JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. **29**: 1031–1052 (2014)
Published online 28 May 2014 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/jae.2395

RARE SHOCKS, GREAT RECESSIONS

VASCO CÚRDIA,[a] MARCO DEL NEGRO[b]* AND DANIEL L. GREENWALD[c]

[a] *Federal Reserve Bank of San Francisco, CA, USA*
[b] *Federal Reserve Bank of New York, NY, USA*
[c] *New York University, NY, USA*

SUMMARY

We estimate a DSGE (dynamic stochastic general equilibrium) model where rare large shocks can occur, by replacing the commonly used Gaussian assumption with a Student's t-distribution. Results from the Smets and Wouters (*American Economic Review* 2007; **97**: 586–606) model estimated on the usual set of macroeconomic time series over the 1964–2011 period indicate that (i) the Student's t specification is strongly favored by the data even when we allow for low-frequency variation in the volatility of the shocks, and (ii)) the estimated degrees of freedom are quite low for several shocks that drive US business cycles, implying an important role for rare large shocks. This result holds even if we exclude the Great Recession period from the sample. We also show that inference about low-frequency changes in volatility—and, in particular, inference about the magnitude of Great Moderation—is different once we allow for fat tails. Copyright © 2014 John Wiley & Sons, Ltd.

Received 16 December 2013; Revised 12 March 2014

Supporting information may be found in the online version of this article.

1. INTRODUCTION

Great Recessions do not happen every decade—this is why they are dubbed 'great' in the first place. To the extent that dynamic stochastic general equilibrium (DSGE) models rely on shocks in order to generate macroeconomic fluctuations, they may need to account for the occurrence of rare but very large shocks to generate an event like the Great Recession. For this reason, we estimate a linearized DSGE model assuming that shocks are generated from a Student's t-distribution, which is designed to capture fat tails, in place of a Gaussian distribution, which is the standard assumption in the DSGE literature. The number of degrees of freedom in the Student's t-distribution, which determines the likelihood of observing rare large shocks (and which we allow to vary across shocks), is estimated from the data.

We show that estimating DSGE models with Student's t-distributed shocks is a fairly straightforward extension of current methods (described, for instance, in An and Schorfheide, 2007). In fact, the Gibbs sampler is a simple extension of Geweke's (1993) Gibbs sampler for a linear model to the DSGE framework. The paper is closely related to Chib and Ramamurthy (2014) who, in independent and contemporaneous work, also propose a similar approach to the one developed here for estimating DSGE models with Student's t-distributed shocks. One difference between this work and that of Chib and Ramamurthy (2014) is that we also allow for low-frequency changes in the volatility of the shocks, in light of the evidence provided by several papers in the DSGE literature (Justiniano and Primiceri, 2008; Fernández-Villaverde and Rubio-Ramírez, 2007; Liu *et al.*, 2011; among others).[1]

* Correspondence to: Marco Del Negro, Federal Reserve Bank of New York, 33 Liberty Street, New York, NY 10045, USA. E-mail: marco.delnegro@ny.frb.org

[1] Another difference between this paper and Chib and Ramamurthy (2014) is that Chib and Ramamurthy use a simple three-equation New Keynesian model, while we use the full Smets and Wouters (2007) model.

We specifically follow the approach in Justiniano and Primiceri (2008; henceforth JP), who postulate a random walk as the law of motion of the log volatilities. We show that ignoring low-frequency movements in volatility biases the results toward finding evidence in favor of fat tails.

We apply our methodology to the Smets and Wouters (2007) model (henceforth SW), estimated on the same seven macroeconomic time series used in SW. Our baseline dataset starts in 1964:Q4 and ends in 2011:Q1, but we also consider a subsample ending in 2004:Q4 to analyze the extent to which our findings depend on the inclusion of the Great Recession in our sample. We use the SW model both because it is a prototypical medium-scale DSGE model and because its empirical success has been widely documented.[2] Models that fit the data poorly will necessarily have large shocks. We therefore chose a DSGE model that is at the frontier in terms of empirical performance to assess the extent to which macro variables have fat tails.

