Measuring the International Dimension in the Evolution of Output Volatility

Gerdie Everaert[1] and Martin Iseringhausen[*2]

[1,2]Ghent University & SHERPPA

November 30, 2016

Abstract

This paper studies output fluctuations in a panel of OECD economies with the aim to decompose the evolution in output volatility into domestic and international factors. To this end we use a factor-augmented dynamic panel model with both domestic and international shocks and spillovers between countries through trade linkages. Changes in the volatility of output growth can be due to a time-varying sensitivity to these shocks, changes in the propagation mechanism or shifts in the variances of shocks. Next to allowing for cross-sectional dependence in the mean equation, through international shocks and spillovers, we explicitly model cross-sectional dependence in the variance equation by specifying a common factor structure in the volatility of domestic shocks. The results show that while the size of international shocks and spillovers does not decrease in most countries, the volatilities of domestic shocks share a clear common decreasing trend. Hence, the 'Great Moderation' appears to be mainly driven by a decline in the volatility of domestic shocks rather than smaller international shocks.

JEL classification: C32, E32, F44

Keywords: Volatility, business cycle, Bayesian model selection

*Correspondence to: Martin Iseringhausen, Ghent University, Faculty of Economics and Business Administration, Campus Tweekerken, Sint Pietersplein 6, 9000 Gent, Belgium. E-mail: Martin.Iseringhausen@UGent.be.

1 Introduction

The sharp decline in output volatility in most advanced economies since the mid 1980s is one of the most striking stylized facts in modern macroeconomics. First documented for the U.S. by Kim and Nelson (1999) and McConnell and Perez-Quiros (2000), the phenomenon has been so widespread and persistent that it was famously named the 'Great Moderation' by Stock and Watson (2003). Although a large literature has already focused on the potential sources and consequences of output volatility, this continues to be an area of lively debate.

One strand of the literature has focused on the fundamentals underlying the observed decline in aggregate volatility such as better monetary policy (Galí and Gambetti, 2009), increased government size and fiscal policy (Fatas and Mihov, 2001), improved inventory management methods (Kahn, McConnell, and Perez-Quiros, 2002), financial innovation and increased global integration (Dynan, Elmendorf, and Sichel, 2006), or demographic changes (Jaimovich and Siu, 2009). Finally, the 'good luck' hypothesis brought forward by Stock and Watson (2003) entails the idea that the period from 1980 onwards has simply been characterized by the absence of large shocks hitting economies. Related to this is the question whether the recent Great Recession marks the end of the Great Moderation. While some authors confirm that this is indeed the case (see e.g. Ng and Wright, 2013), others consider it to be merely a temporary offset of the structural decline in volatility (see e.g. Clark, 2009).

Starting off from Blanchard and Simon (2001) who show that there has been a global decline in output volatility in G7 countries, with magnitude and timing differing across countries, a second strand of the literature tries to explain trends in aggregate volatility in terms of the 'geographic origin', i.e. to what extent these trends are driven by global or country-specific factors. Stock and Watson (2005) estimate a factor-augmented structural VAR where GDP growth is decomposed into common and idiosyncratic shocks as well as spillovers, i.e. shocks that originate in a certain country and subsequently spread to other countries. They find that a decrease in the size of global shocks is responsible for much of the observed decline in business cycle volatility in the G7. Carare and Mody (2012) add evidence that spillovers have become more important since the 1990s and acted as a volatility amplifier during the recent Great Recession. Using a dynamic factor approach, Kose, Otrok, and Whiteman (2003) show that a common world factor is an important source of business cycle volatility in advanced economies. Extending their approach by allowing for time-varying factor loadings and stochastic volatility in the latent factors and idiosyncratic components, Del Negro and Otrok (2008) find that for most G7 countries both the volatility attributable to international business cycles and country-specific fluctuations drop. In fact, their results document that a common drop in the volatility of idiosyncratic fluctuations is an important feature of the Great Moderation, but they leave this aspect unmodeled.

In this paper we set up and estimate a factor-augmented dynamic panel data model with time-varying coefficients and stochastic volatilities to decompose aggregate output growth volatility in international and country-specific factors. More specifically, our encompassing empirical framework allows the moderation in volatility to be driven by (i) smaller international shocks; (ii) a moderation in foreign countries that spills over to the remaining countries; (iii) lower contemporaneous sensitivity to international and foreign shocks; (iv) a milder propagation of shocks over time;

(v) a common and/or idiosyncratic reduction in the volatility of country-specific shocks. Such a general decomposition has not been done before. Disentangling a country's output volatility into its constituent components is also of particular importance for policy makers as it provides information on whether the observed change in output volatility is due to one of the country-specific components, which may be under their control, or due to international factors, which are not.

We contribute to the literature in the following three ways. First, we merge the factor-augmented VAR approach of Stock and Watson (2005), by decomposing output growth shocks into country-specific shocks, common shocks and spillovers, and the dynamic factor approach of Del Negro and Otrok (2008), by allowing for time variation in the variance of shocks and time-varying sensitivities to shocks. Second, we further extend these approaches by explicitly modeling a common factor in the volatility of idiosyncratic shocks. Hence, next to co-movements in countries' GDP through common growth shocks and spillovers, our model is also able to capture co-movement in the size of country-specific shocks. The idea to model a common component in the volatility of otherwise uncorrelated shocks is not entirely new. Kim, Lee, Park, and Yeo (2009) extract macroeconomic uncertainty as the common factor in consumption and dividend growth volatility. Laurini and Mauad (2015) include a common jump factor in a multivariate stochastic volatility model to account for crises and contagion in emerging countries' exchange rates markets. Herskovic, Kelly, Lustig, and Nieuwerburgh (2016) show that not only firms' returns but also their volatilities exhibit a strong common factor structure. However, to the best of our knowledge, we are the first to model a common factor in the volatility of idiosyncratic output growth shocks as one of the potential sources of the Great Moderation. Third, we explicitly address model uncertainty. We start by specifying all coefficients and variance parameters as random walks, but then go on and test which time-varying components are relevant model attributes and fall back to a more parsimonious model when appropriate. This not only avoids over-parameterization but will also provide us with information on which components actually contribute to changes in output volatility.

Using quarterly data on the growth rates of real output for 16 advanced countries over the period 1961:Q1 - 2015:Q4, we obtain the following results. First, the volatility of common shocks clearly varies over time - shooting up around the oils crises of the 1970s, the worldwide recession of the early 1990s and the recent Great Recession - but there is no marked evidence of a declining trend. As individual countries' sensitivity to the common shocks and spillovers has remained stable over the sample period, changes in the volatility of the international business cycle component is not what is driving the Great Moderation. Second, the volatility of domestic shocks shows a clear common downward trend across the 16 advanced economies we consider. We identify this as one of the main drivers of the widespread reduction in volatility. Finally, the Great Recession shows up as a temporary increases in the volatility of common shocks and hence does not mark the end of the Great Moderation.

The remainder of the paper is structured as follows: Section 2 introduces our empirical specification and estimation approach. The main estimation results are presented in Section 3 and further documented by means of variance decompositions in Section 4. Section 5 concludes. A detailed description of the estimation methodology can be found in Appendix A.

2 Model and estimation approach

2.1 Empirical specification

Our starting point is the factor-augmented dynamic panel model proposed by Stock and Watson (2005) extended to allow for time-varying coefficients and stochastic volatilities as in Del Negro and Otrok (2008). More specifically,

$$\Delta y_{it} = \alpha_{it} + \sum_{j=1}^{p} \beta_{it}^{j} \Delta y_{i,t-j} + \sum_{k=1}^{q} \gamma_{it}^{k} \Delta y_{i,t-k}^{*} + \varepsilon_{it}, \tag{1}$$

where Δy_{it} is real GDP growth for country i in quarter t and Δy_{it}^{*} is trade-weighted real GDP growth of the trading partners of country i.

Our model has a number of distinct features. First, as outlined in Stock and Watson (2005), equation (1) can be seen as a vector autoregression (VAR) where the cross-country dimension represents the different variables in the system. The inclusion of Δy_{it}^{*} corresponds to restricting the coefficients on the lags of foreign GDP growth to be proportional to their respective trade shares. Given the medium-size dataset at hand this solves the dimensionality problem which would arise when including the growth rates for each of the foreign countries separately. Moreover, this weighted average offers a convenient spillover measure. Second, the model in equation (1) is structural in the sense that we impose an unobserved component factor structure on the innovations ε_{it},

$$\varepsilon_{it} = \phi_{it}^{\varepsilon} \varepsilon_{t}^{f} + \varepsilon_{it}^{c}, \tag{2}$$

where ε_{t}^{f} are common international shocks with country-specific loadings ϕ_{it}^{ε} and ε_{it}^{c} are country-specific innovations. These are identified through the assumption that spillovers in equation (1) happen with at least one-period lag and imposing that ε_{it}^{c} in equation (2) is an idiosyncratic shock uncorrelated across countries such that all of the contemporaneous cross-country correlation in output growth is induced by the common shock ε_{t}^{f}. Thus, this model makes it possible to quantify both the direct effect of common international shocks ε_{t}^{f} and the indirect effect of spillovers from (domestic and common) shocks ε_{it} in one country to its trading partners.

Third, we specify all coefficients to vary over time according to driftless random walks:

$$\alpha_{it} = \alpha_{i,t-1} + \eta_{it}^{\alpha}, \qquad \eta_{it}^{\alpha} \sim \mathcal{N}(0, \sigma_{\alpha}^{2}), \tag{3}$$

$$\beta_{it}^{j} = \beta_{i,t-1}^{j} + \eta_{it}^{\beta^{j}}, \qquad \eta_{it}^{\beta^{j}} \sim \mathcal{N}(0, \sigma_{\beta^{j}}^{2}), \qquad \text{for } j = 1, \dots, p, \tag{4}$$

$$\gamma_{it}^{k} = \gamma_{i,t-1}^{k} + \eta_{it}^{\gamma^{k}}, \qquad \eta_{it}^{\gamma^{k}} \sim \mathcal{N}(0, \sigma_{\gamma^{k}}^{2}), \qquad \text{for } k = 1, \dots, q, \tag{5}$$

$$\phi_{it}^{\varepsilon} = \phi_{i,t-1}^{\varepsilon} + \eta_{it}^{\phi^{\varepsilon}}, \qquad \eta_{it}^{\phi^{\varepsilon}} \sim \mathcal{N}(0, \sigma_{\phi^{\varepsilon}}^{2}). \tag{6}$$

Hence, we allow for changes in the persistence of shocks as measured by β_{it} as well as changes in a countries sensitivity to both spillovers and common shocks as measured by γ_{it} and ϕ_{it}^{ε}, respectively. We also model the intercept α_{it} as a random walk to capture permanent changes in trend output growth. This avoids that low-frequency drifts, e.g. the productivity slowdown of the early 1970s, bias our decomposition of growth and volatility at the business cycle frequency, a problem first discussed in Fernald (2007).

Finally, the variance of both common and idiosyncratic innovations is allowed to vary stochastically over time:

$$\varepsilon_t^f \sim \mathcal{N}(0, e^{g_t}), \qquad \varepsilon_{it}^c \sim \mathcal{N}(0, e^{h_{it}}). \tag{7}$$

As one of the stylized facts of the Great Moderation is a global reduction in volatility but without an increase in international synchronization of business cycles and with its magnitude and timing varying considerably across countries, country-specific volatilities are likely co-moving even after controlling for common shocks and spillovers. We explicitly model this correlation in the volatilities of country-specific shocks by specifying the log-variance h_{it} as a common factor structure

$$h_{it} = \phi_{it}^h h_t^f + h_{it}^c, \tag{8}$$

where h_t^f is a common factor with country-specific time-varying loading ϕ_{it}^h and h_{it}^c the remaining idiosyncratic part. Again, we assume that the time-varying volatility components follow independent driftless random walk processes:

$$g_t = g_{t-1} + \eta_t^g, \qquad \eta_t^g \sim \mathcal{N}(0, \sigma_g^2), \tag{9}$$

$$h_t^f = h_{t-1}^f + \eta_t^{h^f}, \qquad \eta_t^{h^f} \sim \mathcal{N}(0, \sigma_{h^f}^2), \tag{10}$$

$$\phi_{it}^h = \phi_{i,t-1}^h + \eta_{it}^{\phi^h}, \qquad \eta^{\phi^h} \sim \mathcal{N}(0, \sigma_{\phi^h}^2), \tag{11}$$

$$h_{it}^c = h_{i,t-1}^c + \eta_{it}^{h^c}, \qquad \eta_{it}^{h^c} \sim \mathcal{N}(0, \sigma_{h^c}^2). \tag{12}$$

Note that while all regression and variance parameters are heterogeneous across countries, for the sake of parsimony the variances of their innovations are assumed homogeneous.

For future use, define $\beta_{it} = \left(\beta_{it}^1, \ldots, \beta_{it}^p\right)$ and $\gamma_{it} = \left(\gamma_{it}^1, \ldots, \gamma_{it}^q\right)$. After stacking the unobserved components over cross-sections, i.e. $\beta_t = (\beta_{1t}, \ldots, \beta_{Nt})$ and similarly for the other components, further define the vector of time-varying parameters $\lambda_t = (\alpha_t,\ \beta_t,\ \gamma_t)$, of time-varying factor loadings $\phi_t = (\phi_t^\varepsilon,\ \phi_t^h)$ and of stochastic volatilities $\zeta_t = (g_t,\ h_t^f,\ h_t^c)$. The vectors λ, ϕ and ζ then refer to λ_t, ϕ_t and ζ_t stacked over time. The innovation variances are combined in the vector $\sigma^2 = (\sigma_\alpha^2, \sigma_\beta^2, \sigma_\gamma^2, \sigma_{\phi^\varepsilon}^2, \sigma_{\phi^h}^2, \sigma_g^2, \sigma_{h^f}^2, \sigma_{h^c}^2)$ with $\sigma_\beta^2 = (\sigma_{\beta^1}^2, \ldots, \sigma_{\beta^p}^2)$ and $\sigma_\gamma^2 = (\sigma_{\gamma^1}^2, \ldots, \sigma_{\gamma^q}^2)$. In addition $x_t = (\Delta y_t,\ \Delta y_{t-j},\ \Delta y_{t-k}^*)$ represents the data matrix stacked over cross-sections which is further stacked over time to obtain x.

