## Appendix A: The roles of various model properties

The following subsections discuss details of the role of mean adjustment, regime change, and panel structure in ELB risk estimation. A summary of the appendix can be found in Subsection 5.1.

### A.1 The role of mean adjustment

The extensive literature dealing with equilibrium values of macroeconomic variables in advanced countries allows us to formulate informative priors on the steady-state parameters. Regarding the informativeness of the priors, two extreme cases can be distinguished. First, imposing a very tight prior is similar to the situation where the steady-state values are calibrated. Second, imposing a very loose prior resembles the case where no mean adjustment is conducted and allows us to discuss the role of mean adjustment in the estimation of ELB risk. Examining the two extreme cases can shed some light on, respectively, the effect of calibration and the effect of not imposing any equilibrium values on the ELB risk estimates in the previous literature.

**-- Insert Table A1 about here --**

Table A1 reports the medium-term probabilities of an ELB event for the benchmark case (a 95 percent confidence band for the prior mean on the steady-state parameters of width 2 for GDP growth, inflation, and the interest rate and of width 1.5 for the spread), the case with a very tight prior on the steady state (a 95 percent confidence interval of width 0.1 for all variables), and the case with a loose prior (width 10). The tight prior results in the same or a lower ELB risk. The difference is driven to a great extent by the posterior of the interest rate steady state. With the exception of Norway, the tight prior implies a higher steady-state interest rate (Table A2) and, ceteris paribus, a lower ELB risk. The comparison in Table A1 suggests that the calibration of steady states plays an important role in ELB risk estimation. For example, calibrating the interest rate steady state to 1 percent for Japan implies a reduction in the ELB risk of more than half.[16]

**-- Insert Table A2 about here --**

Comparing the estimates for the loose prior on the steady states with the benchmark setting demonstrates the role of mean adjustment. It turns out that without treating the steady state explicitly and imposing extra-data information on steady-state values, the estimated ELB risk is higher for all countries. For example, the estimate for the US suggests that it would be stuck at the ELB 11 percent of the time. With mean adjustment, the ELB situation is estimated to be observed only 6 percent of the time. In the short term, the effect of mean adjustment is negligible. Within a period of one year, the distance between the ELB event probabilities in the benchmark and the case without mean adjustment does not exceed 0.01 in absolute terms.

**-- Insert Table A3 about here --**

As noted above, the change in the ELB risk estimates for tight priors with respect to the benchmark is influenced by the change in the posterior of the interest rate steady state. In addition, part of the difference can be explained by the fact that the benchmark case allows some uncertainty of the steady state, which is not the case when the steady state is calibrated. The strength of the effect is demonstrated in Table A3. The role of uncertainty in the benchmark estimation is filtered out by employing the tight prior centered on the posterior mean of the benchmark specification. In other words, the steady state is calibrated to the benchmark posterior values and the uncertainty of the steady state is negligible. The table shows that the influence on the ELB risk is at most of order 0.01 in the direction of lower ELB risk if the uncertainty of the parameter is ignored.[17]

**A.2 The role of regime change**

In addition to the theoretical reasoning for allowing regime change due to a structural change in monetary policy conduct, the estimation results can justify the chosen modeling framework ex post. A natural question is whether there are changes in the estimates of the model parameters between regimes. Furthermore, one can ask whether the regime change is driven by a change in shock volatilities, a change in dynamic coefficients, or both.

**-- Insert Figure A1 about here --**

As an illustration, Figure A1 presents the posterior distributions of the parameters at the first lag of the endogenous variables in the equation for the interest rate for the US. The passive conventional monetary policy for regime 1 is manifested by posteriors centered on zero for all parameters at the first lag except for the interest rate. In regime 2, mainly positive numbers are covered by the posterior distributions of the parameters at the first lag of real growth and inflation, which is reminiscent of the standard backward-looking Taylor rule.

A systematic examination of the differences between the two regimes is reported in Figures A2 and A3, which show the dynamic coefficients and diagonal components of the error covariance matrices in both regimes. It turns out that the regime change is driven primarily by a change in shock volatilities, with the posterior distributions covering mostly different parameter values in the two regimes. The volatility of the shocks is lower in regime 1, when the interest rate is stuck at, or close to, the ELB, and the main GFC-related drop in real economy variables happened before 2012, i.e., in regime 2. Significant differences between the regimes can also be observed for some dynamic coefficients. Note that the differences reflect the fact that stability of VAR is imposed in regime 2 but not in regime 1.

**-- Insert Figure A2 about here --**

The estimated shock volatility in the interest rate equation for Japan is an order of magnitude lower than for the other countries. This fact helps explain why the medium-term ELB risk for Japan (0.16) is unrealistically low despite the realistic value of the steady-state interest rate (0.34) (see Table A2). The empirical model reflects the observed interest rate, not the desired interest rate and at the same time is not rich enough to capture all the unconventional measures employed in Japan. This is why such a model is not necessarily appropriate for Japan and provides unrealistic ELB risk estimates. Another issue related to Japan is the heterogeneity in the regime change date by comparison with the rest of the sample, as discussed in the next subsection.

**-- Insert Figure A3 about here --**

Allowing for regime change is one possible explanation for the different ELB risks estimated in the literature. For the US, the difference between estimates based on linear models (Kiley and Roberts, 2017, Hills et al., 2016, and Nakata, 2017b) and those based on nonlinear models (Gust et al., 2017, Richter and Throckmorton, 2016) exceeds 10 percentage points. Table A4 compares the medium-term ELB risk in the benchmark model and the linear model that does not allow for regime change.

Looking at the credible intervals indicated in parenthesis in Table A4, it can be concluded that the medium-term ELB risk is different for Japan and the US only. For the US, the difference goes in the opposite direction than suggested by a plain comparison of estimates found in the literature. The credible intervals overlap for the other countries, suggesting no significant effect of nonlinearity. So, the nonlinearity seems not to be the primary reason for the different ELB risk estimates as suggested by the literature.

**-- Insert Table A4 about here --**

The marginal likelihood of the data estimated for the two versions of the model implies that the data fit of the model without regime change is superior to that of the version with regime change (see the last row in Table A4). The observed differences in the dynamic parameters and volatility components between the regimes suggest that the lower data fit of the nonlinear model could be a consequence of the simple fact that common regime change does not explain the data well for some countries. Country-specific VARs are discussed in the next subsection.

**A.3 The role of panel structure**

Exploiting the panel structure of data with observed ELB spells allows for estimation of the whole set of model parameters despite the short time series available for a single country. The information pooling takes the form of the common mean of the dynamic coefficients across countries. In addition, taking into account the correlation of shocks across countries leads to more efficient estimates, which, in turn, results in more accurate simulation of the ELB risk due to the fact that the parameters' uncertainty enters the simulation procedure. On the other hand, if the heterogeneity between countries is substantial and the interdependence is weak, the panel approach may not be preferable.

**-- Insert Figure A4 about here --**

The degree of commonality of the macroeconomic dynamics across countries in the two regimes is driven by the overall tightness parameter $\lambda^{(r)}$. As shown in Figure A4, the posterior mean of the parameter is higher in regime 1 than in regime 2. A higher $\lambda^{(r)}$ suggests more divergent coefficients across countries or more tightly estimated country coefficients. There are 20 observations for each country in regime 1 and 49 in regime 2 and, as suggested by the posteriors presented in Figure A2 the coefficients in regime 1 are not more tightly estimated. Therefore, the differences in dynamics across countries in regime 1 are probably substantial. This should not come as a surprise, because regime 1 contains different unconventional monetary policy measures employed during the Great Recession.

The posterior mean of $\lambda^{(1)}$ is 0.14 and that of $\lambda^{(2)}$ is 0.05. Following the discussion in Jarocinski (2010), the square roots of the posterior means are comparable to the usual overall tightness used when setting the Minnesota prior. In our case, the square roots are 0.38 and 0.22, close to the interval of 0.1–0.2 covering the usual values used for overall tightness. This gives some confidence that the corresponding prior mean provides valuable information. The estimated $\lambda^{(r)}$ suggests that some pooling of information across countries is present in both regimes and justifies the use of a common mean when the country-specific dynamic parameters are estimated.

