{\phi}}}{1 - \eta (1-\psi) N_t^{1+\frac{1}{\phi}}} \right)^{\psi} \frac{1}{\Xi_{t+1}^{\gamma}} \frac{1}{\Pi_{t+1}},$$

$$\begin{aligned}(\theta - 1) \widetilde{Y}_t = &-\zeta \left( \Pi_t - (\Pi_t^*)^{1-\iota} \Pi_{t-1}^{\iota} \right) \widetilde{Y}_t \Pi_t + \frac{\theta}{\alpha} \frac{\widetilde{w}_t}{P_t} \left( \frac{\widetilde{Y}_t}{Z_t} \right)^{\frac{1}{\alpha}} + \ldots \\ &+ \mathrm{E}_t Q_{t,t+1} \zeta \left( \Pi_{t+1} - (\Pi_{t+1}^*)^{1-\iota} \Pi_t^{\iota} \right) \widetilde{Y}_{t+1} \Xi_{t+1} \Pi_{t+1}.\end{aligned}$$

### A.9 Elasticity of intertemporal substitution

We compute the elasticity of intertemporal substitution of consumption as the elasticity of consumption to a change in the real interest rate holding labour supply constant.

Define the "consumption surplus" $\overleftrightarrow{c}_t \equiv \widetilde{C}_t - h \widetilde{C}_{t-1}$. The first order approximation to the nominal stochastic discount factor

$$Q_{t,t+1} = \beta \left( \frac{\kappa \widetilde{D}_t^{\frac{1}{1-\gamma}}}{\widetilde{J}_{t+1}} \right)^{\gamma-\psi} \left( \frac{\overleftrightarrow{c}_{t+1}}{\overleftrightarrow{c}_t} \right)^{-\psi} \left( \frac{1 - \eta (1-\psi) N_{t+1}^{1+\frac{1}{\phi}}}{1 - \eta (1-\psi) N_t^{1+\frac{1}{\phi}}} \right)^{\psi} \frac{1}{\Xi_{t+1}^{\gamma}} \frac{1}{\Pi_{t+1}}$$

can be written as[1]

$$\widehat{q}_{t,t+1} = -\psi\Delta\widehat{\overleftrightarrow{c}}_{t+1} - \psi\left(1+\frac{1}{\phi}\right)\frac{\overline{n}}{1-\overline{n}}\Delta\widehat{N}_{t+1} - \psi\widehat{\xi}_{t+1} - \widehat{\pi}_{t+1} - (\gamma-\psi)\left(\widehat{\xi}_{t+1} + \widehat{\widehat{j}}_{t+1} - \mathrm{E}_t\left[\widehat{\xi}_{t+1} + \widehat{\widehat{j}}_{t+1}\right]\right),$$

where

$$\widehat{\widehat{j}}_t + \widehat{\xi}_t = \sum_{i=0}^{\infty}\left(\beta\Xi^{1-\psi}\right)^i \mathrm{E}_t\left[\widehat{\xi}_{t+i} + \left(1-\beta\Xi^{1-\psi}\right)\left(\widehat{\overleftrightarrow{c}}_{t+i} - \frac{\psi}{1-\psi}\left(1+\frac{1}{\phi}\right)\frac{\overline{n}}{1-\overline{n}}\widehat{N}_{t+i}\right)\right].$$

As a result,

$$\widehat{q}_{t,t+1} = -\psi\Delta\widehat{\overleftrightarrow{c}}_{t+1} - \psi\left(1+\frac{1}{\phi}\right)\frac{\overline{n}}{1-\overline{n}}\Delta\widehat{N}_{t+1} - \psi\widehat{\xi}_{t+1} - \widehat{\pi}_{t+1},$$

and the real rate is

$$\widehat{r}_t = \psi\mathrm{E}_t\Delta\widehat{\overleftrightarrow{c}}_{t+1} + \psi\left(1+\frac{1}{\phi}\right)\frac{\overline{n}}{1-\overline{n}}\mathrm{E}_t\Delta\widehat{N}_{t+1} + \psi\mathrm{E}_t\widehat{\xi}_{t+1}.$$