2.2 Identification and normalization

As it stands, the model in Section 2.1 is not identified and thus requires properly chosen normalizations. A first issue that arises is that the products $\phi_{it}^\varepsilon \varepsilon_t^f$ and $\phi_{it}^h h_t^f$ in equations (2) and (8) are identified but not the relative scale and sign of their constituent components. Multiplying the loadings by a rescaling constant c while dividing the common factor by the same c would leave the product unchanged. As long as the standard deviations of the innovations to both components are appropriately adjusted, the two models are equivalent. A standard normalization is therefore to constrain the scale of the factor (see e.g. Del Negro and Otrok, 2008). However, while being effective in a model with fixed loadings, time variation brings about a new identification issue as the rescaling term can now be a time-varying sequence c_t rather than a constant c. Del Negro and Otrok (2008) correctly argue that in principle this does not pose a formal identification problem as multiplying the time-varying loadings with c_t implies that the rescaled loadings no longer satisfy the model's assumptions. For instance, when transforming ϕ_{it}^{ε} to $c_t\phi_{it}^{\varepsilon}$ the innovations $c_t\eta_{it}^{\phi^{\varepsilon}}$ would no longer satisfy the homoskedasticity assumption of equation (6). Our experience is that this normalization results in a model that is only weakly identified, though. Hence, we avoid multiplication by a time-varying sequence c_t by restricting the cross-sectional averages of the loadings to be 1 in each period:

$$\bar{\phi}_t^{\varepsilon} = \frac{1}{N}\sum_{i=1}^{N}\phi_{it}^{\varepsilon} = 1, \qquad \bar{\phi}_t^{h} = \frac{1}{N}\sum_{i=1}^{N}\phi_{it}^{h} = 1, \qquad \forall\ t = 1, ..., T. \tag{13}$$

This boils down to dividing the original, unnormalized loadings by their cross-sectional average in every period t and assuming that equations (6) and (11) hold for the rescaled loadings. Note that this normalization scheme implies that a country's loading should be interpreted as being relative to the average loading. Hence, increasing business cycle integration for all countries in the sample should come about through bigger global shocks rather than through an overall increase in the sensitivity to these shocks. A further advantage of the normalization in equation (13) is that the signs of the factor and the loadings are identified and hence avoids the MCMC draws to switch sign along the Markov chain.

A further weak identification issue may arise when separating the constituent components of the idiosyncratic log-variance h_{it} in equation (8). Although, the country-specific random walk processes h_{it}^c are assumed to be uncorrelated across cross-sections, this is not imposed when filtering these sequences using the Kalman filtering approach (outlined below). Hence, there is some scope for h_{it}^c to pick up common volatility trends that should in fact be captured by $\phi_{it}^h h_t^f$. For this reason, we restrict the cross-sectional average of h_{it}^c to 0 in each period:

$$\overline{h}_t^c = \frac{1}{N}\sum_{i=1}^{N} h_{it}^c = 0, \qquad \forall\ t = 1, ..., T. \tag{14}$$

This restriction is consistent with our assumption that all co-movement in countries' log-variances should stem from the common component $\phi_{it}^h h_t^f$ while the remaining idiosyncratic volatilities h_{it}^c should no longer include a common factor.

Note that our normalizations imply that h_t^f will correspond to the panel average log-volatility in each period and hence resembles the Common Correlated Effects (CCE) approach of Pesaran (2006) to estimate panel data models with a common factor structure in the errors. To see this, take cross-sectional averages of h_{it} in equation (8) and solve for the common factor h_t^f to obtain

$$h_t^f = \frac{1}{\overline{\phi}_t^h}\left(\overline{h}_t - \overline{h}_t^c\right), \tag{15}$$

with $\overline{h}_t$ being the cross-sectional average of h_{it}. Plugging this expression for h_t^f back in equation (8), we obtain

$$h_{it} = \frac{\phi_{it}^h}{\overline{\phi}_t^h}\left(\overline{h}_t - \overline{h}_t^c\right) + h_{it}^c = \phi_{it}^{h*}\overline{h}_t + h_{it}^{c*}, \tag{16}$$

where $\phi_{it}^{h*} = \frac{\phi_{it}^h}{\overline{\phi}_t^h}$ and $h_{it}^{c*} = h_{it}^c - \frac{\phi_{it}^h}{\overline{\phi}_t^h}\overline{h}_t^c$. It is easily verified that the cross-sectional averages of ϕ_{it}^{h*} and h_{it}^{c*} are 1 and 0, respectively, for each t. This is exactly the normalization we impose in equations (13) and (14).

2.3 Bayesian estimation and stochastic model specification search

Bayesian estimation

The time-varying dynamic panel model outlined above corresponds to a state space model with the observation equation given by merging equations (1)-(2) and (7)-(8) and the state equations (3)-(6) and (9)-(12) describing the laws of motion for the unobserved random walk components. In a standard linear Gaussian state space model, the Kalman filter can be used to filter the unobserved states from the data and to construct the likelihood function such that the unknown parameters can be estimated using maximum likelihood. However, the inclusion of the stochastic volatilities g_t and h_{it} implies a non-linear estimation problem for which the standard approach via the Kalman filter and maximum likelihood is not feasible. Instead, we use the Bayesian Markov Chain Monte Carlo simulation-based analysis of the stochastic volatility model as outlined in Kim, Shephard, and Chib (1998). This reduces the non-linear estimation problem into a sequence of blocks which are linear conditional on the other blocks. Our Bayesian estimation approach is also convenient to perform stochastic model specification search as we outline below.

Stochastic model specification search

The main aim of this paper is to determine the time-varying sources of output volatility. While previous research has already dealt with this question, most work relies on structural break tests (e.g. McConnell and Perez-Quiros, 2000), ad hoc split-sample regressions (e.g. Stock and Watson, 2005) or impose time-variation on various model components (Del Negro and Otrok, 2008). Instead of merely assuming time variation from the outset, we will test for which model components the time variation is actually relevant and fall back to a more parsimonious specification when appropriate. This not only avoids over-parameterization but will also provide us with information on which components are driving the Great Moderation.

Model selection for UC models is a challenge, though, as it leads to testing problems that are non-regular from a classical point of view. For the autoregressive parameter β_{it}, for example, this implies testing $\sigma_\beta^2 = 0$ versus $\sigma_\beta^2 > 0$, which is non-regular as the null hypothesis lies on the boundary of the parameter space. Our Bayesian approach is convenient to evaluate this null hypothesis. The basic idea of Bayesian stochastic model specification search is to assign a prior probability to each of the possible models and then derive the posterior probability for each model conditional on the data. In particular, Fruehwirth-Schnatter and Wagner (2010) show how to extend Bayesian variable selection in standard regression models to UC models. Their approach relies on a (i) non-centered parameterization of the model in which (ii) binary stochastic indicators for each of the model components are sampled together with the parameters in the MCMC and (iii) the standard inverse Gamma prior for the variances of innovations to the components is replaced by a Gaussian prior centered at zero for the square root of these variances. The exact implementation of this stochastic model specification search is outlined below.

Non-centered parametrization

Fruehwirth-Schnatter and Wagner (2010) start by rewriting the model's time-varying components into a non-centered parametrization. The random walk specification of, for instance, the autoregressive parameters β_{it} in (4) can be reparameterized as

$$\beta^j_{it} = \beta^j_{i0} + \sigma_{\beta^j}\widetilde{\beta}^j_{it}, \tag{17}$$

$$\text{with} \qquad \widetilde{\beta}^j_{it} = \widetilde{\beta}^j_{i,t-1} + \eta^{\widetilde{\beta}^j}_{it}, \qquad \widetilde{\beta}^j_{i0} = 0, \qquad \eta^{\widetilde{\beta}^j}_{it} \sim \mathcal{N}(0,1), \tag{18}$$

for $j = 1, \ldots, p$ and where $\sigma_{\beta^j}\widetilde{\beta}^j_{it}$ is the time-varying part of β^j_{it} and β^j_{i0} the initial value if β^j_{it} varies over time ($\sigma_{\beta^j} > 0$) while being its constant value if there is no time variation ($\sigma_{\beta^j} = 0$). The other random walk components α_{it}, γ_{it}, ϕ^{ε}_{it}, ϕ^h_{it}, g_t, h^f_t and h^c_{it} in the model can be rewritten in a similar way, with α_{i0}, β_{i0}, γ_{i0}, ϕ^{ε}_{i0}, ϕ^h_{i0}, g_0, h^f_0, h^c_{i0} referring to the initial values. For future use let $\lambda_0 = (\alpha_0, \beta_0, \gamma_0)$, $\phi_0 = (\phi^{\varepsilon}_0, \phi^h_0)$ and $\zeta_0 = (g_0, h^f_0, h^c_0)$ be the initial values stacked over cross-sections. Similarly define the time varying parts as $\widetilde{\lambda} = (\widetilde{\alpha}, \widetilde{\beta}, \widetilde{\gamma})$, $\widetilde{\phi} = (\widetilde{\phi}^{\varepsilon}, \widetilde{\phi}^h)$ and $\widetilde{\zeta} = (\widetilde{g}, \widetilde{h}^f, \widetilde{h}^c)$.

The non-centered parameterization offers several features that will prove useful for model selection. First, it is not identified as the signs of σ_{β^j} and $\widetilde{\beta}^j_{it}$ can be changed while leaving their product unchanged. As a result, the likelihood function is symmetric around zero along the σ_{β^j} dimension. When β^j_{it} varies over time (i.e. $\sigma^2_{\beta^j} > 0$) the likelihood function is bimodal with modes $-\sqrt{\sigma^2_{\beta^j}}$ and $\sqrt{\sigma^2_{\beta^j}}$. When β^j_{it} is constant (i.e. $\sigma^2_{\beta^j} = 0$) the likelihood function is unimodal around zero. Hence, non-identification of the sign of σ_{β^j} offers a first view on whether β^j_{it} varies over time or not.

Parsimonious specification

The non-centered parameterization is very useful for model selection as, in contrast to the original component β^j_{it}, the transformed process $\widetilde{\beta}^j_{it}$ does not degenerate to a time-invariant parameter when $\sigma_{\beta^j} = 0$ as the constant part is now represented by β^j_{i0}. This allows us to reformulate the question whether β^j_{it} is time-varying as a more standard variable selection problem. To this end, Fruehwirth-Schnatter and Wagner (2010) define the parsimonious specification as

$$\beta^j_{it} = \beta^j_{i0} + \delta_{\beta^j}\sigma_{\beta^j}\widetilde{\beta}^j_{it}, \tag{19}$$

where δ_{β^j} is a binary indicator that is either 0 or 1. If $\delta_{\beta^j} = 0$, the component $\widetilde{\beta}^j_{it}$ is excluded from the model and σ_{β^j} is set to zero. If $\delta_{\beta^j} = 1$, the time-varying component $\widetilde{\beta}^j_{it}$ is included and σ_{β^j} is estimated.

Defining similar binary indicators for the other time-varying components and collecting all of them in the vector $\mathcal{M} = (\delta_\alpha, \delta_\beta, \delta_\gamma, \delta_{\phi^\varepsilon}, \delta_{\phi^h}, \delta_g, \delta_{h^f}, \delta_{h^c})$, with $\delta_\beta = (\delta_{\beta^1}, \ldots, \delta_{\beta^p})$ and $\delta_\gamma = (\delta_{\gamma^1}, \ldots, \delta_{\gamma^q})$, the specification of the model is described by a combination of the elements in $\mathcal{M}$. Note that we assume the innovation variances in σ^2 and binary indicators in $\mathcal{M}$ to be homogeneous across countries and hence test for panel-wide time-variation in the parameters rather than testing each country separately.

Gaussion prior centered at zero

Our Bayesian estimation procedure requires choosing prior distributions for the time-invariant part of the parameters in λ_0, ϕ_0, and ζ_0, for the innovation variances in σ^2 and for probabilities of the binary indicators in $\mathcal{M}$ being 1. It is well known that when using the standard inverse Gamma prior distribution for the variances in σ^2, the choice of the shape and scale that define this distribution has a strong influence on the posterior, especially when the true value of the variance is close to zero. More specifically, as the inverse Gamma distribution does not have probability mass at zero, using it as a prior distribution tends to push the posterior density away from zero. This is particularly problematic as we want to decide whether the model's parameters are time-varying or not, i.e. whether their innovation variances are zero or not. Due to the fact that in the non-centered parameterization, as outlined above, the standard deviations σ of the innovations to the random walk processes enter as regression parameters, we can replace the commonly used inverse Gamma prior for σ^2 by a Gaussian prior centered at zero for σ. Fruehwirth-Schnatter and Wagner (2010) show that the posterior density is much less sensitive to the hyperparameters of the Gaussian prior distribution for σ compared to using an inverse Gamma prior for σ^2. In particular, the posterior is not pushed away from zero when in fact $\sigma^2 = 0$.

We therefore use a Gaussian prior distribution for all parameters. First, we choose an uninformative prior $\mathcal{N} \sim (0, 1)$ for the time-invariant part of the parameters λ_0 and of the stochastic volatilities ζ_0 while using $\mathcal{N} \sim (1, 1)$ for the time-invariant part of the factor loadings ϕ_0. The latter is consistent with our chosen normalization scheme, introduced in Section 2.2, that imposes the cross-sectional average of the factor loadings to be 1. Second, the prior distributions for the standard deviations σ to the various random walk components are centered around zero with the variance chosen such that the prior distribution has support over the range of relevant parameter values, given the scale of the data and the fact that most of the time-varying parameters capture slow long-run developments. The exact prior choices for σ are provided in the left part of Table 1. Finally, for each of the binary indicators in $\mathcal{M}$ we choose a Bernoulli prior distribution where each indicator has a prior probability $p_0 = 0.5$ of being 1. Robustness of the results will be checked with respect to the prior choices.

MCMC algorithm

Our MCMC scheme is as follows:

1. Sample the binary indicators in $\mathcal{M}$ together with the constant parameters $(\lambda_0, \phi_0, \zeta_0)$ and the standard deviations σ conditional on the time-varying states $\widetilde{\lambda}$, $\widetilde{\phi}$ and $\widetilde{\zeta}$. Restricted elements in σ, i.e. for which the corresponding binary indicator in $\mathcal{M}$ is zero, are set to zero.

2. Sample the time-varying parameters $\widetilde{\lambda}$, the time-varying loadings $\widetilde{\phi}$ and the stochastic volatilities $\widetilde{\zeta}$ conditional on the binary indicators $\mathcal{M}$, the constant parameters $(\lambda_0, \phi_0, \zeta_0)$ and the standard deviations σ.

3. Perform a random sign switch for $\widetilde{\lambda}$, $\widetilde{\phi}$, $\widetilde{\zeta}$ and the corresponding standard deviations σ. Hence, change their sign with probability 0.5.

Starting from an arbitrary set of initial values, sampling from these blocks is iterated J times and after a sufficiently long burn-in period B, the sequence of draws $(B+1,\ldots,J)$ can be taken as a sample from the joint posterior distribution of interest $f(\lambda_0, \phi_0, \zeta_0, \sigma, \widetilde{\lambda}, \widetilde{\phi}, \widetilde{\zeta}, \mathcal{M}|x)$. A detailed description of the exact implementation for each of the blocks can be found in Appendix A. The results reported below are based on 50,000 iterations with 10,000 draws being discarded as burn-in.