The motivation for our work arises from evidence such as that displayed in Figure 1. The top two panels of Figure 1 show the time series of the smoothed 'discount rate' and 'marginal efficiency of investment' shocks (in absolute value) from the SW model estimated under Gaussianity. The shocks are normalized, so that they are expressed in standard deviation units. The solid line is the median, and the dashed lines are the posterior 90% bands. The figure shows that the size of the shocks is above 3.5 standard deviations on several occasions, one of which is the recent recession. The probability of observing such large shocks under Gaussianity is very low. In addition to this DSGE model-based evidence, existing literature shows that the unconditional distribution of macro variables is not Gaussian (for example, see Christiano, 2007, for pre-Great Recession evidence and Ascari *et al.* (2012) for more recent work).

While the visual evidence against Gaussianity is strong, there are several reasons to perform a more careful quantitative analysis. First, these shocks are obtained under the counterfactual assumption of Gaussianity, i.e. using a misspecified model. Second, a quantitative estimate of the fatness of the tails is an obvious object of interest. Third, it is important to disentangle the relative contribution of fat tails from that of (slow-moving) time-varying volatility. The bottom panel of Figure 1, which shows the evolution of the smoothed monetary policy shocks estimated under Gaussianity (again, normalized and in absolute value), provides a case in point, as the clustering of large shocks in the late 1970s and 1980s is quite evident.

In general, studying the importance of fat tails by looking only at the kurtosis in the unconditional distribution of either macro variables (as in Ascari *et al.*, 2012) or of estimated shocks can be misleading, because the evidence against Gaussianity can be due to low-frequency changes in volatility. Conversely, the presence of large shocks can potentially distort the assessment of low-frequency movements in volatility. To see this, imagine estimating a model that allows only for slow-moving time variation in volatility, but no fat tails, in the presence of shocks that fit the pattern shown in the top panel of Figure 1. As the stochastic volatility will try to fit the squared residuals, such a model may produce a time series of volatilities peaking around 1980, and then again during the Great Recession. Put differently, very large shocks may be interpreted as persistent changes in volatility, when they may in fact be rare realizations from a process with a time-invariant distribution. For instance, the extent to which the Great Recession can be interpreted as a permanent rise in macroeconomic volatility may depend on whether we allow for rare large shocks.

Finally, we expect that the evidence provided in this paper will be further motivation for the study of nonlinear models. First, if shocks have fat tails, linearization may simply produce a poor approximation of the full model. Second, nonlinearities may explain away the fat tails: what we capture as large

[2] The forecasting performance of the SW model was found to be competitive in terms of accuracy relative to private forecasters and reduced-form models not only during the Great Moderation period (see Smets and Wouters, 2007; and Edge and Gürkaynak, 2010), but also including data for the Great Recession (Del Negro and Schorfheide, 2014). In earlier drafts of this paper we used a DSGE model with financial frictions along the lines of Bernanke *et al.* (1999) and Christiano *et al.* (2009) and found strong evidence in favor of Student's t-distributed shocks in that setting as well.