Rearranging terms

$$\widehat{\overleftrightarrow{c}}_t = -\frac{1}{\psi}\widehat{r}_t + \mathrm{E}_t\widehat{\overleftrightarrow{c}}_{t+1} + \frac{1}{\psi}\left(1+\frac{1}{\phi}\right)\frac{\overline{n}}{1-\overline{n}}\mathrm{E}_t\Delta\widehat{N}_{t+1} + \mathrm{E}_t\widehat{\xi}_{t+1},$$

so that the long-run elasticity of substitution $\overline{EIS}$, i.e. the elasticity which is obtained after households have adjusted their consumption habits, takes the usual value

$$\overline{EIS} = \frac{1}{\psi}.$$

Note that, in the absence of habits, this expression boils down to the usual value $1/\psi$.

To compute the short-run elasticity, we rewrite the consumption surplus in terms of the underlying consumption levels to obtain

$$\widehat{\widehat{c}}_t = -\frac{1}{\psi}\frac{1-h}{1+h}\widehat{r}_t + \frac{1}{1+h}\mathrm{E}_t\widehat{\widehat{c}}_{t+1} + \frac{h}{1+h}\widehat{\widehat{c}}_{t-1} + \frac{1-h}{1+h}\left(1+\frac{1}{\phi}\right)\frac{\overline{n}}{1-\overline{n}}\mathrm{E}_t\Delta\widehat{N}_{t+1} + \frac{1-h}{1+h}\mathrm{E}_t\widehat{\xi}_{t+1}.$$

The short-run elasticity of substitution $EIS$ is therefore

$$EIS = \frac{1}{\psi}\frac{1-h}{1+h},$$

which again boils down to $1/\psi$ when $h=0$. Note that, since $h>0$, it is always the case that $EIS < \overline{EIS}$.

# B Model estimation

## B.1 Approximate likelihood

Solving the model to second order, we obtain the reduced-form system of equations

$$y^o_{t+1} = k_{y,j} + F\hat{x}_{t+1} + \frac{1}{2}\left(I_{n_y}\otimes\hat{x}'_{t+1}\right)E\hat{x}_{t+1} + Dv_{t+1}, \tag{1}$$

$$\hat{x}_{t+1} = k_{x,i} + P\hat{x}_t + \frac{1}{2}\left(I_{n_x}\otimes\hat{x}'_t\right)G\hat{x}_t + \widetilde{\sigma}\Sigma_i w_{t+1}, \tag{2}$$

$$s_t \backsim MS(Q), \tag{3}$$

[1] In these derivations, $\kappa = 1$.

where

$$
\begin{aligned}
k_{y,j} &= k_{y,s_{t+1}=j},\\
k_{x,i} &= k_{x,s_t=i},\\
\Sigma_i &= \Sigma(s_t = i).
\end{aligned}
$$

and $Q$ is the transition probability matrix associated with the Markov switching (MS) process $s_t$.

The vector $y^o_t$ includes all observable variables, and $v_{t+1}$ and $w_{t+1}$ are measurement and structural shocks, respectively. In this representation, as shown in Amisano and Tristani (2011), the regime switching variables affect the system by changing the intercepts $k_{y,j}$, $k_{x,i}$ and the loadings of the structural innovations $\Sigma_i$ (we indicate here with $i$ the value of the discrete state variables at $t$ and with $j$ the value of the discrete state variables at $t+1$).