3 Estimation results

In this section we present our main estimation results. Following Stock and Watson (2005), all reported results are obtained setting $p = 4$ and $q = 1$ in equation (1). Experimenting with alternative lag structures shows that the results are robust with respect to this choice. After discussing the data used, we start off with the stochastic model specification search to test for time-variation in the various model components. Next, we present the results for the chosen parsimonious specification. We end with a number of robustness tests.

3.1 Data

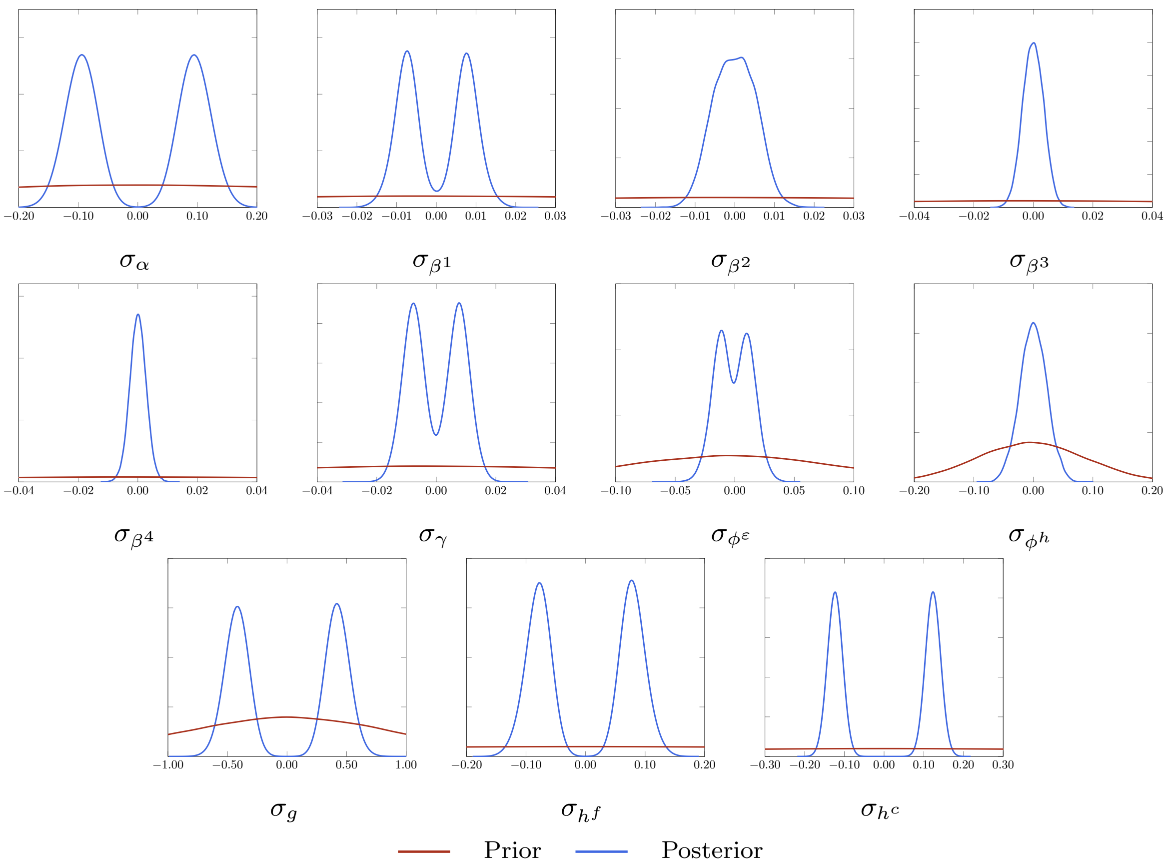
We estimate the model outlined in Section 2 using quarterly data for 16 advanced economies over the period 1961:Q1 - 2015:Q4. The included countries are: Australia, Austria, Belgium, Canada, Finland, France, Germany, Italy, Japan, Netherlands, Portugal, Spain, Sweden, Switzerland, United Kingdom, United States. As our focus is on economic fluctuations over the business cycle horizon, we follow Stock and Watson (2005) and filter out high frequency (quarter-to-quarter) fluctuations by measuring economic growth Δy_{it} as the year-on-year growth rate of real GDP. As a robustness check, we will also report results from using quarter-on-quarter growth rates in Section 3.4. Real GDP is taken from the OECD Quarterly National Accounts database. Trade-weighted growth rates are calculated as $\Delta y^*_{it} = \sum_{j=1,j\neq i}^{N} w^j_{it} \Delta y_{it}\ \forall\ i = 1,...,N,\ t = 1,...,T$, where $w^j_{it} = (\mathrm{EX}^j_{it} + \mathrm{IX}^j_{it})/(\mathrm{EX}_{it} + \mathrm{IM}_{it})$ is the share of country j in total gross trade of country i. Gross trade is taken from the OECD Quarterly International Trade database.

It should be noted, that some caution is needed when using longer series of quarterly real GDP. As early GDP data is only available on an annual basis for the majority of countries, the OECD uses interpolation methods to construct quarterly data. Depending on the length of the interpolated period and the complexity of the applied method, this results in artificial volatility patterns for some countries (most obviously for Belgium and Portugal). Although the country-specific error ε^c_{it} and volatility h^c_{it} terms should be able to capture most of these artificial fluctuations, they may still distort the results of our empirical model. This is especially the case for the relative importance of common and domestic movements in volatility. One option would be to drop all affected countries from the sample, but this would heavily reduce the cross-sectional dimension of our panel. Instead, we estimate the model using the full sample of 16 countries but as a robustness check we also report results obtained when dropping countries with data issues. As an additional robustness check we will also estimate the model using annual instead of quarterly data. It should further be noted that we always replace three larger outliers in the quarterly growth data (France 1968Q2 and 1968Q3, Germany 1991Q1) with their series-specific full sample median growth rate.

3.2 Results stochastic model specification search

We start by estimating an unrestricted model with all binary indicators in $\mathcal{M}$ set to one to generate posterior distributions for the standard deviations σ of the innovations to the 11 non-centered components of interest. As the sign of these standard deviations is not identified, a bimodal posterior distribution is a first indicating of time-variation.

Figure 1: Posterior densities of the standard deviations σ (all binary indicators set to 1)



Posterior densities are plotted in Figure 1, while Table 1 reports the median and percentiles of the absolute value of the standard deviations. A number of interesting features stands out. First, the standard deviations of the innovations to each of the three stochastic volatility components g_t, h_t^f, and h_{it}^c all have a clear-cut bimodal posterior density with no probability mass at zero. This suggests that changes in the size of both common and domestic shocks play an important role in the evolution of aggregate output growth volatility over the last five decades. The variance of global shocks g_{it} is subject to the largest innovations, with a posterior median absolute standard deviation of around 0.5. But also the variance of domestic shocks h_{it} varies over time, induced by time-variation in both the common volatility factor h_t^f and the purely country-specific component h_{it}^c. Second, there appears to be some time-variation in the persistence of output growth. The posterior distribution of the standard deviation of innovations to β_{it}^1 shows clear bimodality, although there is still probability mass left at zero and the absolute posterior median of 0.01 is rather small. For β_{it}^2, β_{it}^3 and β_{it}^4 there is no sign of time-variation. Third, there is also some evidence that the sensitivity to spillovers γ_{it} varies over time. The posterior density of σ_γ is bimodal but also has considerable probability mass at zero. Fourth, accounting for a time-varying mean growth rate α_{it} proves to be necessary as there is clear bimodality in the posterior distribution of σ_α. Finally, we do not find convincing evidence for changes in countries' sensitivity to common growth shocks and to the common volatility factor. The posterior density of $\sigma_{\phi^\varepsilon}$ is bimodal but also has considerable probability mass at zero while that of σ_{ϕ^h} is clearly unimodal.

Table 1: Summary information for the prior and posterior distributions of the standard deviations σ

Parameter		Prior variance V_0	Posterior median	Posterior 5%	Posterior 95%
Std. of long-run growth	σ_α	0.5^2	0.125	0.084	0.174
Std. of AR(1) coefficient	σ_{β^1}	0.1^2	0.010	0.004	0.015
Std. of AR(2) coefficient	σ_{β^2}	0.1^2	0.004	0.000	0.011
Std. of AR(3) coefficient	σ_{β^3}	0.1^2	0.003	0.000	0.008
Std. of AR(4) coefficient	σ_{β^4}	0.1^2	0.002	0.000	0.006
Std. of spillover sensitivity	σ_γ	0.1^2	0.009	0.002	0.016
Std. of loadings common growth shocks	$\sigma_{\phi^\varepsilon}$	0.1^2	0.013	0.002	0.029
Std. of loadings common volatility factor	σ_{ϕ^h}	0.1^2	0.017	0.002	0.049
Std. of SV common growth shocks	σ_g	1.0^2	0.532	0.390	0.703
Std. of common volatility factor	σ_{h^f}	1.0^2	0.103	0.077	0.144
Std. of idiosyncratic volatility	σ_{h^c}	1.0^2	0.154	0.128	0.181

Notes: The prior distribution is $\mathcal{N}(0, V_0)$. The posterior distribution is for the absolute value of the standard deviations.

Inspecting the posterior density of the innovation standard deviations only provides a first idea on the presence of time variation. As a more formal test, we next sample the stochastic binary indicators in $\mathcal{M}$ together with the other parameters in the model. Table 2 reports the posterior probabilities for the binary indicators being one. These probabilities are calculated as the fraction of draws in which the stochastic model specification search prefers a model which allows for time-variation in the corresponding parameter. To check the robustness over alternative prior variances for the innovation standard deviations σ, we multiply V_0 as reported in Table 1 by a scaling factor v_0. Hence, the middle row (where $v_0 = 1$) corresponds to our baseline scenario, while the first two rows (where $v_0 < 1$) and the last two rows (where $v_0 > 1$) imply more and less informative priors respectively. Rows 3 and 5 further check the robustness with respect to alternative values for the prior inclusion probability p_0.

Table 2: Posterior inclusion probabilities for the binary indicators $\mathcal{M}$ over different prior variances for σ and different prior inclusion probabilities p_0

Priors		Posterior inclusion probabilities										
		Intercept	AR coefficients				Spillovers	Loadings		Stochastic volatilities		
p_0	v_0	δ_α	δ_{β^1}	δ_{β^3}	δ_{β^3}	δ_{β^4}	δ_γ	$\delta_{\phi^\varepsilon}$	δ_{ϕ^h}	δ_g	δ_{h^f}	δ_{h^c}
0.5	0.1	1.00	0.84	0.15	0.11	0.08	0.51	0.61	0.41	1.00	1.00	1.00
0.5	0.5	1.00	0.80	0.10	0.05	0.04	0.32	0.42	0.28	1.00	1.00	1.00
0.1	1	1.00	0.21	0.01	0.00	0.00	0.08	0.07	0.03	1.00	1.00	1.00
0.5	1	1.00	0.81	0.11	0.04	0.03	0.28	0.33	0.23	1.00	1.00	1.00
0.9	1	1.00	0.93	0.43	0.28	0.22	0.75	0.78	0.72	1.00	1.00	1.00
0.5	2	1.00	0.69	0.07	0.04	0.02	0.25	0.29	0.21	1.00	1.00	1.00
0.5	10	1.00	0.46	0.03	0.01	0.01	0.16	0.16	0.09	1.00	1.00	1.00

Notes: The prior distribution for each of the elements in σ is $\mathcal{N}(0, v_0 V_0)$ with V_0 the variance in the baseline scenario as reported in Table 1. p_0 is the prior inclusion probability of the binary indicators. The posterior inclusion probabilities are calculated as the average selection frequencies over all iterations of the Gibbs sampler.

Overall, the results in Table 2 confirm our earlier conclusions. Under the baseline prior scenario ($p_0 = 0.5$; $v_0 = 1$) we find strong evidence of time-variation in the stochastic volatility components, the intercept and to a slightly lesser extent in the AR(1) coefficient. Moreover, the inclusion probabilities for the stochastic volatility components and the intercept are completely unaffected by the prior choice. For the AR(1) parameter, the inclusion probability only falls below 50% when using a very loose prior for the innovation standard deviation ($v_0 = 10$) or a low prior inclusion probability ($p_0 = 0.1$). For the other components there is much less evidence in favor of time variation. Only for the spillover parameter and the factor loadings, the posterior inclusion probabilities exceed 50% when using a very informative prior for the innovation standard deviation ($v_0 = 0.1$)[1] or a high prior inclusion probability ($p_0 = 0.9$).

The main conclusions of our tests for time variation are roughly in line with previous findings and discussions in the literature. First, Antolin-Diaz, Drechsel, and Petrella (2016) have recently shown that long-run growth in the U.S. is characterized by a slowly but persistently decreasing pattern. A similar result can be found in Berger, Everaert, and Vierke (2016). Second, the discussion whether persistence as measured by the AR coefficients has changed over time is more controversial. Using rolling regression approaches and different break tests for U.S. data, Blanchard and Simon (2001) and Stock and Watson (2003) do not find evidence for changes in the dynamics of output growth. In contrast, Galí and Gambetti (2009) document changes in the conditional

[1]The increase in the posterior probabilities when using a smaller (and vise versa higher) prior variance for σ may appear counter intuitive, but results from the fact that in this case more weight is given to less extreme (and hence more likely) values of σ when calculating the marginal likelihood of the model with time variation. Moreover, by allowing for less time variation the competing models become more similar in their marginal likelihoods causing a tendency for the posterior inclusion probability to shrink towards the prior $p_0 = 0.5$.

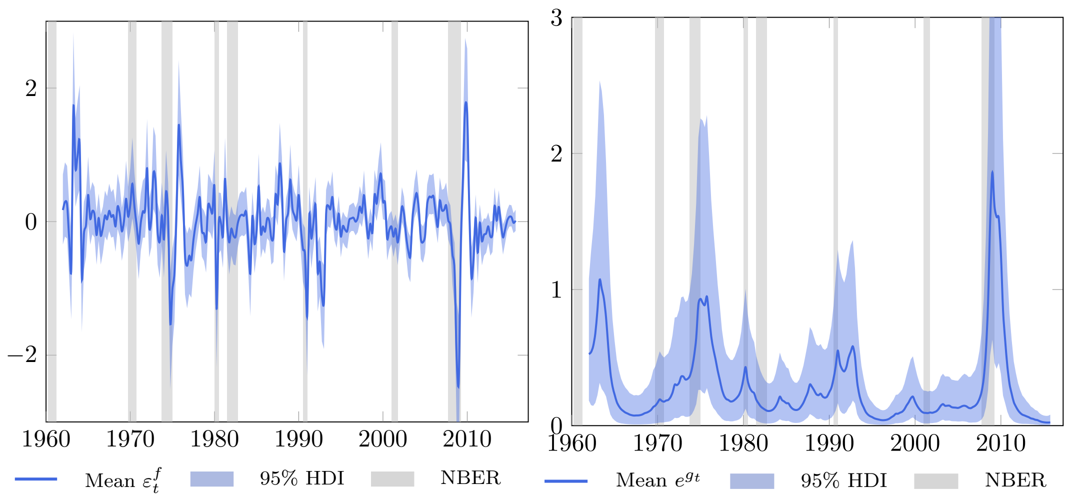
and unconditional moments of several components of output in the United States. Third, also the evidence concerning convergence to a common business cycle is mixed. Kose, Otrok, and Whiteman (2008) find evidence for convergence among industrialized countries while Doyle and Faust (2002) conclude that this has not been the case. Finally, we more formally confirm the observation by Del Negro and Otrok (2008) that there has been a common drop in the volatility of idiosyncratic shocks.