DOI: 10.1002/jae

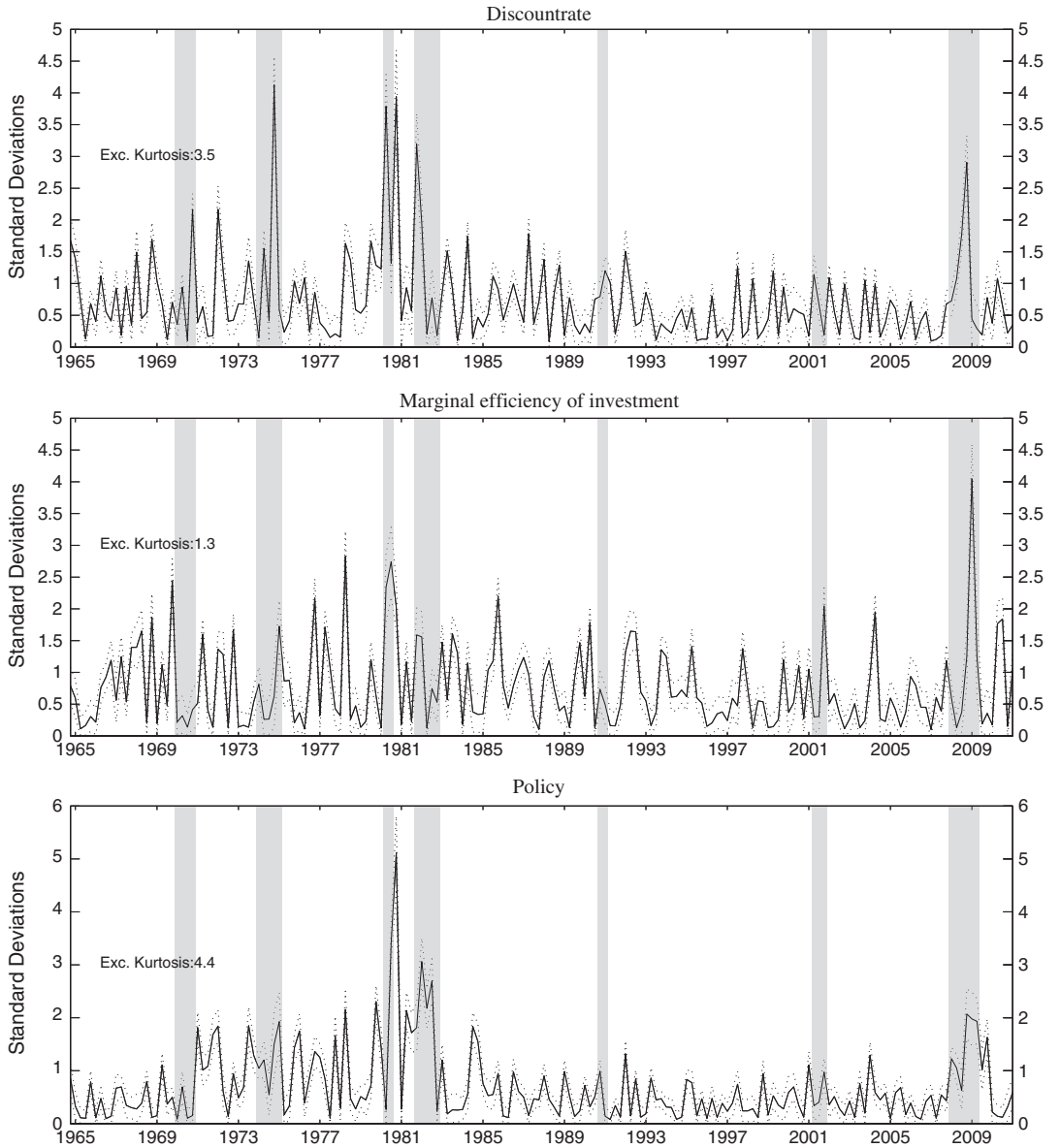

Figure 1. Smoothed shocks under Gaussianity (absolute value, standardized). The solid line is the median and the dashed lines are the posterior 90% bands. The vertical shaded regions identify NBER recession dates

rare shocks may in fact be Gaussian shocks whose effect is amplified through a nonlinear propagation mechanism. In fact, the extent to which nonlinearities can alleviate the need for fat-tailed shocks to explain business cycles could possibly become an additional metric for evaluating their usefulness. In this regard, the results in Dewachter and Wouters (2012) are very promising. These authors solve nonlinearly a model with capital-constrained financial intermediaries and provide very interesting evidence suggesting that the nonlinear propagation of shocks may induce outcomes that resemble those due to fat-tailed shocks in a linear model.