The parameter $\lambda^{(r)}$, however, does not capture the possible heterogeneity in the timing of the regime change, which is assumed to be the same for all countries. If there are significant differences in the regime change timing across countries, the ELB measures could be inaccurate. To set up the model, we face a trade-off between the possibility of pooling information across countries in the model with regime change and inaccuracy due to country differences in the regime change date.

To filter out the heterogeneity in the regime change date in the examination of the role of panel structure in ELB risk estimation, two specifications without regime change are estimated: panel VAR and single-country VARs.[18] Table A5 suggests that the effect of common mean estimation and the improvement in efficiency due to static interdependencies is substantial. The specification that does not use any information from other countries in the estimation yields a medium-term

ELB risk at least twice as high as the panel data specification. The only exception is Japan, which exhibits no significant increase in terms of overlapping 95 percent credible intervals.

In specifications with regime change, the role of information pooling can be expected to be even more profound, because the time series entering the estimation are divided into two series and thus are even shorter than in the estimation without regime change.

**-- Insert Table A5 about here --**

To shed some additional light on the issue of country heterogeneity with respect to regime change, single-country mean-adjusted VARs with regime change can be estimated.[19] It turns out that the regime change dates differ across countries. The most divergent country in this sense is Japan. Regime 2 (normal times) covers 37 quarters, while in the panel VAR the period is 49 quarters.[20] The second country which differs substantially from the rest is Norway, with regime 2 covering 39 quarters. For the other countries, regime 2 consists of at least 45 quarters. Note that if the regime change date differs, its interpretation can change as well, and the assumptions related to the estimation should be revised (for example, the assumption of mean adjustment in one regime only and the assumption of stability of VAR in a regime).

**Appendix B: Bayesian estimation**

The posterior distributions of the model parameters are computed employing the Bayes formula. The prior distribution for $\beta^{(1)}$ and $\beta^{(2)}$ is formulated conditional on hyperparameters $b^{(1)}$, $b^{(2)}$, $\Sigma_c^{b,(1)}$, and $\Sigma_c^{b,(2)}$ for $c = 1,...,N$ and the priors on the hyperparameters enter the formula according to the definition of the conditional probabilities:

$$\begin{aligned}&\pi\left(\beta^{(1)},\beta^{(2)},b^{(1)},b^{(2)},\Sigma_c^{b,(1)},\Sigma_c^{b,(2)},F,\overline{\Sigma}^{(1)},\overline{\Sigma}^{(2)},r \mid \overline{y}\right) \propto \pi\left(\overline{y} \mid \beta^{(1)},\beta^{(2)},F,\overline{\Sigma}^{(1)},\overline{\Sigma}^{(2)},r\right)* \\ &*\pi\left(\beta^{(1)} \mid b^{(1)},\Sigma_c^{b,(1)}\right)\pi\left(\beta^{(2)} \mid b^{(2)},\Sigma_c^{b,(2)}\right)\pi\left(b^{(1)}\right)\pi\left(b^{(2)}\right)\prod_{c=1}^{N}\left[\pi\left(\Sigma_c^{b,(1)}\right)\pi\left(\Sigma_c^{b,(2)}\right)\right]\pi\left(\overline{\Sigma}^{(1)}\right)\pi\left(\overline{\Sigma}^{(2)}\right)\pi(F)\pi(r)\end{aligned} \quad \text{(B1)}$$

In the following subsections, the particular ingredients of the Bayes rule are discussed in turn. Subsection B.4 presents the sampler and Subsection B.5 the marginal likelihood computation.

**B.1 Likelihood**

The likelihood of the model (4) is as follows:

$$\begin{aligned}&\pi\left(\overline{y} \mid \beta^{(1)},\beta^{(2)},F,\overline{\Sigma}^{(1)},\overline{\Sigma}^{(2)},r\right) = \left|\overline{\Sigma}^{(1)}\right|^{-\frac{1}{2}}\left|\overline{\Sigma}^{(2)}\right|^{-\frac{1}{2}} \\ &\exp\left\{-\frac{1}{2}\sum_{r=1}^{2}\left(\overline{y}^{(r)} - \overline{X}^{(r)}\beta^{(r)} - I_t(r=2)\overline{Z}^{(2)}\delta\right)'\left(\overline{\Sigma}^{(r)}\right)^{-1}\left(\overline{y}^{(r)} - \overline{X}^{(r)}\beta^{(r)} - I_t(r=2)\overline{Z}^{(2)}\delta\right)\right\}\end{aligned} \quad \text{(B2)}$$

where the indicator of regime 2, $I(r=2)$, equals one if the system is in regime 2 at time *t* and zero otherwise.

**B.2 Priors**

From (5) it follows that for country *c* and regime *r*, the vector $\beta_c^{(r)}$, made up of parameters relating to the lagged vectors of the endogenous variables (and of a constant term in the case of regime 1), is distributed normally with mean $b^{(r)}$ and country-specific variance $\Sigma_c^{b,(r)}$. The common mean error covariance matrix $\Sigma_c^{b,(r)}$ is treated as a proportion of the $(n^2p+n)\times(n^2p+n)$ diagonal matrix $\Omega_c^{b,(1)}$ and the $(n^2p)\times(n^2p)$ diagonal matrix $\Omega_c^{b,(2)}$, which are defined in the manner of the Minnesota prior: the variance of the parameter in the *i*-th equation at the lagged values of the *j*-th variable is given by:

$$\frac{\sigma_{c,i}^2}{\sigma_{c,j}^2},$$

where $\sigma_{c,i}$ and $\sigma_{c,j}$ are estimated standard errors from univariate AR(*p*) models for the corresponding endogenous variables and serve as scaling parameters to account for the different sizes of the parameters. The standard errors are estimated on the whole sample and are the same for both $\Omega_c^{b,(r)}$. In addition, the matrix $\Omega_c^{b,(1)}$ includes the variance for the intercept, which is set to $10^2\sigma_{c,i}^2$ for the *i*-th equation.

Given $\Omega_c^{b,(r)}$, the common mean covariance matrix is then defined as:

$$\Sigma_c^{b,(1)} = \left(\lambda_1^{(1)} \otimes I_{n^2p+n}\right)\Omega_c^{b,(1)} \text{ and } \Sigma_c^{b,(2)} = \left(\lambda_1^{(2)} \otimes I_{n^2p}\right)\Omega_c^{b,(2)}. \qquad \text{(B3)}$$

The regime-specific parameter $\lambda_1^{(r)}$ plays a role in the prior's overall tightness and its posterior distribution drives the extent to which the vector $\beta_c^{(r)}$ is allowed to change across countries. Moreover, the parameter $\lambda_1^{(r)}$ is the only parameter in $\Sigma_c^{b,(r)}$ that is not fixed and is estimated. The prior on $\lambda_1^{(r)}$ is assumed to be inverse-Gamma distributed:

$$\lambda_1^{(r)} \sim IG\left(\frac{s_0}{2}, \frac{v_0}{2}\right),$$

with $s_0 = 0.001$ and $v_0 = 0.001$.

The prior on the common mean $b^{(r)}$ is assumed to have the standard form of the Minnesota prior. The prior is normal centered around the AR(1) process:

$$b^{(r)} \sim N\left(B^{(r)}, \Xi^{b,(r)}\right),$$

with the prior mean $B^{(r)}$ made up of zeros, except for the coefficient at the first own lag of the variables in levels (interest rate, spread), which is 0.9.[21] The prior variance $\Xi^{b,(r)}$ is set using matrices $\Omega_c^{b,(r)}$, which define the variance of country-specific $\beta_c^{(r)}$ around the common mean. We define

$$\Xi^{b,(r)} = \left(\lambda_1^b \otimes I_{n^2 p}\right)\frac{1}{N}\sum_{c=1}^{N}\Omega_c^{b,(r)} .$$

The overall tightness parameter $\lambda_1^b$ is set to the standard value of 0.01.

The prior on the error covariance matrix $\Sigma^{(r)}$ is assumed be inverse-Wishart:[22]

$$\Sigma^{(r)} \sim W^{-1}\left(0.01 I_{Nn}, Nn+1\right).$$

For regime 2, the prior on the coefficients capturing the steady state for country *c*, $F_c$, is distributed normally with mean $\psi_{c,0}$ and variance $\Lambda_{c,0}$. The parameters of the prior distribution are given by their 95 percent confidence bands for $F_c$ and are reported in Table B1.