To compute the approximate likelihood, at any point in time we first linearize the two quadratic terms around the conditional mean of the continuous state variables conditional on the prevailing regime. As a result, the two equations above can be rewritten as

$$
\begin{aligned}
y^o_{t+1} &= \widetilde{k}^{(i,j)}_{y,t+1} + \widetilde{F}^{(i,j)}_{t+1}\hat{x}_{t+1} + Dv_{t+1},\\
\widehat{x}_{t+1} &= \widetilde{k}^{(i)}_{x,t} + \widetilde{P}^{(i)}_t \widehat{x}_t + \Sigma_i w_{t+1},
\end{aligned}
$$

where

$$
\begin{aligned}
\widetilde{k}^{(i,j)}_{y,t+1} &= \widetilde{k}_{y,j} + \frac{1}{2}\left(I_{n_y} \otimes \hat{x}^{(i)\prime}_{t+1|t}\right) E\hat{x}^{(i)}_{t+1|t} - \Delta_{i,t+1}\hat{x}^{(i)}_{t+1|t},\\
\widetilde{F}^{(i,j)}_{t+1} &= F + \Delta_{i,t+1}\hat{x}^{(i)}_{t+1|t} = E(x_{t+1}|\underline{y}^o_{1:t}, s_t = i, \theta),\\
\Delta_{i,t+1} &= \left[\frac{\partial\left(\frac{1}{2}\left(I_{n_y} \otimes \hat{x}'_{t+1}\right) E\hat{x}_{t+1}\right)}{\partial \hat{x}_{t+1}}\right]_{\hat{x}_{t+1}=\hat{x}^{(i)}_{t+1|t}},\\
\widetilde{k}^{(i)}_{x,t} &= \widetilde{k}_{x,i} + \frac{1}{2}\left(I_{n_x} \otimes \hat{x}^{(i)\prime}_{t|t}\right) G\hat{x}^{(i)}_{t|t} - \Delta_{i,t}\hat{x}^{(i)}_{t|t},\\
\widetilde{P}^{(i)}_t &= P + \Delta_{i,t}\hat{x}^{(i)}_{t|t} = E(\hat{x}_t|\underline{y}^o_{1:t}, s_t = i, \theta),\\
\boldsymbol{\Delta}_{i,t} &= \left[\frac{\partial\left(\frac{1}{2}\left(I_{n_x} \otimes \hat{x}'_t\right) G\hat{x}_t\right)}{\partial \hat{x}_t}\right]_{\hat{x}_t=\hat{x}^{(i)}_{t|t}}
\end{aligned}
$$

for regime-dependent intercepts $\widetilde{k}^{(i,j)}_{y,t+1}$, $\widetilde{k}^{(i)}_{x,t}$ and slope coefficients $\widetilde{F}^{(i,j)}_{t+1}$, $\widetilde{P}^{(i)}_t$.We then apply Kim's (1994) approximate filter to forecast

$$
\begin{aligned}
\hat{x}^{(i,j)}_{t+1|t} &= \widetilde{k}^{(i)}_{x,t} + \widetilde{P}^{(i)}_t \hat{x}^{(i)}_{t|t} = \hat{x}^{(i)}_{t+1|t},\\
Q^{(i,j)}_{t+1|t} &= \widetilde{P}^{(i)}_t Q^{(i,j)}_{t|t} \widetilde{P}^{(i)\prime}_t + \Sigma_i\Sigma_i' = Q^{(i)}_{t+1|t},
\end{aligned}
$$

and update the vector of continuous state variables

$$
\hat{x}^{(j)}_{t+1|t+1} = \sum_{i=1}^{m} \hat{x}^{(i,j)}_{t+1|t+1} \times p(s_t = i|s_{t+1} = j, \underline{y}_{1:t+1}),
$$

$$Q^{(j)}_{t+1|t+1} = \sum_{i=1}^{m} \left[ \left( \hat{x}^{(i,j)}_{t+1|t+1} - \hat{x}^{(j)}_{t+1|t+1} \right) \left( \hat{x}^{(i,j)}_{t+1|t+1} - \hat{x}^{(j)}_{t+1|t+1} \right)' + Q^{(i,j)}_{t+1|t+1} \right] \times$$
$$\times p(s_t = i | s_{t+1} = j, \underline{y}_{1:t+1}),$$