DOI: 10.1002/jae

Our findings are the following. We provide strong evidence that the Gaussianity assumption in DSGE models is counterfactual, even after allowing for low-frequency changes in the volatility of shocks. Such strong evidence is remarkable considering that our sample consists of macro variables only, which implies fat tails are not just a feature of financial data but of macro data as well. This finding is robust to excluding the Great Recession from the sample. We follow two approaches in our analysis: comparing the fit of different specifications using Bayesian marginal likelihoods, and inference on the posterior estimates of the degrees of freedom of Student's t-distribution. We demonstrate that the model fit improves considerably if we allow for Student's t shocks in addition to stochastic volatility. Further, we show that the posterior estimates of Student's t-distribution's degrees of freedom for some shocks are quite low, indicating a substantial degree of fat tails. From our results, we can cluster the shocks in the model into three broad categories. Shocks to productivity, to the household's discount rate, to the marginal efficiency of investment and to the wage markup all have fat tails, even in the case with stochastic volatility. Conversely, shocks to government expenditures and to price markups have posterior means for the degrees of freedom that are somewhat high, indicating that their distribution is not far from Gaussian, regardless of whether we allow for stochastic volatility. Finally, the degrees of freedom for monetary policy shocks are estimated to be extremely low in the case with constant volatility, but shift dramatically toward higher values when we allow for stochastic volatility.

In order to evaluate the importance of fat tails in the study of the business cycle, we consider an experiment in which we shut down the fat tails and recreate a counterfactual path of the economy in their absence. We show that in this case almost all recessions in the sample would have been of roughly the same magnitude, and the Great Recession would have been essentially a 'run-of-the-mill' recession. Finally, we show that allowing for fat tails changes the inference about slow-moving stochastic volatility. Specifically, we re-evaluate the evidence in favor of the Great Moderation hypothesis, discussed for example in JP, and find that when we consider Student's t shocks the magnitude of the reduction in the volatility of output and other variables is smaller. Similarly, we show that the evidence in favor of a permanent increase in volatility following the Great Recession is weaker when we consider the possibility that shocks have a Student's t-distribution.

Although we show that inference about low-frequency movements in volatility is affected by whether we allow for Student's t shocks, we should emphasize that they capture very different features of the data and are therefore far from being perfect substitutes. In other words, this paper is not a horse race between stochastic volatility and fat tails. In fact, we find that the data provide fairly strong evidence in favor of both features as long as the sample includes both pre and post Great Moderation data.

The rest of the paper proceeds as follows. Section 2 discusses Bayesian inference. Section 3 describes the model, as well as our set of observables. Section 4 describes the results. We conclude in Section 5.

2. BAYESIAN INFERENCE

This section begins with a description of the estimation procedure for a DSGE model with both Student's t-distributed shocks and stochastic volatilities (which we will often refer to as SV). The Gibbs sampler combines the algorithm proposed by Geweke (1993) for a linear model with Student's t-distributed shocks (see also Geweke, 1994, 2005, for a textbook exposition) with the approach for sampling the parameters of DSGE models with time-varying volatilities discussed in JP. In essence, our sampler applies to DSGE models of the algorithms proposed by Chib *et al.* (2002) and Jacquier *et al.* (2004) for combining stochastic volatility and fat tails. We mainly follow JP and Chib *et al.* (2002) in using the approach of Kim *et al.* (1998; henceforth KSC) in drawing the stochastic volatilities, but we also check the robustness of our results using the approach developed by Jacquier *et al.* (1994; henceforth, JPR).