3.3 Parameter estimates and unobserved components

In this section we present the posterior distributions of constant parameters and time-varying components in the parsimonious specification. To get further insights into the role played by the various model components, a variance decomposition of the evolution of total volatility will be presented in Section 4.

Based on the results of the model selection presented in the previous section, our parsimonious specification allows for time-variation in the different variance components, the intercept and the AR(1) coefficient. The other parameters are fixed to be invariant over time. We have experimented with models allowing for time-variation in the spillover parameters γ_{it} and the factor loadings ϕ^{ε}_{it}, for which some evidence of time-variation existed, but we found no clear trends in these parameters. Moreover, the behavior of the other components was largely invariant to these changes.

Figure 2: Common growth shocks and their volatility



(a) Common shocks (b) Volatility of common shocks

Note: HDI is the 95% highest density interval. The gray bars indicate National Bureau of Economic Research (NBER) recessions.

Common shocks and spillovers

Figure 2 plots the posterior means and 95% highest density intervals (HDI) of the common shocks ε_t^f along with their time-varying volatility. Several periods characterized by large common shocks correspond to well-known events. We clearly identify the early 1960s recession, the oil price crises of 1973/1974 and 1979/1980, the worldwide recession of the early 1990s and the recent Great Recession of 2007-2009. The timing of most U.S. recessions, which are indicated by the gray bars in Figure 2, coincides with the occurrence of large negative global economic shocks. The only exceptions are the relatively mild recession in 1969/1970, the early 1980s recession due to the Federal Reserve's contractionary monetary policy, as well as the recession in the early 2000s associated with the burst of the dot-com bubble and the September 11th attacks, which are all U.S. recessions that do not show up as global shocks.

In line with the stochastic model specification search in section 3.2, Figure 2(b) shows considerable variation in the volatility of common growth shocks over time. Periods with larger common shocks are followed by more tranquil times and vice versa. There is however no sign of a decreasing trend. This shows that a reduction in the size of common shocks is not a driver of the observed decrease in volatility across advanced economies. This is in line with the different timing in the volatility decline across countries. Also note that the Great Recession shows up as a temporary increases in the volatility of common shocks and hence does not mark the end of the Great Moderation according to our results.

The stochastic model specification search also showed that country-specific sensitivities to the common shocks did not change over the sample period. The left hand side of Table 3 therefore presents the time-invariant factor loadings ϕ_{i0}^{ε}. Most countries exhibit very similar sensitivity to common shocks, i.e. the posterior mean is close to one for the majority of the countries in our sample. Australia constitutes a clear exception. Its median factor loading of 0.37 signals a partial decoupling of the international business cycle. Also Canada, Spain and the U.S. seem to be somewhat less sensitive to global shocks. Other countries like Finland, Germany, and Sweden seem to be particularly sensitive as indicated by a median factor loading clearly exceeding one.

Next, we turn to evaluating the role of spillovers. The middle part of Table 3 presents summary results for the posterior distributions of γ_{i0}, which is the country-specific sensitivity to lagged trade-weighted average growth rates $\Delta y_{i,t-1}^*$. The stochastic model specification search showed these sensitivities to be time invariant. Nevertheless, there are significant differences across countries. On the one hand, some European countries appear to be particularly sensitive to growth spillovers transmitted via the trade channel. Those include small open economies such as Austria, Belgium, Finland, Netherlands and Sweden but also Italy. On the other hand, we find the more closed economies Australia, Japan, the U.K. and the U.S. to be much less affected by spillovers.

Table 3: Posterior distributions of factor loadings and spillover sensitivities

	Growth: Loadings and Spillovers						Volatility: Loadings		
	$\phi^{\varepsilon}_{i,0}$	Percentiles		$\gamma_{i,0}$	Percentiles		$\phi^{h}_{i,0}$	Percentiles	
	median	5%	95%	median	5%	95%	median	5%	95%
Australia	0.37	0.11	0.61	0.11	0.04	0.20	1.11	0.38	1.83
Austria	1.16	0.95	1.38	0.44	0.32	0.57	0.69	-0.04	1.47
Belgium	0.97	0.79	1.15	0.37	0.28	0.47	0.72	-0.02	1.41
Canada	0.75	0.56	0.93	0.24	0.14	0.33	0.57	-0.15	1.30
Finland	1.38	1.03	1.72	0.33	0.17	0.50	1.02	0.25	1.78
France	1.00	0.86	1.13	0.28	0.19	0.39	1.04	0.29	1.76
Germany	1.50	1.27	1.75	0.27	0.11	0.43	1.38	0.71	2.07
Italy	1.06	0.85	1.26	0.40	0.27	0.53	1.38	0.69	2.10
Japan	1.01	0.67	1.34	0.12	0.01	0.23	1.26	0.56	1.99
Netherlands	1.19	0.98	1.43	0.42	0.28	0.56	0.70	0.00	1.45
Portugal	0.93	0.67	1.18	0.18	0.06	0.30	1.14	0.42	1.91
Spain	0.72	0.56	0.87	0.15	0.07	0.24	1.68	0.97	2.38
Sweden	1.39	1.11	1.68	0.35	0.20	0.52	1.09	0.43	1.80
Switzerland	1.01	0.81	1.21	0.23	0.13	0.34	0.87	0.14	1.64
United Kingdom	0.90	0.68	1.11	0.07	-0.04	0.18	1.01	0.19	1.81
United States	0.68	0.48	0.86	0.04	-0.06	0.14	0.32	-0.36	0.99

The left hand side of Table 4 reports average correlation coefficients of the original output growth rates Δy_{it} and the model's residuals $\hat{\varepsilon}^{c}_{it}$. As expected, output growth is positively correlated with an average pairwise correlation coefficient of around 0.5. For the residuals this crumbles to -0.02. This shows that common shocks and spillovers are sufficient to capture most of the cross-country correlation in output growth rates.

Table 4: Cross-sectional correlation in output growth and its volatility

	Growth		Volatility		
	Δy_{it}	$\hat{\varepsilon}^{c}_{it}$	$\Delta\text{Var}(\Delta y_{it})$	$\Delta e^{(\phi^{h}_{it} h^{f}_{t} + h^{c}_{it})}$	$\Delta e^{h^{c}_{it}}$
Avg. corr.	0.53	-0.02	0.74	0.43	-0.04
5th perc.	0.34	-0.19	0.44	0.09	-0.28
95th perc.	0.74	0.14	0.94	0.75	0.24

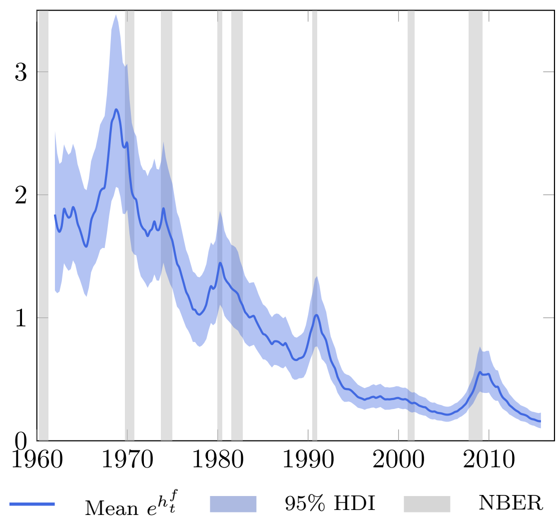
Notes: Reported are the average along with the 5th and 95th percentiles of the country-by-country cross-correlations. The total variance series $\text{Var}(\Delta y_{it})$ is calculated from our model estimates using a simulation-based approach. Details can be found in Section 4. As the estimated stochastic volatilities are non-stationary by construction, we report correlations for first-differenced series to avoid spurious correlations.

Domestic shocks: common and idiosyncratic volatility

The right hand side of Table 4 provides a first view on the correlation structure of output volatility across our 16 advanced economies. First, with a correlation of 0.74, co-movement in output volatility is clearly present. Second, the correlation coefficient of 0.43 for changes in the volatilities of domestic shocks suggest that there is an important role for commonality in the size of country-specific shocks for the overall correlation structure in output volatility. Third, the last column in Table 4 confirms that there is no significant correlation left in the idiosyncratic volatilities h^c_{it}.

We will now take a closer look at the evolution and composition of the volatility of domestic shocks. The model specification search indicated that the country-specific loadings ϕ^h_{it} on the common volatility factor h^f_t do not exhibit time-variation, while both the common volatility factor itself and the remaining idiosyncratic volatility h^c_{it} were found to vary over time. Figure 3 plots the posterior mean and HDI of the common volatility factor. The variances of domestic growth shocks clearly share a common downward trend. This implies that the observed widespread decline in output volatility is induced by a common drop in the size of domestic rather than global shocks or spillovers. Although already suggested by the results in Del Negro and Otrok (2008), this paper is the first to actually model and quantify such a common volatility factor. Next to a clear downward trend, common volatility also shoots up around the time that major global shocks occur. This suggests that the turmoil caused by a large global shock is further amplified through a global increase in country-specific macroeconomic uncertainty.

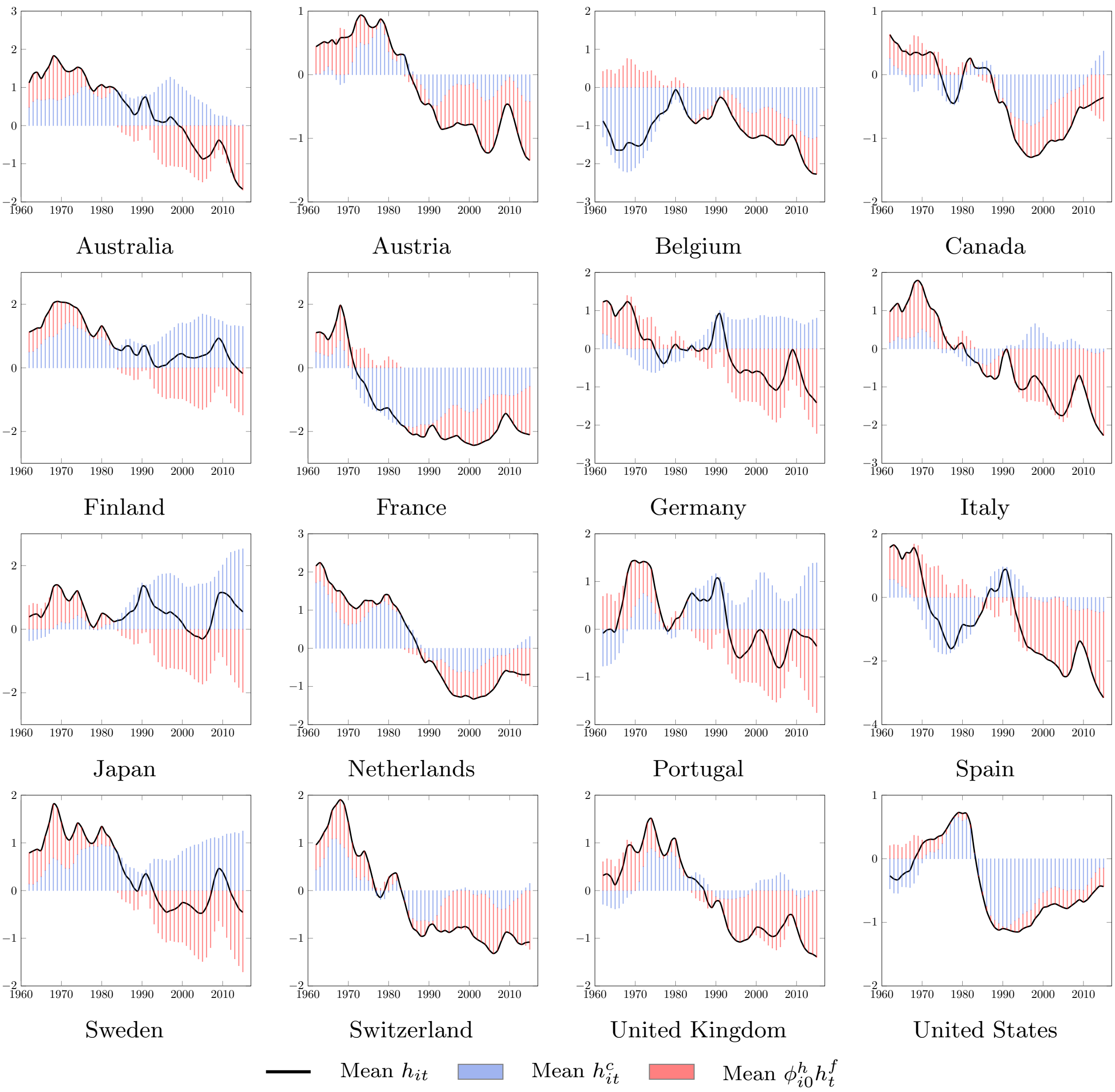
Figure 3: Common factor in the variance of domestic shocks



Results for the loadings ϕ^h_{i0} on the common volatility factor h^f_t are presented in Table 3. Next to exhibiting no significant time variation, as indicated by the results from section 2.3, they are also rather similar across countries, i.e. the posterior distributions include 1 in nearly all countries. Note however that the loadings are not very precisely estimated. The U.S. is the only country for which the factor loadings ϕ^h_{i0} are significantly smaller than 1 with a posterior median of around 0.3.

Figure 4 presents a decomposition of the evolution in the total volatility of domestic shocks into the contribution of the common factor and the idiosyncratic component for each of the 16 considered countries. The relative importance of these two components clearly varies over countries. First, and most striking is the fact that the sharp drop in volatility in the U.S. seems to be purely country-specific in the sense that its particular pronounced pattern is fully captured by the idiosyncratic volatility component. Second, idiosyncratic volatility has remained much more stable in countries like Australia, Finland, Italy or Sweden. Hence, in these countries the observed drop in the volatility of domestic shocks is almost exclusively accounted for by the common volatility factor. Third, in other countries like Germany, Japan and Portugal an increase in the idiosyncratic component has counteracted the decrease in the common component, making the total drop in volatility less pronounced. In Germany for instance, idiosyncratic volatility increased towards the reunion in 1990. In Japan, the idiosyncratic component increased steadily since the beginning of the 1990s, a period of low growth and stagnation commonly called ‘the lost decades’.