**-- Insert Table B1 about here --**

The confidence bands are set to imply prior means that coincide with the long-run equilibrium values available in the literature. Holston et al. (2017) provide estimates of real equilibrium GDP growth and the real equilibrium interest rate for Canada, the euro area, the UK, and the US. Other

European countries are assumed to have the same equilibrium growth and real interest rate as the euro area. Equilibrium growth and real interest rate estimates for Japan can be found in Fujiwara (2016). The inflation steady state is assumed to follow the country's definition of price stability, i.e., except for Norway and Japan a price growth target of 2 percent is assumed. Norway's inflation target is 2.5 percent. Japan introduced an explicit CPI inflation target in January 2013 with two-year time span to achieve it. We set the prior on Japanese steady-state price growth at 1 percent. For the steady state of the nominal interest rate we add the inflation steady state to the equilibrium real rate. The confidence bands for the steady-state interest rate spread are based on the average spread over the period 1999Q1–2016Q4 and the width of the band is assumed to be 1.5.

Finally, the prior distribution assumed for the threshold parameter $r$ is uniform on the interval $[r_{\min}, r_{\max}]$, where the bounds of the interval are defined such that at least 20 observations remain in a regime if the threshold takes an extreme value.

### B.3 Conditional posteriors

Drawing on formula (B1) and ignoring terms that do not involve the parameter of interest, we can derive conditional posterior distributions. To a great extent, the specifications of the conditional posteriors follow the formulas in Dieppe et al. (2016), Villani (2009), Chen and Lee (1995), and Koop and Potter (2003) and derivations are not presented. The main difference from the above-mentioned papers with respect to the derivations is that static interdependencies are allowed for by dealing with data for all countries at once when the vectors of parameters relating to the lagged values of endogenous variables and the error covariance matrix are drawn. Note that mean adjustment is applied to regime 2 only.

The conditional posterior of the common mean $b^{(r)}$ is distributed normally with mean

$$\left(\left(\frac{1}{N^2}\sum_{c=1}^{N}\Sigma_c^{b,(r)}\right)^{-1}+\left(\Xi^{b,(r)}\right)^{-1}\right)^{-1}\left(\left(\frac{1}{N^2}\sum_{c=1}^{N}\Sigma_c^{b,(r)}\right)^{-1}\left(\frac{1}{N}\sum_{c=1}^{N}\beta_c^{(r)}\right)+\left(\Xi^{b,(r)}\right)^{-1}B^{(r)}\right) \tag{B4}$$

and variance

$$\left(\left(\frac{1}{N^2}\sum_{c=1}^{N}\Sigma_c^{b,(r)}\right)^{-1}+\left(\Xi^{b,(r)}\right)^{-1}\right)^{-1}. \tag{B5}$$

The mean of the conditional posterior for $b^{(r)}$ is thus a weighted average of the prior mean $B^{(r)}$ and the average vector of dynamic coefficients across countries, with weights given by the reciprocal of the variances of the two.

Next, the covariance matrix that drives the dispersion of the country-specific vectors of the dynamic coefficients around the common mean, $\Sigma_c^{b,(r)}$, is defined in (B3). The conditional posterior of $\lambda_1^{(r)}$ is inverse-Gamma distributed with the shape parameter

$$\frac{1}{2}\left(h^{(r)}+s_0\right), \text{ where } h^{(1)}=Nn(np+1) \text{ and } h^{(2)}=Nn^2p, \tag{B6}$$

and the scale parameter

$$\frac{1}{2}\left(v_0+\sum_{c=1}^{N}\left\{\left(\beta_c^{(r)}-b^{(r)}\right)'\left(\Omega_c^{b,(r)}\right)^{-1}\left(\beta_c^{(r)}-b^{(r)}\right)\right\}\right). \tag{B7}$$

The conditional posterior of the vector of dynamic coefficients $\beta_c^{(r)}$ is distributed normally with mean

$$\left[\bar{X}^{(r)\prime}\left(\bar{\Sigma}^{(r)}\right)^{-1}\bar{X}^{(r)}+\left(\bar{\Sigma}^{b,(r)}\right)^{-1}\right]^{-1}\left[\bar{X}^{(r)\prime}\left(\bar{\Sigma}^{(r)}\right)^{-1}\bar{X}^{(r)}\left(\bar{X}^{(r)\prime}\bar{X}^{(r)}\right)^{-1}\bar{X}^{(r)\prime}\bar{y}^{(r)}+\left(\bar{\Sigma}^{b,(r)}\right)^{-1}b^{(r)}\right] \quad \text{(B8)}$$

and variance

$$\left[\bar{X}^{(r)\prime}\left(\bar{\Sigma}^{(r)}\right)^{-1}\bar{X}^{(r)}+\left(\bar{\Sigma}^{b,(r)}\right)^{-1}\right]^{-1}, \quad \text{(B9)}$$

where $\bar{\Sigma}^{b,(1)}$ is defined as follows

$$\bar{\Sigma}^{b,(r)}=\begin{pmatrix}\bar{\Sigma}_1^{b,(r)} & 0 & \cdots & 0\\ 0 & \bar{\Sigma}_2^{b,(r)} & \cdots & \vdots\\ \vdots & \vdots & \ddots & 0\\ 0 & \cdots & 0 & \bar{\Sigma}_N^{b,(r)}\end{pmatrix}.$$

Data matrices $\bar{X}^{(r)}$ and vectors $\bar{y}^{(r)}$, $r=1,2$, are defined in (5). Importantly, $\bar{X}^{(2)}$ and $\bar{y}^{(2)}$ are demeaned, i.e., the vector of steady states $F_c$ is subtracted from each $y_{c,t}$ when the data vectors and matrices are constructed.

The mean of the conditional posterior (B8) is a weighted average of the common mean $b^{(r)}$ and the maximum likelihood estimator $\left(\bar{X}^{(r)\prime}\bar{X}^{(r)}\right)^{-1}\bar{X}^{(r)\prime}\bar{y}^{(r)}$, with the weights being the reciprocal of the variances of the two.

The conditional posterior of the error covariance matrix $\Sigma^{(r)}$ is inverse-Wishart distributed with scale parameter

$$0.01 I_{Nn} + \left( y^{(r)} - X^{(r)} \underset{c=1,\dots,N}{blkdiag}\left(\beta_c^{(r)}\right)\right)' * \left( y^{(r)} - X^{(r)} \underset{c=1,\dots,N}{blkdiag}\left(\beta_c^{(r)}\right)\right), \tag{B10}$$

and degrees of freedom

$$Nn + 1 + t^{(r)}, \tag{B11}$$

where $t^{(r)}$ denotes the number of observations in regime $r$. The data matrix $X^{(r)}$ is defined as an $Nt^{(r)} \times Nnp$ block diagonal matrix with $X_c^{(r)}$, $c = 1,\dots,N$, on its diagonal, where only observations from regime $r$ are taken. Similarly, $y^{(r)}$ is an $Nt^{(r)} \times Nn$ block diagonal matrix with $y_c^{(r)}$ on its diagonal. The data vectors and matrices for regime 2 are demeaned by the steady-state vector $F_c$.

Next, the conditional posterior of the vector $F_c$ of steady states in regime 2 is distributed normally with mean

$$\begin{aligned} &\left[\Lambda_{c,0}^{-1} + U_c'\left(Z^{(2)'} Z^{(2)} \otimes \left(\Sigma_c^{(r)}\right)^{-1}\right) U_c\right] * \\ &\left[\Lambda_{c,0}^{-1}\psi_{c,0} + U_c' vec\left(\left(\Sigma_c^{(2)}\right)^{-1}\left(y_c^{(2)} - X_c^{(2)}\left[A_c^{1,(2)},\dots,A_c^{p,(2)}\right]'\right)' Z^{(2)}\right)\right] \end{aligned} \tag{B12}$$

and variance

$$\left[\Lambda_{c,0}^{-1}+U_c'\left(Z^{(2)'}Z^{(2)}\otimes\left(\Sigma_c^{(2)}\right)^{-1}\right)U_c\right]^{-1}, \tag{B13}$$

where

$$U_c=\begin{pmatrix} I_n \\ A_c^{1,(2)} \\ \vdots \\ A_c^{p,(2)} \end{pmatrix}.$$

Data matrices $X_c^{(2)}$ and $y_c^{(2)}$ are not demeaned. Note that draws of steady-state vectors are carried out country-by-country, i.e., the covariance structure relating countries is not exploited. In the case of steady states, this simplifying assumption is probably reasonable.