DOI: 10.1002/jae

The model consists of the standard measurement and transition equations:

$$y_t = D(\theta) + Z(\theta)s_t \tag{1}$$

$$s_{t+1} = T(\theta)s_t + R(\theta)\varepsilon_t \tag{2}$$

for $t = 1, \ldots, T$, where y_t, s_t and ε_t are $n\times 1, k\times 1$, and $\bar{q}\times 1$ vectors of observables, states and shocks, respectively. Let $p(\theta)$ denote the prior on the vector of DSGE model parameters θ. We assume that

$$\varepsilon_{q,t} = \sigma_{q,t}\tilde{h}_{q,t}^{-1/2}\eta_{q,t}, \text{ all } q,t \tag{3}$$

where

$$\eta_{q,t} \sim \mathcal{N}(0,1), \text{ i.i.d. across } q,t \tag{4}$$

$$\lambda_q\tilde{h}_{q,t} \sim \chi^2(\lambda_q), \text{ i.i.d. across } q,t \tag{5}$$

For the prior on the parameters λ_q we assume Gamma distributions with parameters $\underline{\lambda}/\underline{\nu}$ and $\underline{\nu}$:

$$p\left(\lambda_q \left|\underline{\lambda}, \underline{\nu}\right.\right) = \frac{(\underline{\lambda}/\underline{\nu})^{-\underline{\nu}}}{\Gamma(\underline{\nu})}\lambda_q^{\underline{\nu}-1}\exp\left(-\underline{\nu}\frac{\lambda_q}{\underline{\lambda}}\right), \text{ i.i.d. across } q \tag{6}$$

where $\underline{\lambda}$ is the mean and $\underline{\nu}$ is the number of degrees of freedom (Geweke, 2005, assumes a Gamma prior with one degree of freedom). Note that the fat tails in this model affect our structural shocks $\varepsilon_{q,t}$ and hence cannot interpreted as 'outliers' as in standard regression models.[3] Define

$$\tilde{\sigma}_{q,t} = \log(\sigma_{q,t}/\sigma_q) \tag{7}$$

where the parameters $\{\sigma_q\}_{q=1}^{\bar{q}}$ (the non-time-varying component of the shock variances) are included in the vector of DSGE parameters θ. We assume that the $\tilde{\sigma}_{q,t}$ follows an autoregressive process:

$$\tilde{\sigma}_{q,t} = \rho_q\tilde{\sigma}_{q,t-1} + \zeta_{q,t}, \zeta_{q,t} \sim \mathcal{N}\left(0, \omega_q^2\right), \text{ i.i.d. across } q,t \tag{8}$$

The prior distribution for ω_q^2 is an inverse Gamma $\mathcal{IG}\left(\nu_\omega/2, \nu_\omega\underline{\omega}^2/2\right)$, i.e.

$$p\left(\omega_q^2 \left|\nu_\omega, \underline{\omega}^2\right.\right) = \frac{\left(\nu_\omega\underline{\omega}^2/2\right)^{\frac{\nu_\omega}{2}}}{\Gamma(\nu_\omega/2)}\left(\omega_q^2\right)^{-\frac{\nu_\omega}{2}-1}\exp\left[-\frac{\nu_\omega\underline{\omega}^2}{2\omega_q^2}\right], \text{ i.i.d. across } q \tag{9}$$

We consider two types of priors for ρ_q:

$$p\left(\rho_q \left|\omega_q^2\right.\right) = \begin{cases} 1 & \text{SV-UR} \\ \mathcal{N}\left(\bar{\rho}, \omega_q^2\bar{v}_\rho\right)\mathcal{I}(\rho_q), \text{ i.i.d. across } q, \mathcal{I}(\rho_q) = \begin{cases} 1 \text{ if } |\rho_q| < 1 \\ 0 \text{ otherwise} \end{cases} & \text{SV-S} \end{cases} \tag{10}$$

[3] We do not have measurement errors at all in this model, let alone Student's t errors. From looking at the plots of the shocks in Figure 1 (and the posterior estimates of $\tilde{h}_t$ in Figure A.6) it appears that the many large shocks coincide with macro events and are therefore unlikely to just represent measurement error.