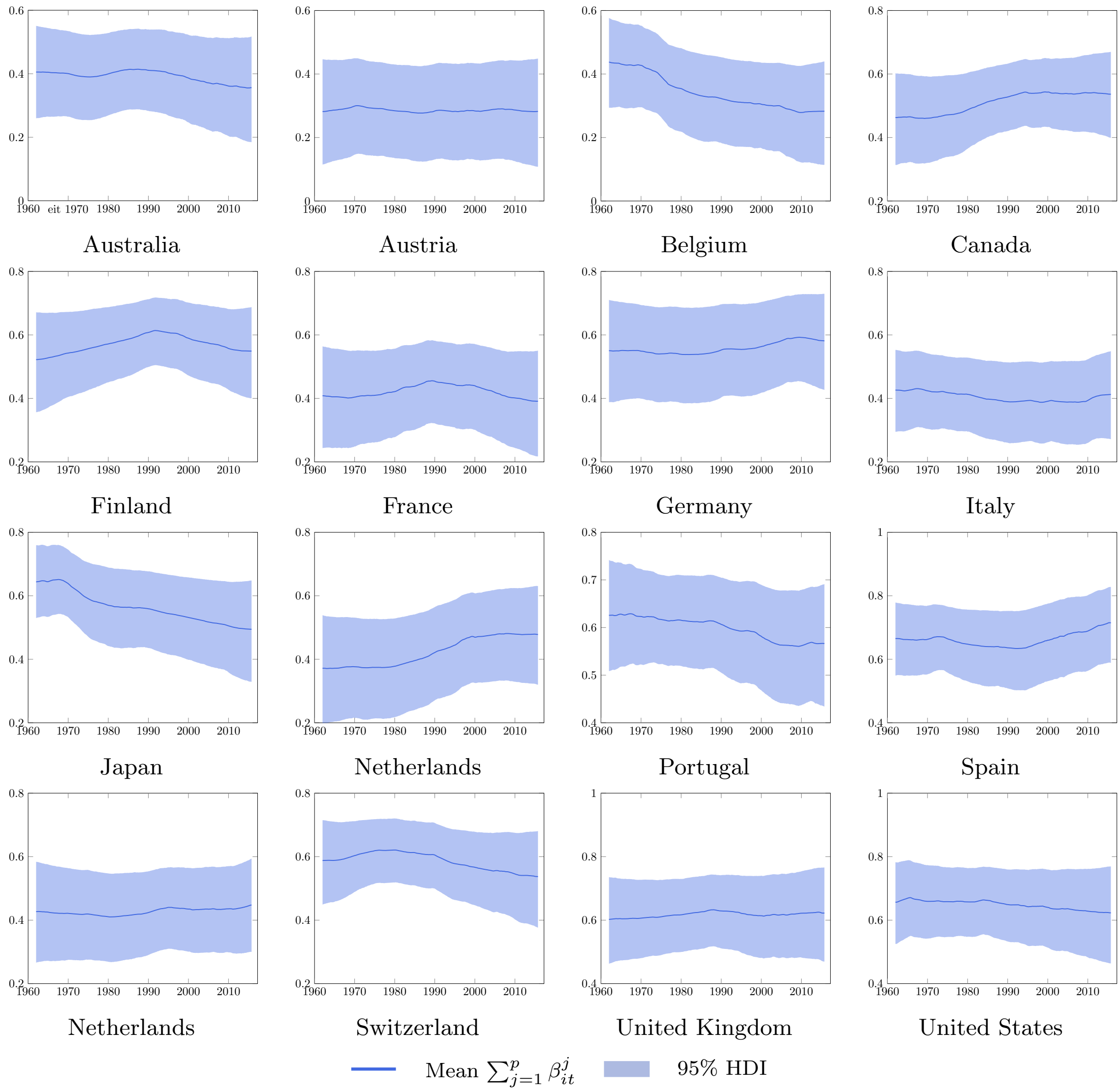
Figure 4: Decomposition of (log)volatility of domestic shocks into common and idiosyncratic part



Propagation of shocks

To get an idea about the time-varying persistence in output growth, Figure 5 plots the sum of the AR coefficients. Based on the results of the stochastic model specification search, only β^1_{it} varies over time while β^2_{it}, β^3_{it} and β^4_{it} are fixed to be constant. Despite some moderate changes, the persistence is relatively stable in most countries, showing that changes in the propagation mechanism in equation (1) are not the main source of the the Great Moderation.

Figure 5: Sum of AR coefficients



3.4 Robustness Checks

In this section, we briefly discuss the outcome of a number of robustness tests. Full results are available from the authors on request.

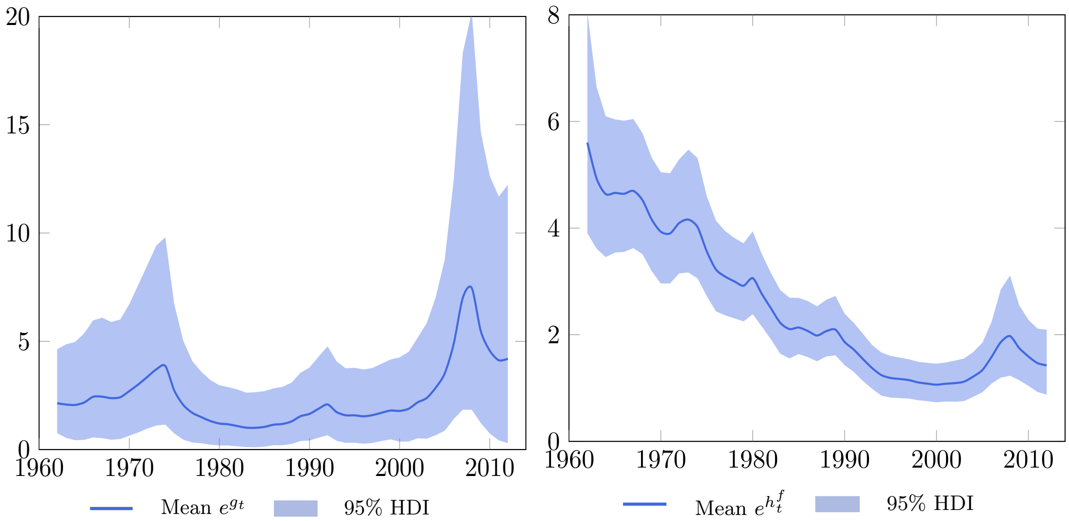
Dropping countries with questionable quality of quarterly data

We start with discussing the robustness of our results with respect to data quality. As mentioned in section 3.1, some countries exhibit rather artificial growth patterns at the beginning of the sample because the OECD constructed the quarterly data from yearly observations. While interpolation methods are used over periods of different length at the beginning of the sample for all countries except Australia, the U.K. and the U.S., the data looks most suspicious for Belgium and Portugal. Hence, we reestimate our model dropping these countries from the sample. While the main results are qualitatively not affected, the correlation in the volatility across countries as well as the persistence of growth rates as measured by the AR coefficients appear to be somewhat sensitive to the cross-sectional dimension of the sample. This is not surprising as excluding countries that, due to interpolation, have less volatile growth rates in the beginning of the sample leads to a strengthening of the cross-country correlation structure. In addition, breaks in the series when the interpolation period ends may affect the AR coefficients implying less evidence of changes in the propagation mechanism when the affected countries are excluded.

Annual data

As an additional robustness check we also use annual data. This greatly improves the quality of the data but also reduces the sample size significantly. Despite the resulting higher estimation uncertainty, Figure 6 demonstrates that the general patterns in the volatility of global shocks and in the common volatility factor still show up. Again no clear trend is visible in the volatility of common shocks, while the volatility of domestic shocks shows a clear common downward trend. For the sake of brevity we do not present results of the tests for time-variation. However, for the two factors plotted in Figure 6 time-variation was found to be relevant whereas this was not the case for the other model components.

Figure 6: Robustness of volatility estimates to annual data



(a) Volatility of common shocks

(b) Common factor in the volatility of domestic shocks

Quarter-on-quarter growth rates

We also test the robustness of our results with respect to the way growth rates are calculated. While we use year-on-year rates in our baseline estimation, other papers rely on annualized quarter-on-quarter rates (e.g. Del Negro and Otrok, 2008; Berger, Grabert, and Kempa, 2016). Overall, results and interpretation remain qualitatively unchanged but we find that the high frequency noise present in quarter-on-quarter data tends to blur the correlation structure in the growth rates, i.e. the role of common shocks and spillovers is found to be smaller. As a result, the correlation structure in the volatility of domestic shocks is even stronger now.

Alternative lag structure

Regarding the lag structure of the model we test a number of different specifications. Our baseline version with four lags of own GDP and one lag of foreign countries GDP follows Stock and Watson (2005). Generally, including less own lags or more lags of foreign GDP does not change the results significantly.

4 Variance decomposition

In this section we use a variance decomposition to illustrate the relative importance of the various model components for explaining the overall evolution in output growth volatility. The variance decomposition approach most commonly used in the dynamic factor model literature (see e.g. Kose, Otrok, and Whiteman, 2003; Del Negro and Otrok, 2008) is based on what is generally known as model-implied variance. This means that for each point in time t, each country i, and each iteration of the Gibbs sampler the model estimates are used to calculate implied (long-run equilibium) variances. Since common and idiosyncratic components are independent by assumption, total variance is additive and the variance shares of interest can be straightforwardly calculated. However, through the inclusion of lagged foreign GDP growth in equation (1), the spillover channel in our empirical specification implies that the different model components are not mutually independent anymore. This requires a slight adjustment of the procedure, as outlined below.

4.1 A simulation-based approach

To calculate the contribution of common shocks, spillovers and idiosyncratic shocks to the overall variance of output growth, first rewrite the model in equations (1)-(2) as follows:

$$\Delta y_{it} - \Delta y_{it}^0 = \Delta y_{it}^1 + \Delta y_{it}^2 + \Delta y_{it}^3, \tag{20}$$

$$\text{with} \quad \Delta y_{it}^0 = \sum_{j=1}^{p} \beta_{it}^j \Delta y_{i,t-j}^0 + \alpha_{it}, \tag{21}$$

$$\Delta y_{it}^1 = \sum_{j=1}^{p} \beta_{it}^j \Delta y_{i,t-j}^1 + \sum_{k=1}^{q} \gamma_{it}^k \Delta y_{i,t-k}^*, \tag{22}$$

$$\Delta y_{it}^2 = \sum_{j=1}^{p} \beta_{it}^j \Delta y_{i,t-j}^2 + \phi_{it}^{\varepsilon} \varepsilon_t^f, \tag{23}$$

$$\Delta y_{it}^3 = \sum_{j=1}^{p} \beta_{it}^j \Delta y_{i,t-j}^3 + \varepsilon_{it}^c. \tag{24}$$

Conditional on the model estimates we can use equations (21)-(24) to calculate Δy_{it}^0 and generate samples for Δy_{it}^1, Δy_{it}^2 and Δy_{it}^3 by (i) drawing ε_t^f and ε_{it}^c from their distributions in equation (7) and (ii) calculating the spillover terms $\Delta y_{i,t-k}^*$ using lagged simulated growth rates.[2] Doing this in each draw of the Gibbs sampler, we obtain $J - B$ simulated samples of Δy_{it} and its constituent components Δy_{it}^0, Δy_{it}^1, Δy_{it}^2 and Δy_{it}^3.

Next, for each component, each country and each point in time we compute the sample variance over the $J - B$ draws. Note that by construction, our spillover component Δy_{it}^1 is not independent of the common and domestic shock components Δy_{it}^2 and Δy_{it}^3. This is because both a global shock f_{t-1} and a domestic shock $\varepsilon_{i,t-1}^c$ will still be present in the components Δy_{it}^2 and Δy_{it}^3 but at the same time feed into the spillover component Δy_{it}^1 through their impact on trade-weighted growth $\Delta y_{i,t-1}^*$. Although the resulting covariance terms $Cov(\Delta y_{it}^1, \Delta y_{it}^2)$ and $Cov(\Delta y_{it}^1, \Delta y_{it}^3)$ are small, we assign them to the spillovers component to make sure that the components' variances sum up to the total variance. Note that the covariance between the common and domestic shock components is zero by assumption. Hence, our variance decomposition is given by

$$\begin{aligned} \underbrace{Var(\Delta y_{it} - \Delta y_{it}^0)}_{\text{Total}} &= \underbrace{Var(\Delta y_{it}^1) + 2Cov(\Delta y_{it}^1, \Delta y_{it}^2) + 2Cov(\Delta y_{it}^1, \Delta y_{it}^3)}_{\text{Spillovers}} \\ &+ \underbrace{Var(\Delta y_{it}^2)}_{\text{Common shocks}} + \underbrace{Var(\Delta y_{it}^3)}_{\text{Idiosyncratic shocks}} . \end{aligned} \tag{25}$$

Our simulation-based approach offers several advantages. First, it allows to separate the contribution of spillovers from that of global and domestic shocks. Second, the dynamics of the model are fully taken into account. This is not the case when calculating model-implied variances as these are typically long-run equilibrium measures ignoring short-run dynamics. Third, by simulating the model in every draw of the Gibbs sampler, parameter uncertainty is explicitly taken into account.

[2] Each of the components was initialized at zero with a burn-in period of 10 quarters. The parameter values used over the burn-in period are set equal to their mean values over the first 5 years of the sample.

4.2 Simulated model-based versus rolling window volatility

To assess the adequacy of our model and simulation approach, Figure 7 plots the simulated total variance of GDP growth along with the commonly used 10-year centered rolling window variance. With a decreasing trend in all countries, starting either at the beginning of the sample or somewhere in the 1970s or 1980s, the Great Moderation clearly shows up in the two alternative volatility measures. However, our model-based approach seems to be much more accurate in the timing of changes. Although the centered rolling window measure is able to pick up the timing of the long-run decline in volatility in most countries, by partly relying on future realized volatility more sudden events like the Great Recession are predated by a number of years. Using an entirely backward-looking window as an alternative will tend to backdate most events. Also note that the rolling window variance remains high(er) at the end of the sample for most countries and hence is not yet able to show that the Great Recession induced only a temporary increases in volatility.