and then update the regime probabilities

$$p(s_{t+1} = j, s_t = i | \underline{y}_{1:t}) = p_{ij,t+1|t} = p_{ij} \times p(s_t = i | \underline{y}_{1:t}),$$

and

$$p(s_{t+1} = j, s_t = i | \underline{y}_{t+1}) = p_{ij,t+1|t} \times \frac{p(y_{t+1} | \underline{y}_t, s_{t+1} = j, s_t = i)}{p(y_{t+1} | \underline{y}_t)},$$

$$p(s_{t+1} = j | \underline{y}_{1:t+1}) = \sum_{i=1}^{m} p(s_{t+1} = j, s_t = i | \underline{y}_{1:t+1}),$$

$$p(s_t = i | s_{t+1} = j, \underline{y}_{1:t+1}) = \frac{p(s_{t+1} = j, s_t = i | \underline{y}_{1:t+1})}{p(s_{t+1} = j | \underline{y}_{1:t+1})},$$

$$p(y_{t+1} | \underline{y}_{1:t}) = \sum_{i=1}^{m} \sum_{j=1}^{m} p(y_{t+1} | \underline{y}_{1:t}, s_{t+1} = j, s_t = i) \times p(s_{t+1} = j, s_t = i | \underline{y}_{1:t}).$$

The conditional log-likelihood is obtained as

$$\log L = \sum_{t=1}^{T} \log p(y_{t+1} | y_{1:t}).$$

## B.2 MCMC simulation

We start by computing the mode of the posterior distribution of the parameters by using a two step approach:

1. we compute a reasonable approximation to the mode by using a simulated annealing algorithm (Goffe, Ferrier and Rogers, 1994);
2. using the result from the first step as initial value, we then run a gradient based method (C. Sims's `csminwell`) to find the posterior mode.

Having found the posterior mode, we compute the Hessian of the log posterior distribution at the mode and we use minus the inverse of this matrix as covariance matrix for a Gaussian distribution in a random walk Metropolis-Hastings algorithm, as customarily done in Bayesian estimation of DSGE models (as described in An and Schorfheide, 2007). This covariance matrix is scaled to achieve acceptance rates of 50%.

The MCMC algorithm is run to obtain 300,000 draws, the first 100,000 are discarded and the remaining ones are thinned (i.e. one every ten draws is recorded), resulting in a final posterior sample of 20,000 draws, which is then used in all the computations reported in the paper.

We find that the resulting posterior sample has good properties in terms of acceptance rate and low correlation across draws.

### B.3 Unconditional moments

To compute first and second order moments, we apply a pruning approach, i.e. we take into consideration only linear terms for the computation of second order moments, and linear and quadratic terms for the computation of first order moments. The computation of unconditional moments works as follows: for each draw of the parameter vector from the posterior distribution, we compute the state space representation (1), (2) and (3). From the state space representation we obtain the unconditional covariance matrix of state vector shocks as

$$\Omega_{ww} = \widetilde{\sigma}\sum_{i=1}^{m}\Sigma_i\Sigma_i^{'}\pi_i,$$

where $\pi_i$ are the ergodic state probabilities associated with the transition probability matrix $Q$. Taking the state equation stripped of its second order term, we can obtain $\Omega_{xx,0}$,the static covariance matrix of $\widehat{x}_t$, as solution of

$$Cov(\widehat{x}_t) = \Omega_{xx,0} = P\Omega_{xx,0}P^{'} + \Omega_{ww}.$$

Dynamic covariance matrices are obtained by applying the recursion

$$Cov(\widehat{x}_t, \widehat{x}_{t-j}) = \Omega_{xx,j} = P\Omega_{xx,j-1}, j = 1, 2, 3, ...$$

and the covariance matrices for the variables in $y_t^o$ are obtained using the corresponding linear measurement equation:

$$\begin{aligned}\Omega_{yy,0} &= Cov(y_t^o) = F\Omega_{xx,0}F^{'} + DD^{'},\\ \Omega_{yy,j} &= Cov(y_t^o, y_{t-j}^o) = F\Omega_{xx,j}F^{'}, j = 1, 2, 3, ...\end{aligned}$$