DOI: 10.1002/jae

In the SV-UR case $\tilde{\sigma}_{q,t}$ follows a random walk as in JP, while in the SV-S it follows a stationary process as in Fernández-Villaverde and Rubio-Ramírez (2007). In both cases the $\sigma_{q,t}$ process is persistent: in the SV-UR case the persistence is wired into the assumed law of motion for $\tilde{\sigma}_{q,t}$, while in the SV-S case it is enforced by choosing the hyperparameters $\bar{\rho}$ and $\bar{\sigma}_\rho$ in such a way that the prior for ρ_q puts most mass on relatively high values of ρ_q. As a consequence, $\sigma_{q,t}$ and $\tilde{h}_{q,t}$ play different roles in equation (3): $\sigma_{q,t}$ allows for slow-moving trends in volatility, while $h_{q,t}$ allows for large transitory shocks. In essence, this model postulates that there are two coexisting processes driving the volatility dynamics in the macro data: one that is very persistent $\left(\tilde{\sigma}_{q,t}\right)$ and one that is high frequency—in our case, i.i.d. $\left(\tilde{h}_{q,t}\right)$. The literature so far has focused on the former and largely ignored the latter.[4] Of course, the i.i.d. assumption for the high-frequency movements in volatility is very stark. In future research it will be worth trying to relax this assumption. Considering two overlapping Markov-switching processes for the volatilities—one persistent and one transitory—is an option.

Finally, to close the model we make the following distributional assumptions on the initial conditions $\tilde{\sigma}_{q,0}$:

$$p\left(\tilde{\sigma}_{q,0}\left|\rho_q,\omega_q^2\right.\right)=\begin{cases}0 & \text{SV-UR}\\ \mathcal{N}\left(0,\omega_q^2/\left(1-\rho_q^2\right)\right), \text{ i.i.d. across } q & \text{SV-S}\end{cases} \tag{11}$$

where the restriction under the SV-UR case is needed to obtain identification. In the stationary case we have assumed that $\tilde{\sigma}_{q,0}$ is drawn from the ergodic distribution. In the remainder of the paper we will use the notation $x_{1:\bar{q},1:T}$ to denote the sequence $\left\{x_{1,1},\ldots,x_{1,T},\ldots,x_{\bar{q},T}\right\}$ and, to further simplify notation if no ambiguity arises, we will often omit the $1:\bar{q}$ subscript. So, for instance, $\tilde{h}_{1:T}$ will denote $\left\{\tilde{h}_{1,1},\ldots,\tilde{h}_{1,T},\ldots,\tilde{h}_{\bar{q},T}\right\}$.

2.1. The Gibbs Sampler

The joint distribution of data and unobservables (parameters and latent variables) is given by

$$\begin{aligned}&p(y_{1:T}|s_{1:T},\theta)p(s_{1:T}|\varepsilon_{1:T},\theta)p\left(\varepsilon_{1:T}\left|\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\theta\right.\right)p\left(\tilde{h}_{1:T}\left|\lambda_{1:\bar{q}}\right.\right)\\&p\left(\tilde{\sigma}_{1:T}\left|\rho_{1:\bar{q}},\omega_{1:\bar{q}}^2\right.\right)p\left(\lambda_{1:\bar{q}}\right)p\left(\rho_{1:\bar{q}}\left|\omega_{1:\bar{q}}^2\right.\right)p\left(\omega_{1:\bar{q}}^2\right)p(\theta)\end{aligned} \tag{12}$$

where $p\left(y_{1:T}\left|s_{1:T},\theta\right.\right)$ and $p(s_{1:T}|\varepsilon_{1:T},\theta)$ come from the measurement and transition equation, respectively, $p\left(\varepsilon_{1:T}\left|\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\theta\right.\right)$ obtains from equations (3) and (4):

$$p\left(\varepsilon_{1:T}\left|\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\theta\right.\right)\propto\prod_{q=1}^{\bar{q}}\left(\prod_{t=1}^{T}\tilde{h}_{q,t}^{-1/2}\sigma_{q,t}\right)\exp\left[-\sum_{t=1}^{T}\tilde{h}_{q,t}\varepsilon_{q,t}^2/2\sigma_{q,t}^2\right] \tag{13}$$