Figure 7: Simulated model-based versus rolling window variance

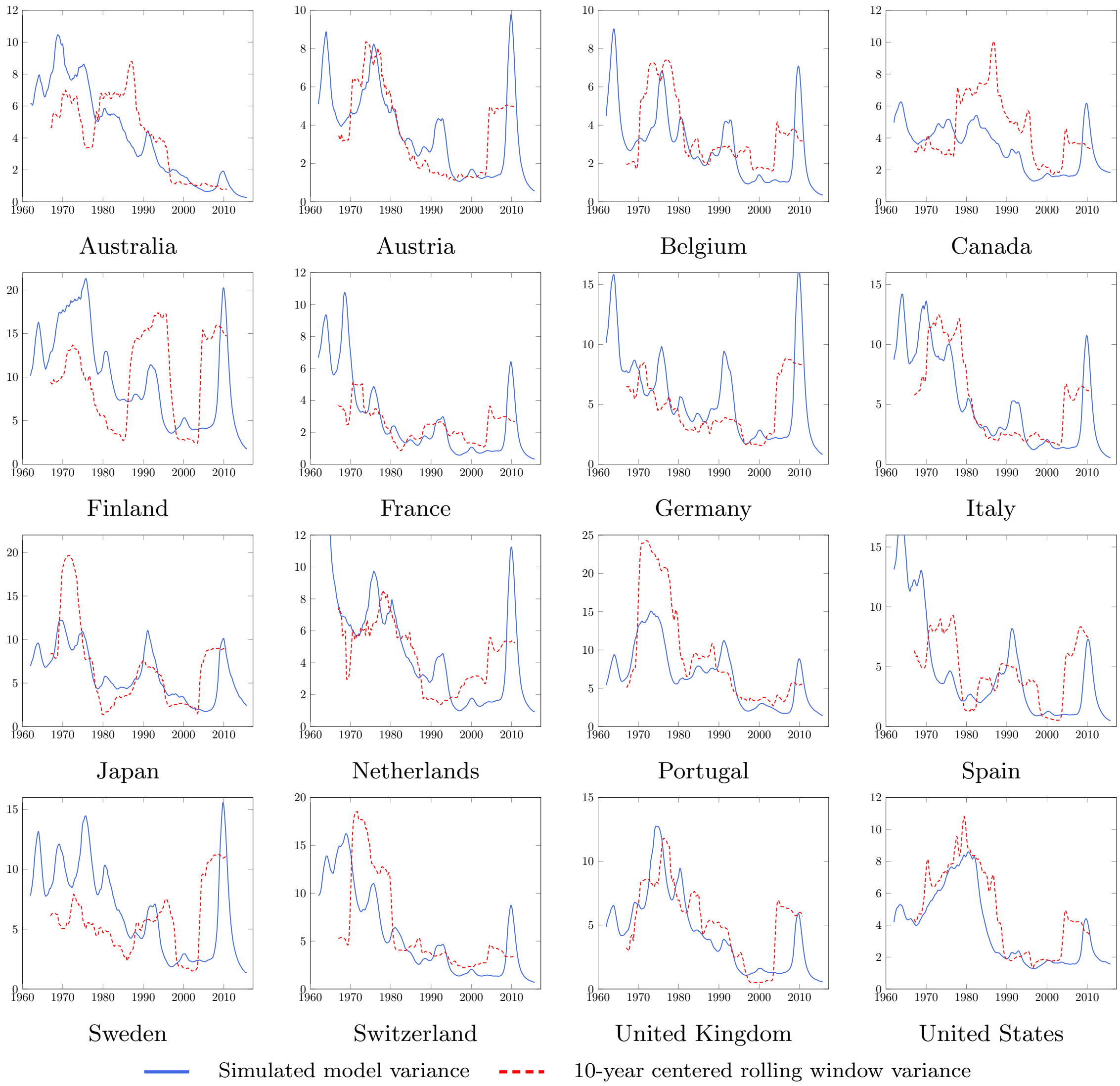
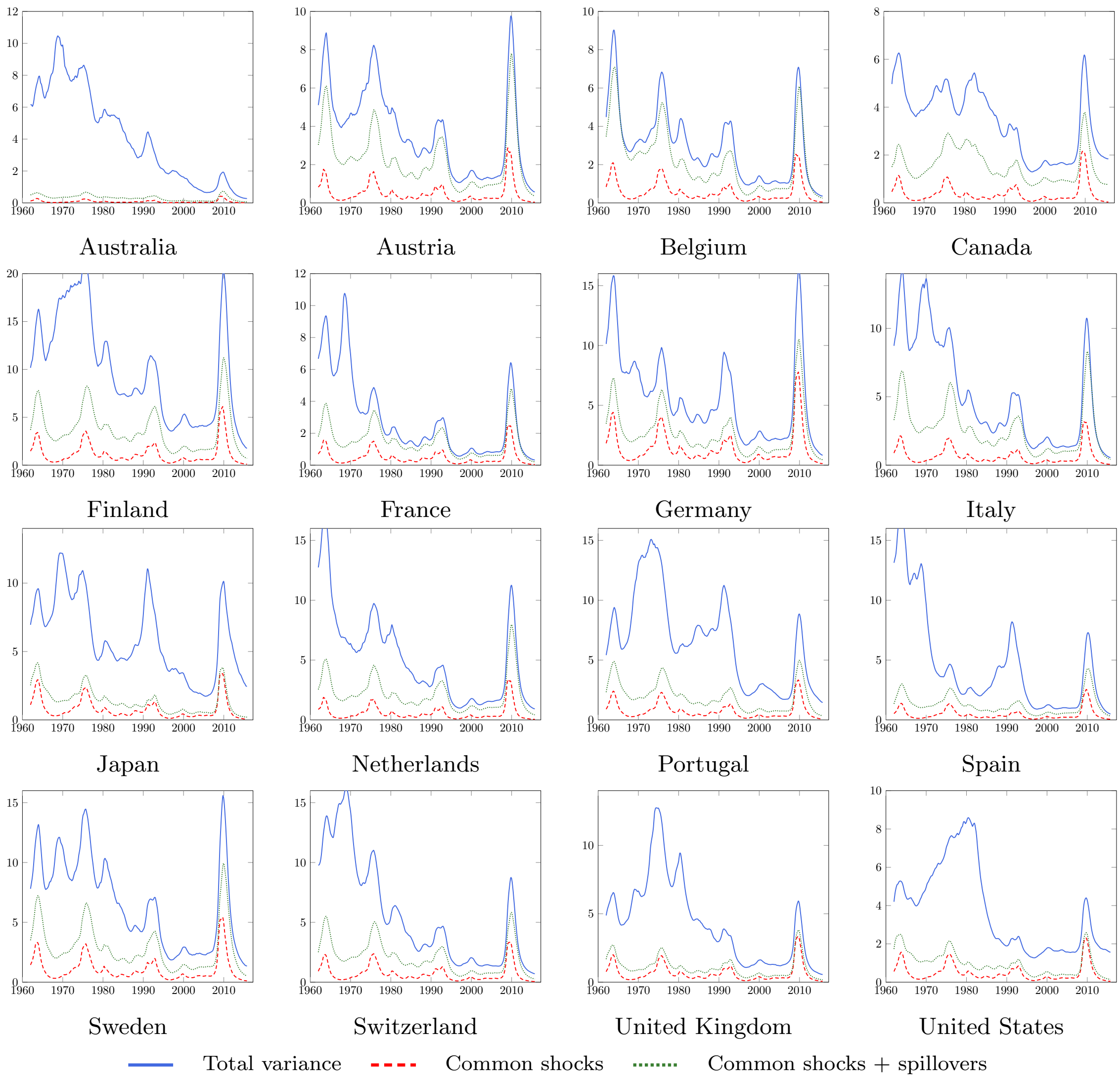


Figure 8: Simulation-based variance decomposition



4.3 Results variance decomposition

Figure 8 decomposes total volatility into the contributions of global shocks, spillovers and domestic shocks. The plots reveal large differences across countries and time with respect to the importance of these three components. First, driven by considerable cross-sectional variation in the sensitivities reported in Table 3, the contribution of common shocks and spillovers differs widely across countries. Output volatility in small open economies like Austria and Belgium is almost entirely driven by these international shocks, leaving only a minor role for domestic shocks. Figure 8 reveals that in these countries the decline in total volatility mainly spills over from the less volatile output growth of their trading partners. To a lesser extent, a similar pattern emerges in Canada and even in larger economies like France, Germany and Italy. At the other end of the spectrum, output volatility in Australia is almost entirely driven by domestic shocks. Also in Japan, Spain, Switzerland, the U.K. and the U.S. international shocks are relatively less important compared to the other countries in the sample. Second, for most countries, a decline in the volatility of domestic shocks is an important source of the overall volatility decline. This is most prominently the case for Australia, Finland, France, Germany, Italy, the Netherlands, Spain, Sweden, Switzerland, the U.K. and the U.S. Third, because of the decline in the volatility of domestic shocks in most countries, international shocks contribute more to overall output volatility towards the end of the sample. This can also be observed from Table 5 where we report the average variance shares for the 1962-1983 and 1984-2015 sub-samples.

Table 5: Average variance shares of the different components over two subsamples (in %)

	Global shocks		Spillovers		Domestic shocks	
	1962-1983	1984-2015	1962-1983	1984-2015	1962-1983	1984-2015
Australia	1.38	4.66	4.75	10.82	93.87	84.52
Austria	9.87	15.20	42.02	53.31	48.11	31.49
Belgium	15.60	20.24	58.90	47.78	25.50	31.98
Canada	8.43	12.33	34.24	38.82	57.33	48.85
Finland	8.97	13.47	22.87	30.32	68.17	56.21
France	14.55	24.79	36.31	51.18	49.14	24.02
Germany	20.15	25.65	24.95	26.46	54.90	47.89
Italy	9.58	17.70	34.82	51.85	55.60	30.46
Japan	12.57	11.74	14.05	8.44	73.38	79.82
Netherlands	7.27	15.85	24.26	46.76	68.48	37.40
Portugal	10.46	13.54	19.91	19.41	69.62	67.05
Spain	10.81	15.81	21.95	29.38	67.24	54.81
Sweden	11.76	19.07	23.74	32.70	64.50	48.23
Switzerland	9.70	19.25	22.00	34.96	68.29	45.79
United Kingdom	10.04	22.60	9.66	15.35	80.30	62.05
United States	9.61	17.40	14.35	13.42	76.04	69.18

5 Conclusion

This paper has investigated the sources of output volatility within a time-varying factor-augmented dynamic panel model with stochastic volatility for a panel of 16 OECD countries over the period 1961:Q1 - 2015:Q4. Our empirical specification allows output growth in a particular country to be driven by global shocks, spillovers and domestic shocks. Changes in the volatility of output growth can stem from a time-varying sensitivity to each of these shocks, changes in the propagation mechanism or shifts in the variances of shocks. As a novel model component we allow for a common factor in the volatility of domestic shocks. We first use a Bayesian stochastic model specification search approach to test for which of the model's components the time variation is actually relevant. The results clearly indicate that both the volatility of global and domestic shocks vary over time. There is some evidence of time variation in the propagation mechanism, while the sensitivities to spillovers and global shocks are found to be constant. Next, we estimate the parsimonious model specification, restricting parameters for which no relevant time variation was found to be constant over the sample period. The results show that although the volatility of global shocks varies over time it does not exhibit a clear downward trend. It mainly reflects periods of worldwide turmoil, shooting up around the oils crises of the 1970s, the worldwide recession of the early 1990s and the recent Great Recession. As individual countries' sensitivity to the common shocks and spillovers has also remained stable over the sample period, changes in the volatility of the international business cycle component is not what is driving the Great Moderation. In contrast, the volatilities of domestic shocks show a clear common downward trend. We identify this as one of the main drivers of the widespread reduction in volatility. We further find that the recent Great Recession induced only a temporary increase in the volatility of common shocks and hence does not mark the end of the Great Moderation.

Our findings have important implications for panel studies explaining output volatility by a set of country-specific regressors. If the unobserved common volatility factor is correlated with the regressors, ignoring this factor will render standard estimation such as fixed effects regressions inconsistent. Therefore, future research on output volatility and its macroeconomic determinants needs to take cross-sectional dependence in the variance equation into account. One possibility is to apply the well-known common correlated effects (CCE) estimation approach of Pesaran (2006) when regressing a volatility measure on its determinants. Alternatively, the model presented here can be augmented by replacing the various random walk components directly by macroeconomic fundamentals.

6 Acknowledgements

The authors would like to thank the members of the Macroeconomics, Policy, and Econometrics Research Group at Ghent University for helpful comments and suggestions. The computational resources (Stevin Supercomputer Infrastructure) and services used in this work were provided by the Flemish Supercomputer Center, funded by Ghent University; the Hercules Foundation; and the Economy, Science, and Innovation Department of the Flemish Government. Martin Iseringhausen gratefully acknowledges financial support from Ghent University's Special Research Fund (BOF).

References

Antolin-Diaz, J., T. Drechsel, and I. Petrella (2016): "Tracking the slowdown in long-run GDP growth," Bank of England working papers 587, Bank of England.

Berger, T., G. Everaert, and H. Vierke (2016): "Testing for time variation in an unobserved components model for the U.S. economy," *Journal of Economic Dynamics and Control*, 69, 179 – 208.

Berger, T., S. Grabert, and B. Kempa (2016): "Global and Country-Specific Output Growth Uncertainty and Macroeconomic Performance," *Oxford Bulletin of Economics and Statistics*, 78(5), 694–716.

Blanchard, O., and J. Simon (2001): "The long and large decline in US output volatility," *Brookings papers on economic activity*, 2001(1), 135–164.

Carare, A., and A. Mody (2012): "Spillovers of Domestic Shocks: Will They Counteract the Great Moderation?," *International Finance*, 15(1), 69–97.

Carter, C. K., and R. Kohn (1994): "On Gibbs sampling for state space models," *Biometrika*, 81(3), 541–553.

Clark, T. E. (2009): "Is the Great Moderation over? An Empirical Analysis," *Economic Review*, 2009, Q4–5.

Del Negro, M., and C. Otrok (2008): "Dynamic factor models with time-varying parameters: measuring changes in international business cycles," Staff Reports 326, Federal Reserve Bank of New York.

Del Negro, M., and G. E. Primiceri (2015): "Time varying structural vector autoregressions and monetary policy: a corrigendum," *The Review of Economic Studies*, 82(4), 1342–1345.

Doran, H. E. (1992): "Constraining Kalman Filter and Smoothing Estimates to Satisfy Time-Varying Restrictions," *The Review of Economics and Statistics*, 74(3), 568–72.

Doyle, B. M., and J. Faust (2002): "An investigation of co-movements among the growth rates of the G-7 countries," *Federal Reserve Bulletin*, (Oct), 427–437.

Durbin, J., and S. J. Koopman (2012): *Time Series Analysis by State Space Methods: Second Edition*, no. 9780199641178 in OUP Catalogue. Oxford University Press.

Dynan, K., D. Elmendorf, and D. Sichel (2006): "Can Financial Innovation Help to Explain the Reduced Volatility of Economic Activity?," *Journal of Monetary Economics*, 53, 123–150.

Fatas, A., and I. Mihov (2001): "Government Size and Automatic Stabilizers: International and Intranational Evidence," *Journal of International Economics*, 55(1), 3–28.

Fernald, J. G. (2007): "Trend breaks, long-run restrictions, and contractionary technology improvements," *Journal of Monetary Economics*, 54(8), 2467–2485.

Fruehwirth-Schnatter, S., and H. Wagner (2010): "Stochastic model specification search for Gaussian and partial non-Gaussian state space models," *Journal of Econometrics*, 154(1), 85–100.