Finally, regarding the threshold parameter $r$, the conditional posterior is not available in analytical form. We therefore employ a Metropolis step based on the conditional posterior probability of the threshold:

$$\begin{aligned} &p\left(r\mid\beta^{(1)},\beta^{(2)},\overline{\Sigma}^{(1)},\overline{\Sigma}^{(2)},F,\overline{y},\overline{X}\right)\propto \\ &\left|\overline{\Sigma}^{(1)}\right|^{-\frac{t^{(1)}}{2}}\left|\overline{\Sigma}^{(2)}\right|^{-\frac{t^{(2)}}{2}}\exp\left\{-\frac{1}{2}tr\left[\sum_{r=1}^{2}\left(\overline{y}^{(r)}-\overline{X}^{(r)}\beta^{(r)}\right)'\left(\overline{\Sigma}^{(r)}\right)^{-1}\left(\overline{y}^{(r)}-\overline{X}^{(r)}\beta^{(r)}\right)\right]\right\}. \end{aligned} \tag{B14}$$

### B.4 The Gibbs sampler with a Metropolis step

The sample from the joint posterior of the model parameters is obtained by sampling from the conditional posteriors in the following steps:

1) Initial values: the country-specific vectors of dynamic parameters $\beta_c^{(r)}$ are initialized at their respective OLS estimates, the common mean $b^{(r)}$ at the prior mean, the overall tightness for the common mean error covariance matrix $\lambda^{(r)}$ at the value of 0.01, and the error covariance matrix $\Sigma^{(r)}$ at the OLS estimate based on pooled data for all countries in a regime given by the initial draw of the threshold parameter. The threshold is initialized at a random draw from the uniform distribution on the interval $[r_{\min}, r_{\max}]$, the vector of steady states $F$ at the respective prior means.

2) Given $\beta_c^{(r)}$, $c = 1, \ldots, N$, and $\Sigma_c^{b,(r)}$ from the previous iteration, the conditional posterior of the common mean $b^{(r)}$ is normally distributed with the mean given by (B4) and the variance given by (B5). For a given draw of the common mean $b^{(2)}$ a check is done of whether the eigenvalues of VAR with $b^{(2)}$ in the companion form are less than or equal to one. If they are not, another draw is taken. The maximum number of tries to obtain a stable VAR with $b^{(2)}$ is set to 20.

3) Given $\beta_c^{(r)}$, $c = 1, \ldots, N$, and $b^{(r)}$ from the previous iteration, the covariance matrix $\Sigma_c^{b,(r)}$ equals $(\lambda_1^{(1)} \otimes I_{n^2p+n})\Omega_c^{b,(1)}$ or $(\lambda_1^{(2)} \otimes I_{n^2pn})\Omega_c^{b,(2)}$, where the overall tightness parameter $\lambda_1^{(r)}$ is inverse-gamma distributed with the shape parameter defined in (B6) and the scale parameter defined in (B7).

4) Given $\Sigma_c^{b,(r)}$, $b^{(r)}$, $\overline{\Sigma}^{(r)}$, and $F$ from the previous iteration, the country-specific vectors of the dynamic parameters for regime $r$, $\beta_c^{(r)}$, are multivariate-normal distributed with mean (B8) and variance (B9). Similarly to $b^{(2)}$ in the first step, a stability check of the VAR structure for the drawn $\beta_c^{(2)}$ is conducted. The draws are taken until a stable VAR is

implied, the maximum number of tries being 200. The total number of tries is driven by the aim to get the ratio of unstable draws below 1 percent of the total draws.

5) Given $\Sigma_c^{b,(r)}$, $b^{(r)}$, $\beta_c^{(r)}$, and $F_c$ from the previous iteration, the error covariance matrix is inverse-Wishart distributed with scale parameter (B10) and degrees of freedom (B11).

6) Given $\Sigma_c^{b,(r)}$, $b^{(r)}$, and $\Sigma^{(r)}$ from the previous iteration, the vector $F_c$ is distributed normally with mean (B12) and variance (B13).

7) Metropolis step:

   a. Draw a proposed value $r^*$ from the prior distribution for the parameter.

   b. Compare the log of the conditional probability (B14) of the proposed value with the log of the conditional probability (B14) for the original value $r$ from the previous iteration.

   c. Accept the proposed value with a probability of $\min\{1, p(r^* \mid \ldots)/p(r \mid \ldots)\}$, where $p(\cdot \mid \ldots)$ is defined in (B14) (i.e., if the difference between the two logs of the conditional probabilities is larger than the logarithm of a draw from a standard uniform).

8) Repeat steps 2–7 until convergence is achieved.

## B.5 Marginal likelihood

The marginal likelihood of the model given the data is computed according to Chib (1995) and Chib and Jeliazkov (2001) adjusted for the existence of hyperparameters. The posterior distribution of the model parameters $\Theta \equiv \left\{b^{(1)}, b^{(2)}, \Sigma_c^{b,(1)}, \Sigma_c^{b,(2)}, \beta^{(1)}, \beta^{(2)}, \bar{\Sigma}^{(1)}, \bar{\Sigma}^{(2)}, F, r\right\}$ is given according to the Bayes rule:

$$p(\Theta|\bar{y}) = \frac{Lik(\bar{y}|\Theta)p(\Theta)}{MLik}. \qquad \text{(B15)}$$

The log of the marginal likelihood is then computed as follows:

$$ln(MLik) = ln\big(Lik(\bar{y}|\Theta)\big) + ln\big(p(\Theta)\big) - ln\big(p(\Theta|\bar{y})\big). \qquad \text{(B16)}$$

Formula (B16) is evaluated at a high-density point under the posterior. Here, the posterior mean of the parameter vector, $\Theta^*$, is taken. For the model formulated in (4), the terms in (B16) are computed as follows:

1) The log-likelihood of the data at $\Theta^*$is given by the log of formula (B2).

2) The log of the joint prior at $\Theta^*$:

$$ln\big(p(\Theta^*)\big) = \sum_{r,c}\left\{ln\left(p\left(\beta_c^{(r),*}|b^{(r),*},\lambda_r^*\right)\right)\right\} + \sum_r\left\{ln\left(p\left(b^{(r),*}\right)\right) + ln\big(p(\lambda_r^*)\big)\right\} + \sum_r ln\left(p\left(\bar{\Sigma}^{(r),*}\right)\right) + \sum_c ln\big(p(F^{c,*})\big) + ln\big(p(r^*)\big), \qquad \text{(B17)}$$

where the prior distributions are described in Subsection B.2.

3) The log of the posterior density of the parameters at $\Theta^*$:

$$ln\big(p(\Theta^*|\bar{y})\big) = \sum_r ln\left(p\left(\beta^{(r),*}|b^{(r),*},\lambda_r^*,\bar{\Sigma}^{(r),*},F^*,r^*,\bar{y}\right)\right) + \sum_r ln\left(p\left(b^{(r),*}|\lambda_r^*,\bar{\Sigma}^{(r),*},F^*,r^*,\bar{y}\right)\right) + \sum_r ln\left(p\left(\lambda_r^*|\bar{\Sigma}^{(r),*},F^{c,*},r^*,\bar{y}\right)\right) + \sum_r ln\left(p\left(\bar{\Sigma}^{(r),*}|F^{c,*},r^*,\bar{y}\right)\right) + ln\big(p(F^*|r^*,\bar{y})\big) + ln\big(p(r^*|\bar{y})\big). \qquad \text{(B18)}$$

The first RHS term represents the sum of the full conditional posterior density ordinates of the AR parameters, which are distributed normally with parameters defined in (B8) and (B9).