When computing first moments, we take into consideration both first and second order. To show how first moments are obtained, we re-write the state space representation in equivalent form as

$$y_{t+1}^o = k_{y,j} + F\hat{x}_{t+1} + \frac{1}{2}\overline{E}vec\left(\hat{x}_{t+1}\hat{x}_{t+1}^{'}\right) + Dv_{t+1}, \tag{4}$$

$$\hat{x}_{t+1} = k_{x,i} + P\hat{x}_t + \frac{1}{2}\overline{G}vec\left(\hat{x}_t\hat{x}_t^{'}\right) + \widetilde{\sigma}\Sigma_i w_{t+1}, \tag{5}$$

where $\overline{E}$ and $\overline{G}$ are obtained by suitably re-arranging the elements of the matrices $E$ and $G$, respectively. Taking the unconditional expected value of the two expressions above yields the first moments:

$$\begin{aligned}\mu_y &= E(y_t^o) = k_y + F\mu_x + \frac{1}{2}\overline{E}vec\left(\Omega_{xx,0}\right), k_y = \sum_{j=1}^{m}k_{y,j}\pi_j,\\ \mu_x &= E(\widehat{x}_t) = [I_{n_x} - P]^{-1}[k_x + \frac{1}{2}\overline{G}vec\left(\Omega_{xx,0}\right)], k_x = \sum_{i=1}^{m}k_{x,i}\pi_i.\end{aligned}$$

### B.4 Impulse response functions

To compute impulse response functions (IRFs) we follow Koop, Pesaran and Potter (1996). IRFs can be computed with respect to all shocks hitting the model, either continuous (the shocks in the state vector $w_t$) or discrete, i.e. the shocks that lead to a change in the discrete Markov switching process that affects the model. We define $\varepsilon_t$ as the vector containing all the shocks affecting continuous and discrete states. We compute IRFs to a shock $\varepsilon_{jt}$ of size $\delta_j$ occurring at time $t$,using the following algorithm:

- draw $\theta^{(i)}, i = 1, 2, ..., M$, from the posterior distribution of the parameters;
- compute the state space representation corresponding to $\theta^{(i)}$, run the Kalman filter and draw $\widehat{x}_t^{(i)}, s_t^{(i)}$ from their joint posterior distribution conditional on $\theta^{(i)}$;
- draw two histories of shocks $\varepsilon_{t+h}^{(i,1)} \varepsilon_{t+h}^{(i,2)}, h = 0, 1, 2, ..., H$, which are totally identical but differ only for the shock $\varepsilon_{jt}$, such that

$$\varepsilon_{jt}^{(i,2)} = \varepsilon_{jt}^{(i,1)} + \delta_j;$$

- feed these two histories of shocks to state and measurement equations starting from $\widehat{x}_t^{(i)}, s_t^{(i)}$, and generate 2 paths

$$y_{t+h}^{(i,1)}, y_{t+h}^{(i,2)}, h = 0, 1, 2, ..., H;$$

  the difference between these two path traces the dynamic response of shock $\delta_j$;
- the empirical distribution of this difference across draws $\theta^{(i)}$ gives the posterior distribution of the IRFs.

Note that IRFs reported in Figure (7) in the paper are obtained by fixing the state $s_{t+h}, h = 0, 1, 2.., H$ at the value corresponding to low volatility for all the shocks. IRFs reported in Figure (5) are obtained by contemplating a one-off shift in volatility, i.e. forcing the process to move to the high volatility state only once at time $t$.