Galí, J., and L. Gambetti (2009): "On the sources of the great moderation," *American Economic Journal: Macroeconomics*, 1(1), 26–57.

George, E. I., and R. E. McCulloch (1993): "Variable selection via Gibbs sampling," *Journal of the American Statistical Association*, 88(423), 881–889.

Herskovic, B., B. Kelly, H. Lustig, and S. V. Nieuwerburgh (2016): "The common factor in idiosyncratic volatility: Quantitative asset pricing implications," *Journal of Financial Economics*, 119(2), 249 – 283.

Jaimovich, N., and H. E. Siu (2009): "The Young, the Old, and the Restless: Demographics and Business Cycle Volatility," *American Economic Review*, 99(3), 804–26.

Kahn, J., M. McConnell, and G. Perez-Quiros (2002): "On the Causes of the Increased Stability of the U.S. Economy," *Federal Reserve Bank of New York, Economic Policy Review*, May, 183–202.

Kim, C.-J., and C. R. Nelson (1999): "Has the US Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle," *Review of Economics and Statistics*, 81(4), 608–616.

Kim, H., H. I. Lee, J. Y. Park, and H. Yeo (2009): "Macroeconomic uncertainty and asset prices: A stochastic volatility model," in *AFA 2010 Atlanta Meetings Paper*.

Kim, S., N. Shephard, and S. Chib (1998): "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *Review of Economic Studies*, 65(3), 361–393.

Kose, M. A., C. Otrok, and C. H. Whiteman (2003): "International Business Cycles: World, Region, and Country-Specific Factors," *American Economic Review*, 93(4), 1216–1239.

——— (2008): "Understanding the evolution of world business cycles," *Journal of International Economics*, 75(1), 110 – 130.

Laurini, M. P., and R. B. Mauad (2015): "A common jump factor stochastic volatility model," *Finance Research Letters*, 12, 2–10.

McConnell, M. M., and G. Perez-Quiros (2000): "Output Fluctuations in the United States: What Has Changed since the Early 1980's?," *American Economic Review*, 90(5), 1464–1476.

Ng, S., and J. H. Wright (2013): "Facts and Challenges from the Great Recession for Forecasting and Macroeconomic Modeling," *Journal of Economic Literature*, 51(4), 1120–54.

Omori, Y., S. Chib, N. Shephard, and J. Nakajima (2007): "Stochastic Volatility With Leverage: Fast and Efficient Likelihood Inference," *Journal of Econometrics*, 140(2), 425–449.

Pesaran, M. H. (2006): "Estimation and inference in large heterogeneous panels with a multifactor error structure," *Econometrica*, 74(4), 967–1012.

Stock, J. H., and M. W. Watson (2003): "Has the Business Cycle Changed and Why?," in *NBER Macroeconomics Annual 2002, Volume 17*, NBER Chapters, pp. 159–230. National Bureau of Economic Research, Inc.

——— (2005): "Understanding Changes In International Business Cycle Dynamics," *Journal of the European Economic Association*, 3(5), 968–1006.

Appendix A Gibbs sampling algorithm

In this appendix we provide details on the Gibbs sampling algorithm used in subsection 2.3 to jointly sample the binary indicators $\mathcal{M}$, the time-varying regression parameters $\lambda_{it} = (\alpha_{it}, \beta_{it}, \gamma_{it})$, the time-varying factor loadings $\phi_{it} = (\phi^{\varepsilon}_{it}, \phi^{h}_{it})$, the stochastic volatilities $\zeta_{it} = (g_t, h^f_t, h^c_{it})$, the common growth shocks ε^f_t, and the mixture indicators s and l. The structure of our Gibbs sampling approach is based on Fruehwirth-Schnatter and Wagner (2010). The description draws heavily from Berger, Everaert, and Vierke (2016) but is extended to account for the cross-sectional dimension.

Block 1: Sampling the binary indicators δ_α, δ_β, δ_γ, the time-varying regression parameters α, β, γ, and the common growth shocks ε^f

For notational convenience, let us define a general regression model

$$w = z^{\mathcal{M}} b^{\mathcal{M}} + e, \qquad\qquad e \sim \mathcal{N}(0, \Sigma), \tag{A-1}$$

where w is a $NT \times 1$ vector including observations on a dependent variable w_{it} stacked over time and cross-sections and z an unrestricted predictor matrix. The corresponding unrestricted parameter vector is denoted b. $z^{\mathcal{M}}$ and $b^{\mathcal{M}}$ are then the restricted predictor matrix and restricted parameter vector that excludes those elements in z and b for which the corresponding indicator in $\mathcal{M}$ is 0. Furthermore, Σ is a diagonal matrix matrix with elements σ^2_{it} that may vary over time to allow for heteroscedasticity of a known form.
As in Fruehwirth-Schnatter and Wagner (2010) we marginalize over the parameters in b when sampling $\mathcal{M}$ and next draw the parameters in b conditional on the indicators $\mathcal{M}$. The posterior distribution can be obtained using Bayes' theorem as

$$f(\mathcal{M}|\lambda, \phi, h, w, x) \propto f(w|\mathcal{M}, \lambda, \phi, h, x) p(\mathcal{M}), \tag{A-2}$$

with $p(\mathcal{M})$ being the prior probability of $\mathcal{M}$ and $f(w|\mathcal{M}, \lambda, \phi, h, x)$ being the marginal likelihood of the regression model (A-1) where the effect of the parameters $b^{\mathcal{M}}$ and σ^2 has been integrated out. The closed form solution of the marginal likelihood depends on whether the error term e_{it} is homoscedastic or heteroscedastic. In our case exclusively the latter case will be relevant. More specifically, in the case of heteroscedasticity, i.e. if $\Sigma = \mathrm{diag}(\sigma^2_{1,t}, ..., \sigma^2_{1,T}, ..., \sigma^2_{N,t}, ..., \sigma^2_{N,T})$ and under the normal conjugate prior $b^{\mathcal{M}} \sim N(a^{\mathcal{M}}_0, A^{\mathcal{M}}_0)$ the closed form solution for the marginal likelihood $f(w|\mathcal{M}, \lambda, \phi, h, x)$ is

$$\begin{aligned} f(w|\mathcal{M}, \lambda, \phi, h, x) \propto \frac{|\Sigma|^{-0.5} |A^{\mathcal{M}}_T|^{0.5}}{|A^{\mathcal{M}}_0|^{0.5}} \exp \Bigg(-\frac{1}{2}(w'\Sigma^{-1}w + (a^{\mathcal{M}}_0)'(A^{\mathcal{M}}_0)^{-1} a^{\mathcal{M}}_0 \\ -(a^{\mathcal{M}}_T)'(A^{\mathcal{M}}_T)^{-1} a^{\mathcal{M}}_T) \Bigg), \end{aligned} \tag{A-3}$$

with

$$a_T^{\mathcal{M}} = A_T^{\mathcal{M}}\left((z^{\mathcal{M}})'\Sigma^{-1}w + (A_0^{\mathcal{M}})^{-1}a_0^{\mathcal{M}}\right), \tag{A-4}$$

$$A_T^{\mathcal{M}} = \left((z^{\mathcal{M}})'\Sigma^{-1}z^{\mathcal{M}} + (A_0^{\mathcal{M}})^{-1}\right)^{-1}. \tag{A-5}$$

Following George and McCulloch (1993), instead of using a multi-move sampler in which all elements in $\mathcal{M}$ are sampled simultaneously, we use a single-move sampler in which each of the binary indicators δ_r (for $r = \alpha, \beta, \gamma$) in $\mathcal{M}$ is sampled from $f(\delta_r|\delta_{/r}, \lambda, \phi, h, w, x)$. Given these general definitions, Block 1 of the Gibbs sampling scheme splits up as follows:

Block 1(a): Sampling the binary indicators δ_α, δ_β, and δ_γ

In this block we sample the binary indicators δ_r (for $r = \alpha, \beta, \gamma$) conditional on the states λ, ϕ, and ζ. Using equation (19), the mean equation (1) can be written in the general linear regression format (A-1) where a single observation w_{it} is defined as

$$\underbrace{y_{it} - \phi_{it}^{\varepsilon} e^{g_t/2}\varepsilon_t^f}_{w_{it}} = \underbrace{\begin{bmatrix}1 & y_{i,t-1} & y_{i,t-1}^* & \widetilde{\alpha}_{it} & \widetilde{\beta}_{it} y_{i,t-1} & \widetilde{\gamma}_{it} y_{i,t-1}^*\end{bmatrix}}_{z_{it}^{\mathcal{M}}} \underbrace{\begin{bmatrix}\alpha_{i0}\\ \beta_{i0}\\ \gamma_{i0}\\ \sigma_\alpha\\ \sigma_\beta\\ \sigma_\gamma\end{bmatrix}}_{b_{it}^{\mathcal{M}}} + \varepsilon_{it}^c, \tag{A-6}$$

where in both the restricted vector $z_{it}^{\mathcal{M}}$ and the restricted parameter vector $b_{it}^{\mathcal{M}}$ some or all of the elements corresponding to the shock standard deviations are excluded depending on the indicator values $\delta = (\delta_\alpha, \delta_\beta, \delta_\gamma)$. The covariance matrix Σ contains the heteroscedastic error variances obtained from the stochastic volatility process in equation (8). The marginal likelihood $f(w|\delta, \lambda, \phi, h, x)$ can than be calculated as in (A-3). Finally, the binary indicators $\delta = (\delta_\alpha, \delta_\beta, \delta_\gamma)$ can be sampled from the Bernoulli distribution with probability

$$p(\delta_r = 1|\alpha, \beta, \gamma, w, x) = \frac{f(\delta_r = 1|\alpha, \beta, \gamma, w, x)}{f(\delta_r = 0|\alpha, \beta, \gamma, w, x) + f(\delta_r = 1|\alpha, \beta, \gamma, w, x)}. \tag{A-7}$$

If a certain indicator is 0, the corresponding standard deviation will be set to 0. Otherwise, the standard deviation will be sampled from the posterior as described in the next subblock.

Block 1(b): Sampling the time-invariant parameters α_0, β_0, γ_0, σ_α, σ_β, σ_γ

In this subblock we sample the time-invariant parts of the parameters α_{i0}, β_{i0}, γ_{i0} and the (unrestricted) shock standard deviations σ_α, σ_β, σ_α from their respective posterior distribution. We use the regression model outlined in (A-6). Given the normal conjugate prior $N \sim (a_0^{\mathcal{M}}, A_0^{\mathcal{M}})$ the parameters are sampled from the posterior $N \sim (a_T^{\mathcal{M}}, A_T^{\mathcal{M}})$ whereas the posterior moments are given by A-4 and A-5, respectively.

Block 1(c): Sampling the states $\widetilde{\alpha}$, $\widetilde{\beta}$, $\widetilde{\alpha}$, and the common growth shocks f

In this block we use the forward-filtering and back-sampling approach of Carter and Kohn (1994) to sample the states $\widetilde{\alpha}$, $\widetilde{\beta}$, $\widetilde{\gamma}$. To this end, we employ a general state space model of the following form as given in Durbin and Koopman (2012)

$$w_t = Z_t \kappa_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, H_t), \tag{A-8}$$

$$\kappa_{t+1} = T_t \kappa_t + R_t \eta_t, \qquad \eta_t \sim \mathcal{N}(0, Q_t), \tag{A-9}$$

where w_t again is a vector of observations (stacked over all cross-sections) and κ_t an unobserved state vector. The matrices Z_t, T_t, H_t, Q_t are assumed to be known (conditioned upon). The error terms ϵ_t and η_t are assumed to be serially uncorrelated and independent of each other at all points in time. As equations (A-8) and (A-9) constitute a linear Gaussian state space model, the unknown state variables κ_t can be filtered using the standard Kalman filter. Sampling $\kappa = [\kappa_1, ..., \kappa_T]$ can then be done using the algorithm outlined in Carter and Kohn (1994). To be more specific, the state space model to be estimated in this block is

$$\underbrace{(y_t - \alpha_0 - \beta_0 y_{t-1} - \gamma_0 y^*_{t-1})}_{w_t} = \underbrace{\begin{bmatrix} \sigma_\alpha I_N & \sigma_\beta I_N y_{t-1} & \sigma_\gamma I_N y^*_{t-1} & \phi^\varepsilon_t e^{g_t/2} \end{bmatrix}}_{Z_t} \underbrace{\begin{bmatrix} \widetilde{\alpha}_t \\ \widetilde{\beta}_t \\ \widetilde{\gamma}_t \\ \varepsilon^f_t \end{bmatrix}}_{\kappa_t} + \varepsilon^c_t, \tag{A-10}$$

$$\underbrace{\begin{bmatrix} \widetilde{\alpha}_{t+1} \\ \widetilde{\beta}_{t+1} \\ \widetilde{\gamma}_{t+1} \\ \varepsilon^f_{t+1} \end{bmatrix}}_{\kappa_{t+1}} = \underbrace{\begin{bmatrix} I_N & 0 & 0 & 0 \\ 0 & I_N & 0 & 0 \\ 0 & 0 & I_N & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{T_t} \underbrace{\begin{bmatrix} \widetilde{\alpha}_t \\ \widetilde{\beta}_t \\ \widetilde{\gamma}_t \\ \varepsilon^f_t \end{bmatrix}}_{\kappa_t} + \underbrace{\begin{bmatrix} I_N & 0 & 0 & 0 \\ 0 & I_N & 0 & 0 \\ 0 & 0 & I_N & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{R_t} \underbrace{\begin{bmatrix} \eta^{\widetilde{\alpha}}_t \\ \eta^{\widetilde{\beta}}_t \\ \eta^{\widetilde{\gamma}}_t \\ \eta^{\varepsilon^f}_t \end{bmatrix}}_{\eta_t}, \tag{A-11}$$

$$H_t = I_N e^{(\phi^h_t h^f_t + h^c_t)}, \qquad Q_t = I_{3N+1}, \tag{A-12}$$

where I_N denotes the identity matrix of dimension N. Finally, after sampling the time-varying states we perform a random sign switch on σ_α and $\widetilde{\alpha}$, σ_β and $\widetilde{\beta}$, and σ_γ and $\widetilde{\gamma}$ as suggested by Fruehwirth-Schnatter and Wagner (2010).