The term $p\left(b^{(r),*}|\lambda_r^*, \bar{\Sigma}^{(r),*}, F^{c,*}, r^*, \bar{\mathrm{y}}\right)$ is approximated by the sum:

$$\frac{1}{K}\sum_k\left\{p\left(b^{(r),*}|\lambda_r^*, \bar{\Sigma}^{(r),*}, F^{c,*}, r^*, \beta^{(r),(k)}, \bar{\mathrm{y}}\right)\right\},$$

where a reduced conditional density of the common mean in a regime is generated by a separate reduced Gibbs MCMC run produced for a fixed $\lambda_1^*, \bar{\Sigma}^{(r),*}, F^{c,*}$, and $r^*$.

Similarly, the term $p\left(\lambda_r^*|\bar{\Sigma}^{(r),*}, F^{c,*}, r^*, \bar{\mathrm{y}}\right)$ can be approximated by:

$$\frac{1}{K}\sum_k\left\{p\left(\lambda_r^*|\bar{\Sigma}^{(r),*}, F^{c,*}, r^*, \beta^{(r),(k)}, b^{(r),k}, \bar{\mathrm{y}}\right)\right\},$$

the term $p\left(\bar{\Sigma}^{(r),*}|F^{c,*}, r^*, \bar{\mathrm{y}}\right)$ by

$$\frac{1}{K}\sum_k\left\{p\left(\bar{\Sigma}^{(r),*}|F^{c,*}, r^*, \beta^{(r),(k)}, b^{(r),k}, \lambda_1^{(k)}, \bar{\mathrm{y}}\right)\right\},$$

and the term $p(F^*|r^*, \bar{\mathrm{y}})$ by

$$\frac{1}{K}\sum_k\left\{p\left(F^{c,*}|r^*, \beta^{(r),(k)}, b^{(r),k}, \lambda_1^{(k)}, \bar{\Sigma}^{(r),(k)}, \bar{\mathrm{y}}\right)\right\}.$$

Finally, the term $p(r^*|\bar{\mathrm{y}})$ is approximated using the draws from the original chain of the Gibbs sampler. The approximation draws on Chib and Jeliazkov (2001) with the following ratio:

$$\frac{\sum_k \alpha\left(r^{(k)}, r^*|\beta^{(r),(k)}, b^{(r),k}, \lambda_1^{(k)}, \bar{\Sigma}^{(r),(k)}, F^{c,(k)}, \bar{y}\right)p(r^*)}{\sum_k \alpha\left(r^*, r^{(k)}|\beta^{(r),(k)}, b^{(r),k}, \lambda_1^{(k)}, \bar{\Sigma}^{(r),(k)}, F^{c,(k)}, \bar{y}\right)},$$

where

$$\alpha\left(r, r' | \beta^{(r),(k)}, b^{(r),k}, \lambda_1^{(k)}, \bar{\Sigma}^{(r),(k)}, F^{c,(k)}, \bar{y}\right) = min\left\{1, \frac{p\left(\bar{y}|\beta^{(r),(k)}, b^{(r),k}, \lambda_1^{(k)}, \bar{\Sigma}^{(r),(k)}, F^{c,(k)}, r'\right)}{p\left(\bar{y}|\beta^{(r),(k)}, b^{(r),k}, \lambda_1^{(k)}, \bar{\Sigma}^{(r),(k)}, F^{c,(k)}, r\right)}\right\}.$$

The marginal likelihood is estimated with 1,000 draws, where each draw is taken after a burn-in period of 5,000. For other versions of the model (for example, panel VAR without regime change), the marginal likelihood is computed analogously.

The marginal likelihood for the benchmark panel VAR with different regime break dates is reported in Figure B1. The highest value is attained for the break date with regime 1 containing 20 observations.

**-- Insert Figure B1 about here –**

**Appendix C: Post-estimation diagnostics**

Two measures of convergence of the Gibbs part of the sampler are used: autocorrelation of the draws at a distance of 10, and the Raftery and Lewis (1992) estimate of the number of draws from the conditional posteriors needed to obtain a stationary distribution.[23] For the Metropolis step, the usual acceptance ratio – the ratio of the number of draws accepted to the total number of attempts – is reported.

For parameters $\lambda^{(1)}$, $\lambda^{(2)}$, and $F_c$, the autocorrelation of the draws is less than 0.02 in absolute terms and the number of suggested runs is less than 1,000. Similar results are obtained for the rest of the parameters, as reported in Figures C1–C3 for a selection of the parameter set.

**-- Insert Figure C1 about here --**

**-- Insert Figure C2 about here --**

**-- Insert Figure C3 about here --**

Finally, the acceptance ratio for the threshold parameter $r$ is 0.24. The posterior distribution of the threshold parameter is presented in Figure C4. The posterior distributions of the steady-state parameters are shown in Figure C5.

**-- Insert Figure C4 about here --**

**-- Insert Figure C5 about here --**

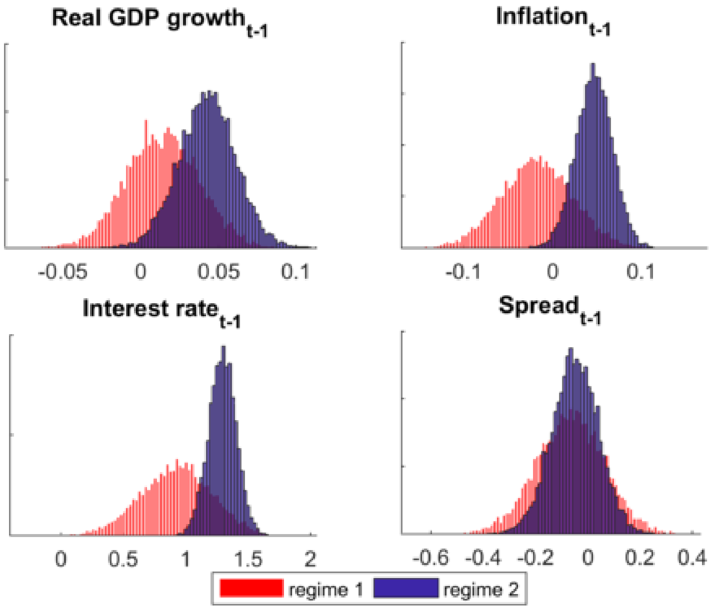


Figure A1: The posterior distribution of the parameters at the first lags of the endogenous variable in the interest rate equation for the US.