### B.5 Variance decomposition

Forecast Error Variance decomposition (FEVD) is a measure of the importance of the model's orthogonal shocks in determining the observed behavior of each variable in the model at different horizons. The procedure to compute variance decomposition is straightforward in linear models and a bit more complicated in non linear ones, such as the quadratic MS-DSGE model used in our paper. In particular, difficulties arise since:

1. the model is non-linear;
2. there are shocks in the variances, i.e. discrete shocks, beside the usual continuous shocks;
3. there is uncertainty around the latent states, even conditioning on parameter values.

It is important to notice though, that the non-linearities generated by quadratic terms in the model's solution do not play any role if second order moments are computed using an appropriate pruning procedure, i.e. taking into consideration only linear terms. In order to describe how the variance decomposition results contained in the paper are computed, we define

$$v^{(i)}(j,h,\{S\}) \;=\; V(y_{it+h}|y_{1:t},\theta^{(i)}_{\{S\}})$$

the conditional variance of $y_{it+h}$ conditional on the parameter vector $\theta^{(i)}$ drawn from the joint posterior distribution, and $\{S\}$ denotes the set of shocks or sources of randomness being allowed to be active in the system from $t+1$ to $t+H$, the end of the projection period.

As an example, setting $S = \{0\}$ means that all all shocks in the system (continuous and discrete) are being switched off, and this is achieving by modifying from $\theta^{(i)}$ to $\theta^{(i)}_{\{0\}}$ setting all shocks standard deviations to zero. In this case the conditional variance is determined only by the conditional variance of all the latent variables (continuous and discrete) at time $t$, what is usually referred to as "initial condition". When we instead define $S$ to be the full set of shocks, we compute conditional variances using $\theta^{(i)}$. These variances are determined by the full structure of shocks in the model.

In order to describe the portion of forecast variances attributable to each shock, let us call $\varepsilon^k_t, k = 1,2,3,4,5$, the continuous shocks in the model, the first three of them having Markov switching variances.

FEVD coefficients are computed as follows:

- for each value of the parameters $\theta^{(i)}, i = 1,2,...,M$, drawn from the posterior distribution, we compute the solution for the model and the theoretical $h$-step ahead forecast variances of all observed series. This is done using the appropriate pruning, i.e. considering only linear terms. These conditional variances, generated when all shocks are active, enter in the denominator of the FEVD coefficients. This is the denominator of any FEVD coefficient and it is indicated as

$$v^{(i)}(j,h,\{all\}) \;=\; V(y_{it+h}|y_{1:t},\theta^{(i)}).$$

- Starting from $\theta^{(i)}$, we set the standard deviations of all shocks (and measurement errors) to zero, and we obtain $\theta^{(i)}_{\{0\}}$. We then compute the associated forecast variances. This is the portion of variances due to uncertainty around initial conditions:

$$v^{(i)}(j,h,\{0\}) \;=\; V(y_{it+h}|y_{1:t},\theta^{(i)}_{\{0\}}),$$

$$FEVD^{(i)}(j,h,\{0\}) \;=\; 100 \times \frac{v^{(i)}(j,h,\{0\})}{v^{(i)}(j,h,\{all\})}.$$

- Starting from $\theta^{(i)}_{\{0\}}$, for each of the continuous shocks with Markov switching variances ($\varepsilon^k_t$,$k = 1,2,3$),we first consider the contribution of the shock by setting its two variances both equal to its low volatility regime value, i.e. $\sigma_{1,k}$ $k = 1,2,3$ , therefore obtaining the vector $\theta^{(i)}_{\{0,k\}}$. In this way we introduce only the $k^{th}$ continuous shock, but we zero out its variance jumps. We indicate the corresponding variance as

$$v^{(i)}(j,h,\{0,k\}) \;=\; V(y_{it+h}|y_{1:t},\theta^{(i)}_{\{0,k\}})$$

and we isolate the contribution of that shock by netting out the effect of the initial condition as follows

$$v^{(i)}(j,h,\{k\}) \;=\; v^{(i)}(j,h,\{0,k\}) - v^{(i)}(j,h,\{k\})$$

and

$$FEVD^{(i)}(j,h,\{k\}) \quad = \quad 100 \times \frac{v^{(i)}(j,h,\{k\})}{v^{(i)}(j,h,\{all\})}.$$