Block 2: Sampling the binary indicator δ_g, the stochastic volatility of common shocks g, and the mixture indicator l

Consider the following stochastic volatility model

$$y_{it} = e^{h_{it}/2} \epsilon_{it}, \qquad \epsilon_{it} \sim \mathcal{N}(0, 1), \tag{A-13}$$

where h_{it} denotes the log-variance process which is assumed to evolve according to a random walk process. Following Kim, Shephard, and Chib (1998) this expression can be linearized by taking the natural-log of the squares

$$ln(y_{it}^2 + c) = h_{it} + \widetilde{\epsilon}_{it}, \tag{A-14}$$

where $c = 0.001$ is an offset constant and $\widetilde{\epsilon}_{it} = \ln(\epsilon_{it}^2)$. The last term follows a log-chi-square distribution and can be approximated by a mixture of normal distributions as follows

$$f(\widetilde{\epsilon}_{it}) = \sum_{j=1}^{M} q_j f_{\mathcal{N}}(\widetilde{\epsilon}_{it}|m_j - 1.2704, v_j^2), \tag{A-15}$$

where q_j is the component probability of a specific normal distribution with mean $m_j - 1.2704$ and variance v_j^2. This mixture can equivalently by expressed in terms of component probabilities

$$\widetilde{\epsilon}_{it}|(l_{it} = j) \sim \mathcal{N}(m_j - 1.2704, v_j^2), \quad \text{with} \quad \Pr(l_{it} = j) = q_j. \tag{A-16}$$

We follow Omori, Chib, Shephard, and Nakajima (2007) and use a mixture of $M = 10$ normal distributions to proxy the log-chi-square distribution. With this linearization at hand, block 2 of the Gibbs sampling algorithm consists of the following steps.

Block 2(a): Sampling the mixture indicator l

Following Del Negro and Primiceri (2015), the mixture indicators are sampled before the stochastic volatility component. Specifically, we use the general expression (A-14) where the index i drops as the component is common across cross-sections and sample the indicator from its conditional probability mass

$$p(l_t = j|g_t, \widetilde{\epsilon}_t) \propto q_j f_{\mathcal{N}}(\widetilde{\epsilon}_t|g_t + m_j - 1.2704, v_j^2), \tag{A-17}$$

with the values for q_j, m_j, and v_j^2 taken from table 1 in Omori, Chib, Shephard, and Nakajima (2007). It should be noted that their reported expected values m_j are already adjusted for the mean of the log-chi-squared distribution 1.2704.

Block 2(b): Sampling the binary indicators δ_g

The binary indicator δ_g is drawn from a Bernoulli distribution as described above whereas the general regression model in (A-1) now takes the following form

$$\underbrace{\left(ln\left((e^{g_t/2}\varepsilon_t^f)^2 + 0.001\right) - (m_t^l - 1.2704)\right)}_{w_t} = \underbrace{\begin{bmatrix}1 & \widetilde{g}_t\end{bmatrix}}_{z^{\mathcal{M}}} \underbrace{\begin{bmatrix}g_0 \\ \sigma_g\end{bmatrix}}_{b^{\mathcal{M}}} + \widetilde{\epsilon}_t. \tag{A-18}$$

Block 2(c): Sampling the time-invariant parameters h_0^f and σ_{h^f}

In this subblock we sample the time-invariant part of the stochastic volatility process g_t, g_0 and the (unrestricted) shock standard deviation σ_g from their respective posterior distributions. We use the regression model outlined in (A-18). In this case, heteroscedasticity of the error arises due to the normal mixture approximation.

Block 2(d): Sampling the state $\widetilde{g}$

In order to sample the time-varying state $\widetilde{g}_t$, the following state space model is set up

$$\underbrace{\left(ln\left((e^{g_t/2}\varepsilon_t^f)^2+0.001\right)-(m_t^l-1.2704)\right)-g_0}_{w_t}=\underbrace{\begin{bmatrix}\sigma_g\end{bmatrix}}_{Z_t}\underbrace{\begin{bmatrix}\widetilde{g}_t\end{bmatrix}}_{\kappa_t}+\epsilon_t, \tag{A-19}$$

$$\underbrace{\widetilde{g}_{t+1}}_{\kappa_{t+1}}=\underbrace{\begin{bmatrix}1\end{bmatrix}}_{T_t}\underbrace{\begin{bmatrix}\widetilde{g}_t\end{bmatrix}}_{\kappa_t}+\underbrace{\begin{bmatrix}1\end{bmatrix}}_{R_t}\underbrace{\begin{bmatrix}\eta_t^{\widetilde{g}}\end{bmatrix}}_{\eta_t}, \tag{A-20}$$

$$H_t=v_t^l, \qquad Q_t=1, \tag{A-21}$$

where estimates are again obtained using the forward-filtering and backward-sampling approach. A random sign switch is performed on σ_g and $\widetilde{g}$.

Block 3: Sampling the binary indicators δ_{h^c}, δ_{ϕ^h}, the stochastic volatility h^c, the time-varying factor loadings ϕ^h, and the mixture indicators s

In order to implement the normalizations on ϕ_{it}^h and h_{it}^c as described in section 2.2, we follow the approach outlined in Doran (1992) which augments the Kalman filter such that the estimates satisfy chosen restrictions. However, for the sake of brevity, the following description only outlines the general estimation procedure but does not elaborate on the normalizations.

Block 3(a): Sampling the mixture indicators s

Sampling the mixture indicators s is done as shown before. It should be noted, that the mixture indicators s_{it} of the idiosyncratic stochastic volatility process vary over cross-sections.

Block 3(b): Sampling the binary indicators δ_{ϕ^h} and δ_{h^c}

The binary indicators δ_{h^c} and δ_{ϕ^h} are drawn from a Bernoulli distribution as described above whereas the general regression model in (A-1) now takes the following form for a single observation

$$\widetilde{y}_{it} = y_{it} - \alpha_{it} - \beta_{it} y_{i,t-1} - \gamma_{it} y^*_{i,t-1} - \phi^{\varepsilon}_{it} e^{g_t/2} \varepsilon^f_t, \tag{A-22}$$

$$\underbrace{ln(\widetilde{y}^2_{it} + 0.001)}_{w_{it}} = \underbrace{\begin{bmatrix} h^f_t & 1 & \widetilde{\phi}^h_{it} h^f_t & \widetilde{h^c}_{it} \end{bmatrix}}_{z^{\mathcal{M}}} \underbrace{\begin{bmatrix} \phi^h_{i0} \\ h^c_{i0} \\ \sigma_{\phi^h} \\ \sigma_{h^c} \end{bmatrix}}_{b^{\mathcal{M}}} + \widetilde{\epsilon}_{it}. \tag{A-23}$$

Block 3(c): Sampling the time-invariant parameters ϕ^h_0, h^c_0, σ_{ϕ^h}, and σ_{h^c}

In this subblock we sample the time-invariant parts of the idiosyncratic stochastic volatility process h^c_{i0} and the factor loadings ϕ^h_{i0} as well as the (unrestricted) shock standard deviations σ_{h^c} and σ_{ϕ^h} from their respective posterior distributions. For this purpose, the regression model in (A-22) is used.

Block 3(d): Sampling the states $\widetilde{\phi}^h$ and $\widetilde{h}^c$

The linear state space model employed in this block to sample the time-varying states $\widetilde{\phi}^h_{it}$ and $\widetilde{h}^c_{it}$ can be written as

$$\widetilde{y}_t = y_t - \alpha_t - \beta_t y_{t-1} - \gamma_t y^*_{t-1} - \phi^{\varepsilon}_t e^{g_t/2} \varepsilon^f_t, \tag{A-24}$$

$$\underbrace{\left(ln(\widetilde{y}_t{}^2 + 0.001) - (m^s_t - 1.2704) - \phi^h_0 h^f_t - h^c_0\right)}_{w_t} = \underbrace{\begin{bmatrix} \sigma_{\phi^h} h^f_t I_N & \sigma_{h^c} I_N \end{bmatrix}}_{Z_t} \underbrace{\begin{bmatrix} \widetilde{\phi}^h_t \\ \widetilde{h}^c_t \end{bmatrix}}_{\kappa_t} + \epsilon_t, \tag{A-25}$$

$$\underbrace{\begin{bmatrix} \widetilde{\phi}^h_{t+1} \\ \widetilde{h}^c_{t+1} \end{bmatrix}}_{\kappa_{t+1}} = \underbrace{\begin{bmatrix} I_N & 0 \\ 0 & I_N \end{bmatrix}}_{T_t} \underbrace{\begin{bmatrix} \widetilde{\phi}^h_t \\ \widetilde{h}^c_t \end{bmatrix}}_{\kappa_t} + \underbrace{\begin{bmatrix} I_N & 0 \\ 0 & I_N \end{bmatrix}}_{R_t} \underbrace{\begin{bmatrix} \eta^{\widetilde{\phi}^h}_t \\ \eta^{\widetilde{h}^c}_t \end{bmatrix}}_{\eta_t}, \tag{A-26}$$

$$H_t = v^s_t, \qquad Q_t = I_{2N}. \tag{A-27}$$

After sampling the states, a random sign switch is performed on σ_{ϕ^h} and $\widetilde{\phi}^h$, and σ_{h^c} and $\widetilde{h}^c$.

Block 4: Sampling the binary indicator δ_{h^f} and the common volatility factor h^f

Block 4(a): Sampling the binary indicator δ_{h^f}

The binary indicator δ_{h^f} is drawn from a Bernoulli distribution as described above whereas the general regression model in (A-1) reads

$$\widetilde{y}_{it} = y_{it} - \alpha_{it} - \beta_{it} y_{i,t-1} - \gamma_{it} y^*_{i,t-1} - \phi^{\varepsilon}_{it} e^{g_t/2} \varepsilon^f_t, \tag{A-28}$$

$$\underbrace{\left(ln\left((\widetilde{y}^2_{it} + 0.001\right) - (m^s_{it} - 1.2704) - h^c_{it}\right)}_{w_{it}} = \underbrace{\begin{bmatrix}\phi^h_{it} & \phi^h_{it}\widetilde{h}^f_t\end{bmatrix}}_{Z^{\mathcal{M}}} \underbrace{\begin{bmatrix} h^f_0 \\ \sigma_{h^f}\end{bmatrix}}_{b^{\mathcal{M}}} + \widetilde{\epsilon}_{it}. \tag{A-29}$$

Block 4(b): Sampling the time-invariant parameter h^f_0 and σ_{h^f}

In this subblock we sample the time-invariant part of the common volatility factor h^f_0 and the (unrestricted) shock standard deviations σ_{h^f}. We use the regression model outlined in (A-28).

Block 4(c): Sampling the state $\widetilde{h}^f$

The linear state space model used in this block to sample the time-varying state $\widetilde{h}^f$ is

$$\widetilde{y}_t = y_t - \alpha_t - \beta_t y_{t-1} - \gamma_t y^*_{t-1} - \phi^{\varepsilon}_t e^{g_t/2} \varepsilon^f_t, \tag{A-30}$$

$$\underbrace{\left(ln\left((\widetilde{y}^2_t + 0.001\right) - (m^s_t - 1.2704) - h^c_t - \phi^h_t h^f_0\right)}_{w_t} = \underbrace{\begin{bmatrix}\phi^h_t \sigma_{h^f}\end{bmatrix}}_{Z_t} \underbrace{\begin{bmatrix}\widetilde{h}^f_t\end{bmatrix}}_{\kappa_t} + \epsilon_t, \tag{A-31}$$

$$\underbrace{\begin{bmatrix}\widetilde{h}^f_{t+1}\end{bmatrix}}_{\kappa_{t+1}} = \underbrace{\begin{bmatrix}1\end{bmatrix}}_{T_t} \underbrace{\begin{bmatrix}\widetilde{h}^f_t\end{bmatrix}}_{\kappa_t} + \underbrace{\begin{bmatrix}1\end{bmatrix}}_{R_t} \underbrace{\begin{bmatrix}\eta^{\widetilde{h}^f}_t\end{bmatrix}}_{\eta_t}, \tag{A-32}$$

$$H_t = v^s_t, \qquad Q_t = 1. \tag{A-33}$$

Again, a random sign switch is performed on σ_{h^f} and $\widetilde{h}^f$

Block 5: Sampling the binary indicators $\delta_{\phi^{\varepsilon}}$ and the factor loadings of common growth shocks ϕ^{ε}

In order to implement the normalization on ϕ^{ε}_{it} as described in section 2.2, we follow the approach outlined in Doran (1992) which augments the Kalman filter such that the estimates satisfy chosen restrictions. However, for the sake of brevity, the following description only outlines the general estimation procedure but does not elaborate on the normalizations.

Block 5(a): Sampling the binary indicator $\delta_{\phi^\varepsilon}$

The binary indicator is once again sampled from the Bernoulli distribution using the procedure outlined above where the general regression model in (A-1) now takes the form

$$\underbrace{y_{it} - \alpha_{it} - \beta_{it} y_{i,t-1} - \gamma_{it} y^*_{i,t-1}}_{w_{it}} = \underbrace{\begin{bmatrix} e^{g_t/2}\varepsilon^f_t & e^{g_t/2}\varepsilon^f_t \widetilde{\phi}^\varepsilon_{it} \end{bmatrix}}_{z^{\mathcal{M}}} \underbrace{\begin{bmatrix} \phi^\varepsilon_{i0} \\ \sigma_{\phi^\varepsilon} \end{bmatrix}}_{b^{\mathcal{M}}} + \varepsilon^c_{it}. \tag{A-34}$$

As in block 1, heteroscedasticity of the errors arises due to the stochastic volatility process of idiosyncratic shocks.

Block 5(b): Sampling the time-invariant parameter ϕ^ε_0 and $\sigma_{\phi^\varepsilon}$

In this subblock we sample the time-invariant part of the common volatility factor ϕ^ε_{i0} and the (unrestricted) shock standard deviations $\sigma_{\phi^\varepsilon}$. We use the regression model in (A-34).

Block 5(c): Sampling the state $\widetilde{\phi}^\varepsilon$

The linear state space model used in this block to sample the time-varying state $\widetilde{\phi}^\varepsilon$ takes the following form

$$\underbrace{(y_t - \alpha_t - \beta_t y_{t-1} - \gamma y^*_{t-1} - \phi^\varepsilon_0 e^{g_t/2}\varepsilon^f_t)}_{w_t} = \underbrace{\begin{bmatrix} \sigma_{\phi^\varepsilon} e^{g_t/2}\varepsilon^f_t I_N \end{bmatrix}}_{Z_t} \underbrace{\begin{bmatrix} \widetilde{\phi}^\varepsilon_t \end{bmatrix}}_{\kappa_t} + \varepsilon^c_t, \tag{A-35}$$

$$\underbrace{\begin{bmatrix} \widetilde{\phi}^\varepsilon_{t+1} \end{bmatrix}}_{\kappa_{t+1}} = \underbrace{\begin{bmatrix} 1 \end{bmatrix}}_{T_t} \underbrace{\begin{bmatrix} \widetilde{\phi}^\varepsilon_t \end{bmatrix}}_{\kappa_t} + \underbrace{\begin{bmatrix} 1 \end{bmatrix}}_{R_t} \underbrace{\begin{bmatrix} \eta^{\widetilde{\phi}^\varepsilon}_t \end{bmatrix}}_{\eta_t}, \tag{A-36}$$

$$H_t = I_N e^{(\phi^h_t h^f_t + h^c_t)}, \qquad Q_t = 1. \tag{A-37}$$

Finally, a random sign switch is performed on $\sigma_{\phi^\varepsilon}$ and $\widetilde{\phi}^\varepsilon$.