$p\left(\tilde{h}_{1:T}\left|\lambda_{1:\bar{q}}\right.\right)$ obtains from equation (5):

$$p\left(\tilde{h}_{1:T}\left|\lambda_{1:\bar{q}}\right.\right)=\prod_{q=1}^{\bar{q}}\prod_{t=1}^{T}\left(2^{\lambda_q/2}\Gamma(\lambda_q/2)\right)^{-1}\lambda_q^{\lambda_q/2}\tilde{h}_{q,t}^{(\lambda_q-2)/2}\exp\left(-\lambda_q\tilde{h}_{q,t}/2\right) \tag{14}$$

[4] Whether the low-frequency movements in volatility should be modeled using auto-regressive processes or regime-switching models is an ongoing debate in empirical macroeconomics (e.g. Liu *et al.*, 2011).

DOI: 10.1002/jae

and $p\left(\tilde{\sigma}_{1:T}\left|\omega^2_{1:\bar{q}}\right.\right)$ obtains from expressions (8) and (11):

$$p\left(\tilde{\sigma}_{1:T}\left|\rho_{1:\bar{q}},\omega^2_{1:\bar{q}}\right.\right)\propto\prod_{q=1}^{\bar{q}}\left(\omega_q^2\right)^{-(T-1)/2}\exp\left[-\sum_{t=2}^{T}\left(\tilde{\sigma}_{q,t}-\rho_q\tilde{\sigma}_{q,t-1}\right)^2/2\omega_q^2\right]p\left(\tilde{\sigma}_{q,1}\left|\rho_q,\omega_q^2\right.\right) \tag{15}$$

where

$$p\left(\tilde{\sigma}_{q,1}\left|\rho_q,\omega_q^2\right.\right)\propto\begin{cases}\left(\omega_q^2\right)^{-1/2}\exp\left(-\frac{\tilde{\sigma}_{q,1}^2}{2\omega_q^2}\right) & \text{SV-UR}\\ \left(\omega_q^2\left(1-\rho_q^2\right)\right)^{-1/2}\exp\left(-\frac{\tilde{\sigma}_{q,1}^2}{2\omega_q^2\left(1-\rho_q^2\right)}\right) & \text{SV-S}\end{cases} \tag{16}$$

Finally, $p\left(\lambda_{1:\bar{q}}\right)=\prod_{q=1}^{\bar{q}}p\left(\lambda_q\left|\underline{\lambda}\right.\right)$, $p\left(\omega^2_{1:\bar{q}}\right)=\prod_{q=1}^{\bar{q}}p\left(\omega_q^2\left|\nu,\underline{\omega}^2\right.\right)$.

The sampler is slightly different depending on the approach for drawing the stochastic volatilities, which are obtained from

$$p\left(\varepsilon_{1:T}\left|\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\theta\right.\right)p\left(\tilde{\sigma}_{1:T}\left|\rho_{1:\bar{q}},\omega^2_{1:\bar{q}}\right.\right) \tag{17}$$

Section A.3 in the online Appendix (supporting information) describes these differences, and provides the details of the implementation under both the KSC and the JPR approaches.[5] Call $\vartheta=\left\{\theta,s_{1:T},\tilde{h}_{1:T},\lambda_{1:\bar{q}},\rho_{1:\bar{q}},\omega^2_{1:\bar{q}},\varepsilon_{1:T}\right\}$. The next section shows how to draw ϑ given $\tilde{\sigma}_{1:T}$.