- For each of the shocks with switching variances, we define $\theta^{(i)}_{\{0,k,s_k\}}$ the modification of $\theta^{(i)}_{\{0,k\}}$ where the two variances of shock $\varepsilon^k_t$ are set to their respective high and low values. In this way that is allowed to be heteroskedastic. The corresponding conditional variances are then

$$v^{(i)}(j,h,\{0,k,s_k\}) \quad = \quad V(y_{it+h}|y_{1:t},\theta^{(i)}_{\{0,k,s_k\}}),$$

  and we isolate the contribution of the $k^{th}$ shock variance jumps by subtracting the portion of variance jointly due to the initial condition and to the $k^{th}$ shock when assumed to be homoskedastic:

$$v^{(i)}(j,h,\{s_k\}) \quad = \quad v^{(i)}(j,h,\{0,k,s_k\}) - v^{(i)}(j,h,\{0,k\});$$

  the FEVD of the $k^{th}$ shock Markov switching jumps is hence computed as follows

$$FEVD^{(i)}(j,h,\{s_k\}) \quad = \quad 100 \times \frac{v^{(i)}(j,h,\{s_k\})}{v^{(i)}(j,h,\{all\})}.$$

- For each of the 8 continuous shocks without Markov-switching variance, i.e. the mark-up shock $\varepsilon^{\mu}_t$, the permanent technology shock $\varepsilon^{\xi}_t$ and the 6 measurement errors ($l = 1, 2, ..., 8$), we measure the FEVD contribution by defining $\theta^{(i)}_{\{0,l\}}$, i.e. the parameter vector obtained by modifying $\theta^{(i)}_{\{0\}}$ to allow the standard deviation of the $l^{th}$ shock shock to be equal to the corresponding value of $\theta^{(i)}$. We then compute

$$v^{(i)}(j,h,\{0,l\}) \quad = \quad V(y_{it+h}|y_{1:t},\theta^{(i)}_{\{0,l\}})$$

  and we isolate the effect of shock $l$ by subtracting the effect of the initial condition as follows

$$v^{(i)}(j,h,\{l\}) \quad = \quad v^{(i)}(j,h,\{0,l\}) - v^{(i)}(j,h,\{0\})$$

  and the corresponding FEVD coefficients are

$$FEVD^{(i)}(j,h,\{l\}) \quad = 100\times \quad \frac{v^{(i)}(j,h,\{l\})}{v^{(i)}(j,h,\{all\})}.$$

- The FEVD coefficients describe above by construction sum to 1 across all sources of uncertainty for each variable (initial condition, continuous shocks, measurement errors and variance switches).

- These computations are repeated for all draws from the posterior distribution and results are averaged across draws. In Table (2) of the paper we report the posterior means of each FEVD coefficient (one for each variable and for each shock) and the corresponding 5% and 95% quantiles at different horizons.

## References


[1] An, S. and F. Schorfheide (2007): "Bayesian analysis of DSGE models", *Econometric Reviews*, 26, 223-172.

[2] Goffe, W., G. Ferrier and J. Rogers (1994): "Global optimization of statistical functions with simulated annealing", *Journal of Econometrics*, 60, 65-99.

[3] Koop, G., ,M. Pesaran and S. Potter (1996): "Impulse response analysis in nonlinear multivariate models", *Journal of Econometrics*, 74, 119-147.