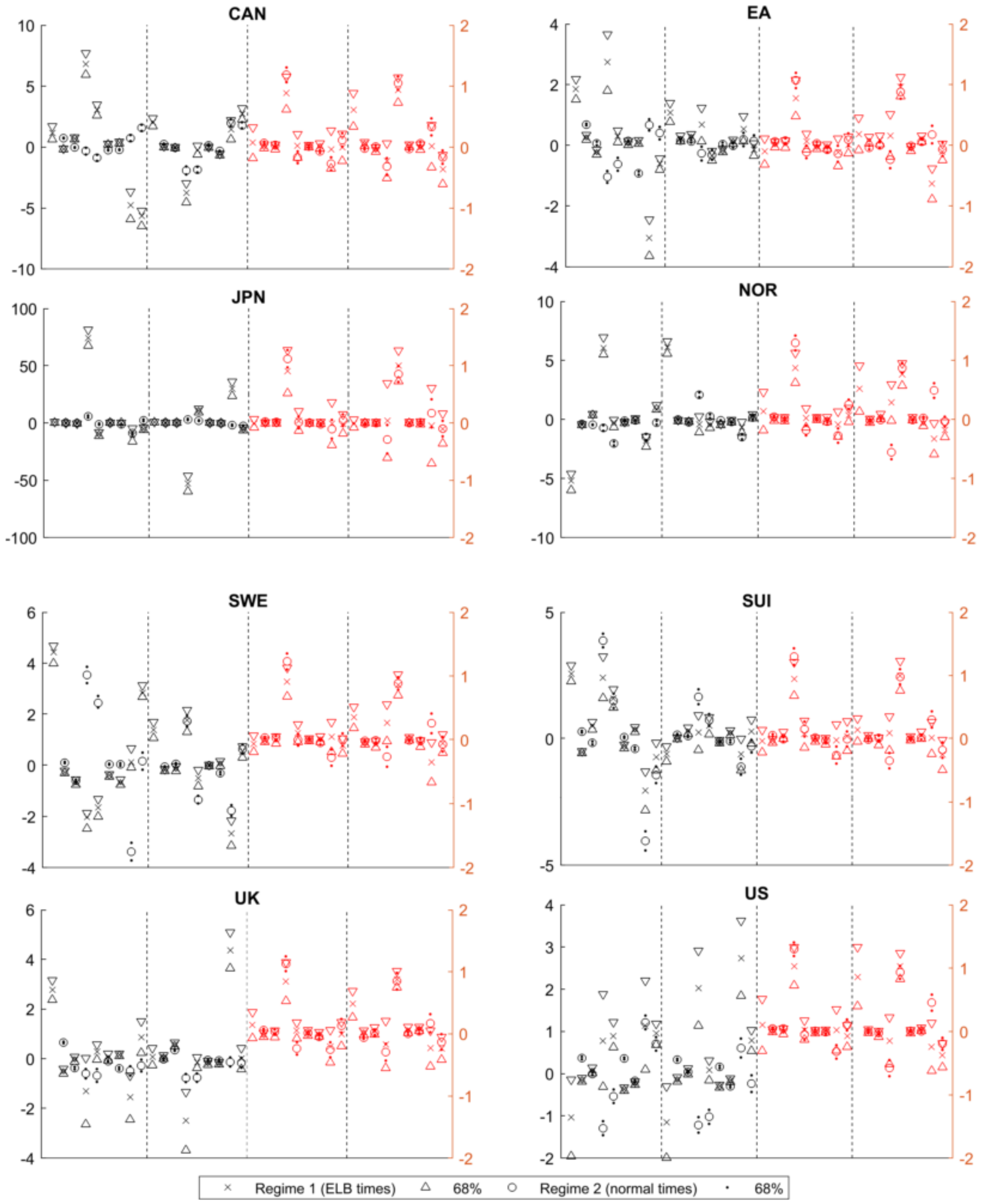


Figure A2: The dynamic coefficients and the respective 68 percent credible intervals in the two regimes.

Note: The four areas indicated by the dashed vertical lines refer to the four equations in the model (the equations for real GDP growth, inflation, the interest rate, and the spread). The third and fourth areas indicated in red relate to the right-hand y-axis. The first coefficient in each area is reported for regime 1 only because the intercept is not part of the dynamic coefficients vector in regime 2.

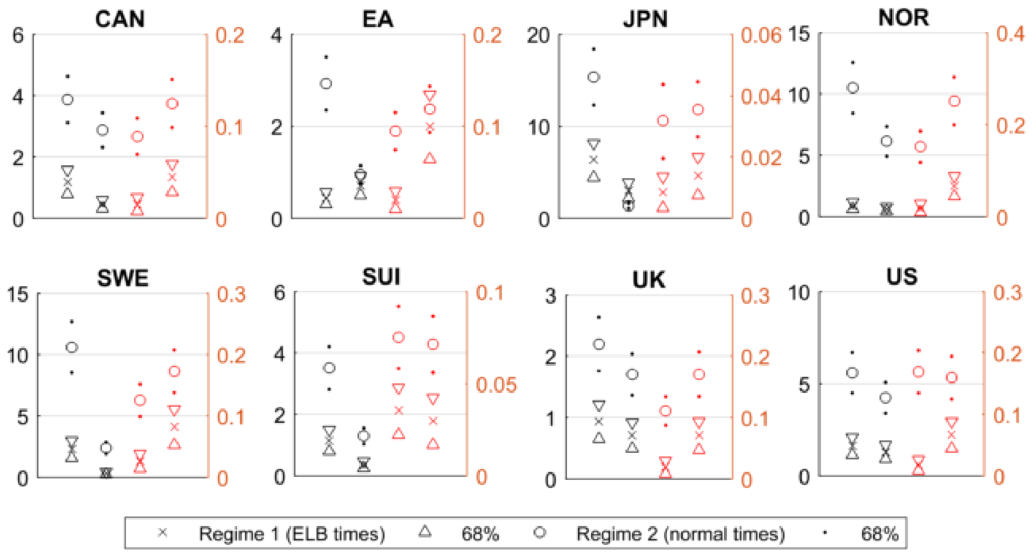


Figure A3: The diagonal elements of the error covariance matrices and the respective 68 percent credible intervals.

Note: The first two elements in each panel relate to the left-hand y-axis and the last two elements to the right-hand y-axis as indicated by different colors. The order of the LHS variables in the model is as follows: real GDP growth, inflation, the interest rate, and the spread.

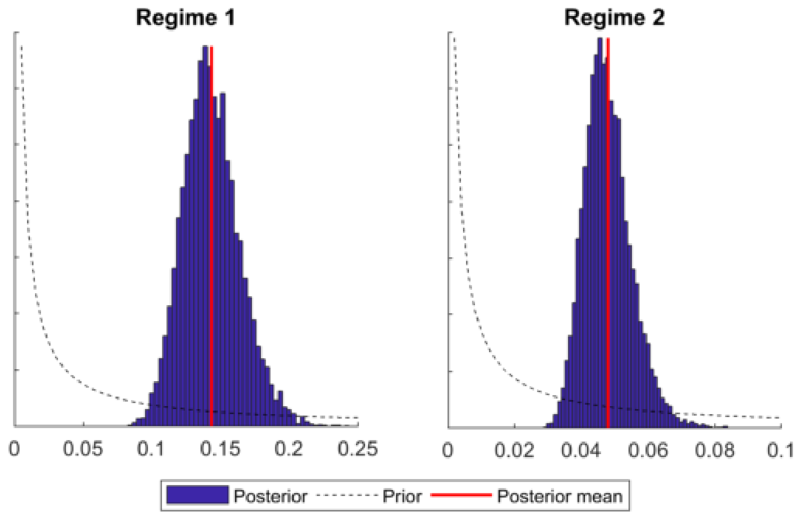


Figure A4: The posterior distribution of parameter $\lambda$ in the two regimes.

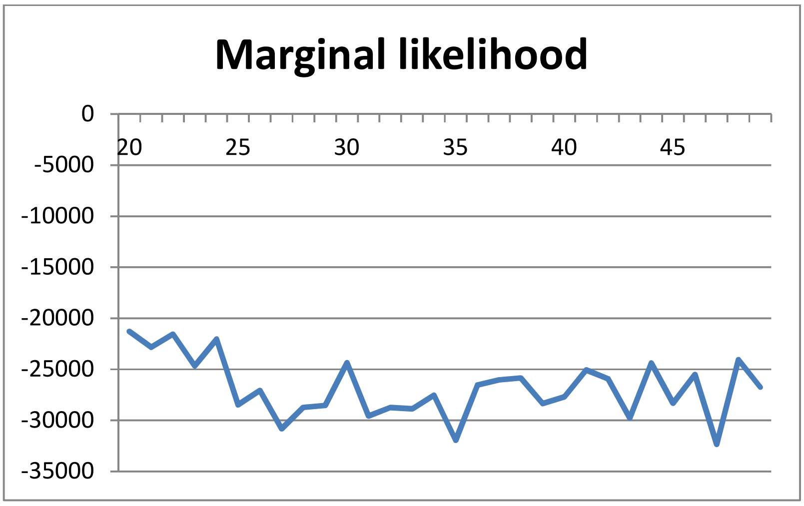


Figure B1: The marginal likelihood of the benchmark model.

Note: The horizontal axis denotes the number of observations in regime 1.

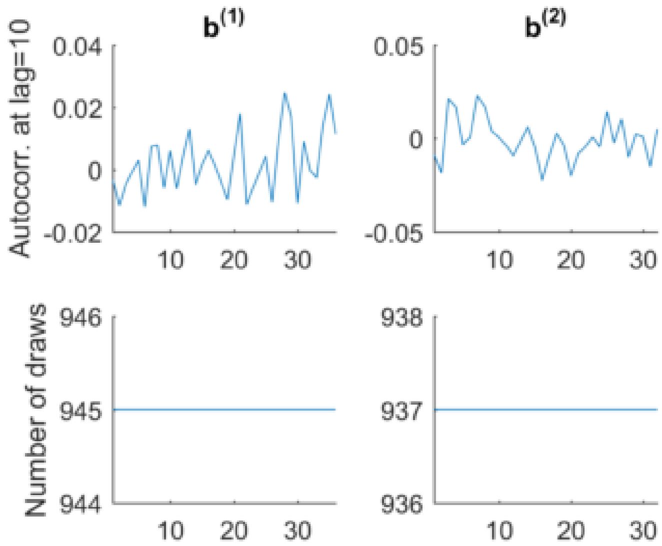


Figure C1: Convergence diagnostics for the common mean $b^{(r)}$.

Note: The parameters are stacked on the x-axis (36 parameters in regime 1 and 32 parameters in regime 2).

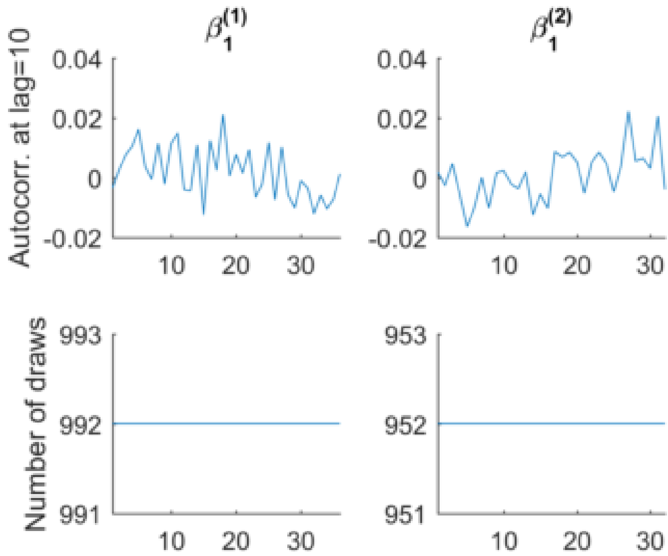


Figure C2: Convergence diagnostics for the vector of dynamic coefficients for Canada $\beta_1^{(r)}$.

Note: The parameters are stacked on the x-axis (36 parameters in regime 1 and 32 in regime 2).

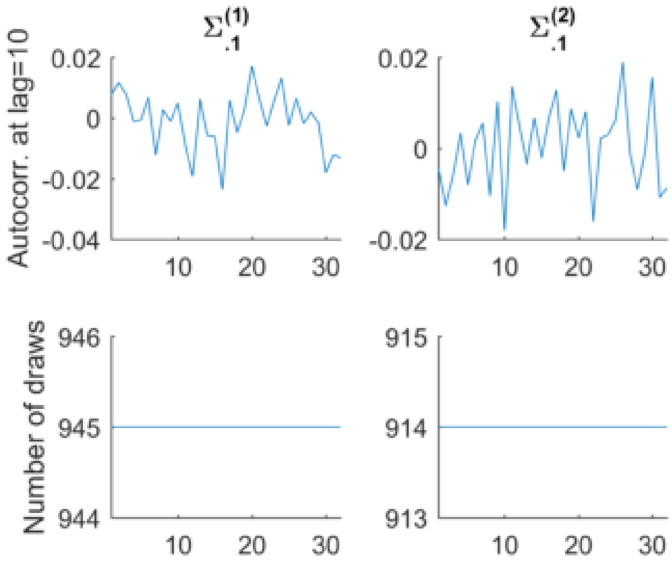


Figure C3: Convergence diagnostics for the first column of the error covariance matrix $\Sigma^{(r)}$.

Note: The parameters are stacked on the x-axis (32 parameters in each regime).

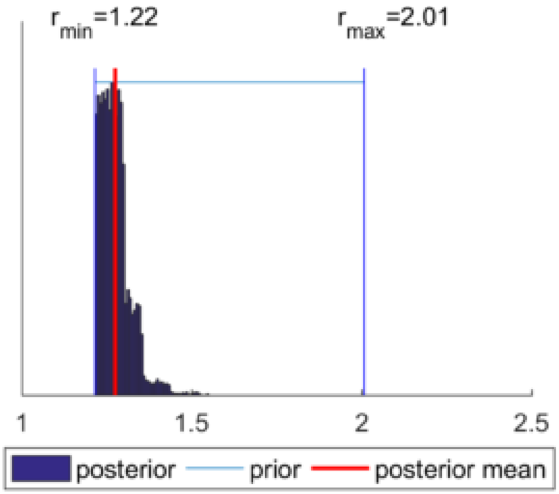


Figure C4: The posterior distribution of the threshold parameter.

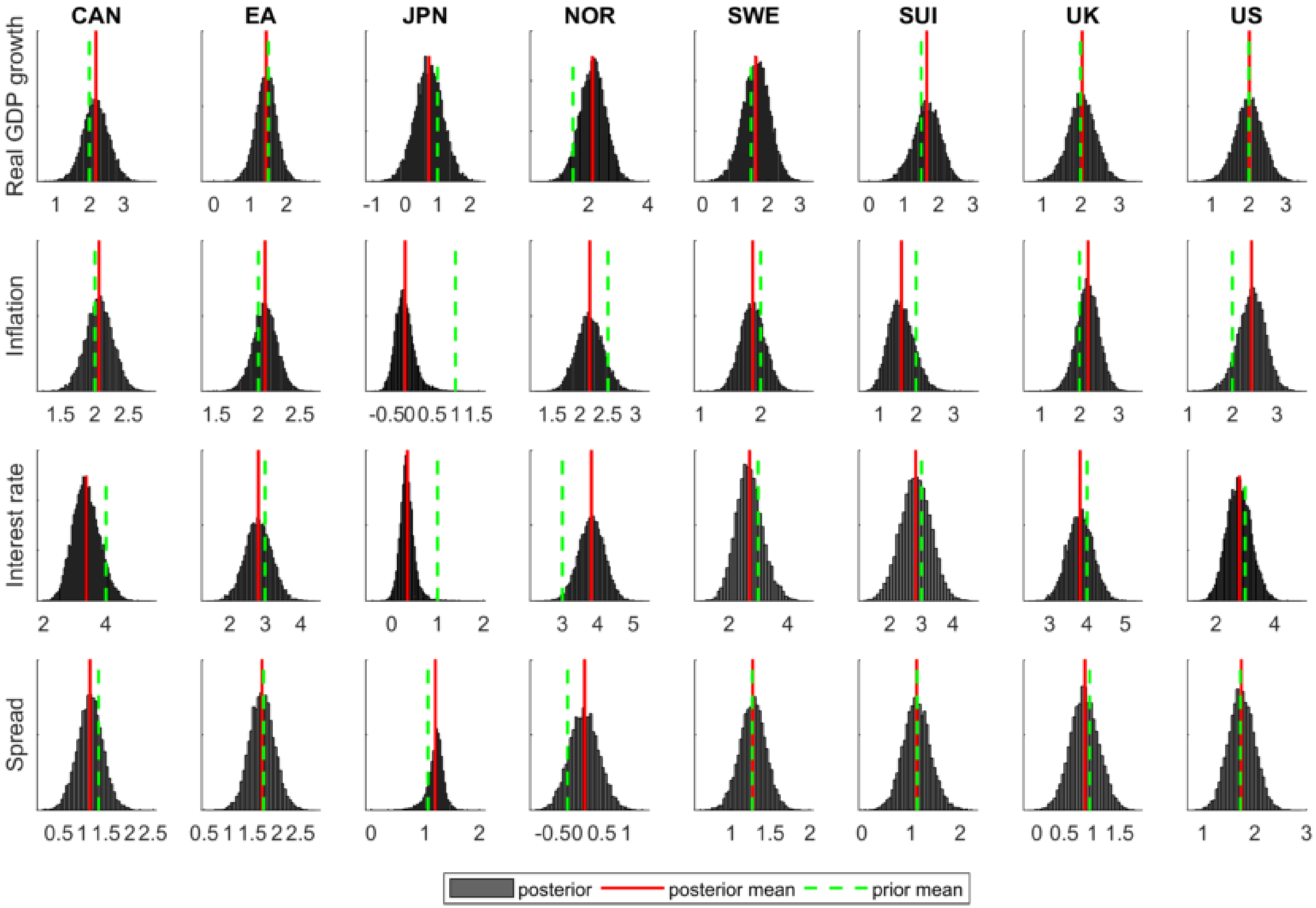


Figure C5: Posterior distributions of the steady-state parameters.

**Table A2: Prior mean and posterior means of the steady-state interest rate.**

| | Prior | | Posterior | |
|---|---|---|---|---|
| | | Benchmark | Tight | Loose |
| Canada | 4.00 | 3.38 | 4.00 | 2.21 |
| | (3.00, 5.00) | (2.57, 4.26) | (3.95, 4.05) | (-0.23, 4.57) |
| Euro area | 3.00 | 2.81 | 3.00 | 2.46 |
| | (2.00, 4.00) | (2.08, 3.53) | (2.95, 3.05) | (0.21, 4.49) |
| Japan | 1.00 | 0.34 | 1.00 | 0.25 |
| | (0.00, 2.00) | (0.05, 0.71) | (0.95, 1.05) | (-0.55, 1.02) |
| Norway | 3.00 | 3.83 | 3.00 | 4.02 |
| | (2.00, 4.00) | (3.14, 4.52) | (2.95, 3.05) | (1.81, 6.10) |
| Sweden | 3.00 | 2.70 | 3.00 | 1.61 |
| | (2.00, 4.00) | (1.84, 3.63) | (2.95, 3.05) | (-1.37, 4.70) |
| Switzerland | 3.00 | 2.82 | 3.00 | 1.50 |
| | (2.00, 4.00) | (1.89, 3.75) | (2.95, 3.05) | (0.16, 3.31) |
| United Kingdom | 4.00 | 3.81 | 3.99 | 3.15 |
| | (3.00, 5.00) | (3.16, 4.48) | (3.94, 4.04) | (0.79, 5.36) |
| United States | 3.00 | 2.79 | 3.00 | 2.00 |
| | (2.00, 4.00) | (1.99, 3.62) | (2.95, 3.05) | (-0.77, 4.62) |

Note: 95 percent credible intervals reported in parenthesis.

**Table A1: ELB risk in the medium term for different priors on the steady-state parameters.**

| | Benchmark | Tight | Loose |
|---|---|---|---|
| Canada | 0.01 | 0.01 | 0.04 |
| | (0.011, 0.015) | (0.007, 0.013) | (0.033, 0.041) |
| Euro area | 0.03 | 0.02 | 0.08 |
| | (0.028, 0.035) | (0.015, 0.021) | (0.065, 0.085) |
| Japan | 0.16 | 0.07 | 0.18 |
| | (0.152, 0.166) | (0.061, 0.071) | (0.169, 0.183) |
| Norway | 0.03 | 0.04 | 0.05 |
| | (0.027 ,0.040) | (0.034, 0.043) | (0.041, 0.051) |
| Sweden | 0.02 | 0.01 | 0.06 |
| | (0.019 ,0.026) | (0.012, 0.020) | (0.053, 0.063) |
| Switzerland | 0.03 | 0.02 | 0.06 |
| | (0.030, 0.040) | (0.021, 0.033) | (0.055, 0.069) |
| United Kingdom | 0.03 | 0.02 | 0.06 |
| | (0.022, 0.028) | (0.015, 0.021) | (0.050, 0.061) |
| United States | 0.06 | 0.04 | 0.11 |
| | (0.055, 0.064) | (0.038, 0.047) | (0.100, 0.113) |

Note: 95 percent credible intervals reported in parenthesis. The tight prior is given by a 95 percent confidence band of width 0.1 and the loose prior by a 95 percent confidence band of width 10.

**Table A3: ELB risk in the medium term for the benchmark and the tight prior centered on the estimated steady state.**

| | Benchmark | Tight on SS |
|---|---|---|
| Canada | 0.01 | 0.01 |
| | (0.011, 0.015) | (0.008, 0.012) |
| Euro area | 0.03 | 0.03 |
| | (0.028, 0.035) | (0.023, 0.028) |
| Japan | 0.16 | 0.15 |
| | (0.152, 0.166) | (0.146, 0.161) |
| Norway | 0.03 | 0.02 |
| | (0.027, 0.040) | (0.021, 0.034) |
| Sweden | 0.02 | 0.02 |
| | (0.019, 0.026) | (0.015, 0.022) |
| Switzerland | 0.03 | 0.03 |
| | (0.030, 0.040) | (0.025, 0.036) |
| United Kingdom | 0.03 | 0.02 |
| | (0.022, 0.028) | (0.018, 0.023) |
| United States | 0.06 | 0.05 |
| | (0.055, 0.064) | (0.047, 0.056) |

Note: 95 percent credible intervals reported in parenthesis. “Tight on SS” means that the prior on the steady state is centered on the posterior mean of the benchmark case and the prior is tight (the width of the 95 percent confidence band is 0.01).

**Table A4: ELB risk in the medium term for the benchmark and the model without regime change.**

| | Regime change | No regime change |
|---|---|---|
| Canada | 0.01 | 0.01 |
| | (0.011, 0.015) | (0.007, 0.014) |
| Euro area | 0.03 | 0.04 |
| | (0.028, 0.035) | (0.032, 0.043) |
| Japan | 0.16 | 0.20 |
| | (0.152, 0.166) | (0.183, 0.209) |
| Norway | 0.03 | 0.03 |
| | (0.027, 0.040) | (0.025, 0.033) |
| Sweden | 0.02 | 0.02 |
| | (0.019, 0.026) | (0.013, 0.028) |
| Switzerland | 0.03 | 0.02 |
| | (0.030, 0.040) | (0.020, 0.038) |
| United Kingdom | 0.03 | 0.03 |
| | (0.022, 0.028) | (0.025, 0.036) |
| United States | 0.06 | 0.04 |
| | (0.055, 0.064) | (0.039, 0.048) |
| Marginal likelihood of the model: | | |
| | -21284 | -18485 |

Note: 95 percent credible intervals reported in parenthesis.

**Table A5: ELB risk in the medium term for the panel VAR model without regime change and single-country VAR without regime change.**

| | Panel VAR | Single-country VAR |
|---|---|---|
| Canada | 0.01 | 0.07 |
| | (0.007, 0.014) | (0.068, 0.077) |
| Euro area | 0.04 | 0.08 |
| | (0.032, 0.043) | (0.076, 0.086) |
| Japan | 0.20 | 0.22 |
| | (0.183, 0.209) | (0.203, 0.234) |
| Norway | 0.03 | 0.10 |
| | (0.025, 0.033) | (0.095, 0.109) |
| Sweden | 0.02 | 0.06 |
| | (0.013, 0.028) | (0.060, 0.069) |
| Switzerland | 0.02 | 0.07 |
| | (0.020, 0.038) | (0.064, 0.073) |
| United Kingdom | 0.03 | 0.08 |
| | (0.025, 0.036) | (0.073, 0.083) |
| United States | 0.04 | 0.12 |
| | (0.039, 0.048) | (0.112, 0.128) |

Note: 95 percent credible intervals reported in parenthesis.

**Table B1: Steady-state priors – 95 percent confidence bands.**

| | Real GDP growth | | Inflation | | Interest rate | | Spread | |
|---|---|---|---|---|---|---|---|---|
| | left | right | left | right | left | right | left | right |
| Canada | 1 | 3 | 1 | 3 | 3 | 5 | 0.6 | 2.1 |
| Euro area | 0.5 | 2.5 | 1 | 3 | 2 | 4 | 0.98 | 2.48 |
| Japan | 0 | 2 | 0 | 0 | 0 | 2 | 0.3 | 1.8 |
| Norway | 0.5 | 2.5 | 1.5 | 3.5 | 2 | 4 | -0.95 | 0.55 |
| Sweden | 0.5 | 2.5 | 1 | 3 | 2 | 4 | 0.52 | 2.02 |
| Switzerland | 0.5 | 2.5 | 1 | 3 | 2 | 4 | 0.38 | 1.88 |
| United Kingdom | 1 | 3 | 1 | 3 | 3 | 5 | 0.2 | 1.7 |
| United States | 1 | 3 | 1 | 3 | 2 | 4 | 0.98 | 2.48 |