2.1.1. Drawing from $\vartheta\left|\tilde{\sigma}_{1:T},y_{1:T}\right.$

We draw from $\vartheta\left|\tilde{\sigma}_{1:T},y_{1:T}\right.$ using a sequence of conditional distributions (i.e. a Gibbs-within-Gibbs step). These are as follows:

(i) Draw from $p\left(\theta,s_{1:T},\varepsilon_{1:T}\left|\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\lambda_{1:\bar{q}},\rho_{1:\bar{q}},\omega^2_{1:\bar{q}},y_{1:T}\right.\right)$. This is accomplished in two steps:

(a) Draw from the marginal $p\left(\theta\left|\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\lambda_{1:\bar{q}},\rho_{1:\bar{q}},\omega^2_{1:\bar{q}},y_{1:T}\right.\right)$, where

$$p\left(\theta\left|\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\lambda_{1:\bar{q}},\rho_{1:\bar{q}},\omega^2_{1:\bar{q}},y_{1:T}\right.\right)\propto p\left(y_{1:T}\left|\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\theta\right.\right)p(\theta) \tag{18}$$

where

$$p\left(y_{1:T}\left|\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\theta\right.\right)=\int p(y_{1:T}|s_{1:T},\theta)p(s_{1:T}|\varepsilon_{1:T},\theta)p\times\left(\varepsilon_{1:T}\left|\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\theta\right.\right)\cdot\mathrm{d}(s_{1:T},\varepsilon_{1:T})$$

is computed using the Kalman filter with (1) as the measurement equation and (2) as transition equation, with

$$\varepsilon_t\left|\tilde{h}_{1:T},\tilde{\sigma}_{1:T}\right.\sim\mathcal{N}(0,\Delta_t) \tag{19}$$

[5] Note that under the KSC approach we structure the sampler as in Del Negro and Primiceri (2013). Moreover, we follow Omori *et al.* (2007) in using a 10-mixture approximation, as opposed to the seven-mixture approximation adopted in Kim *et al.* (1998).

DOI: 10.1002/jae

where Δ_t are $\bar{q} \times \bar{q}$ diagonal matrices with $\sigma_{q,t}^2 \cdot \tilde{h}_{q,t}^{-1}$ on the diagonal. The draw is obtained from a random-walk Metropolis–Hastings step.[6]

(b) Draw from the conditional $p\left(s_{1:T}, \varepsilon_{1:T} \left| \theta, \tilde{h}_{1:T}, \tilde{\sigma}_{1:T}, \lambda_{1:\bar{q}}, \rho_{1:\bar{q}}, \omega_{1:\bar{q}}^2, y_{1:T} \right.\right)$. This is accomplished using the simulation smoother of Durbin and Koopman (2002).

(ii) Draw from $p\left(\tilde{h}_{1:T} \left| \theta, s_{1:T}, \varepsilon_{1:T}, \tilde{\sigma}_{1:T}, \lambda_{1:\bar{q}}, \rho_{1:\bar{q}}, \omega_{1:\bar{q}}^2, y_{1:T} \right.\right)$. This is accomplished by drawing from

$$\begin{aligned} & p\left(\varepsilon_{1:T} \left| \tilde{h}_{1:T}, \tilde{\sigma}_{1:T}, \theta \right.\right) p\left(\tilde{h}_{1:T} \left| \lambda_{1:\bar{q}} \right.\right) \\ & \quad \propto \prod_{q=1}^{\bar{q}} \prod_{t=1}^{T} \tilde{h}_{q,t}^{(\lambda_q - 1)/2} \exp\left(-\left[\lambda_q + \varepsilon_{q,t}^2 / \sigma_{q,t}^2\right] \tilde{h}_{q,t} / 2\right) \end{aligned} \tag{20}$$

which implies

$$\left[\lambda_q + \varepsilon_{q,t}^2 / \sigma_{q,t}^2\right] \tilde{h}_{q,t} \left| \theta, \varepsilon_{1:T}, \tilde{\sigma}_{1:T}, \lambda_q \sim \chi^2(\lambda_q + 1) \right. \tag{21}$$

(iii) Draw from $p\left(\lambda_{1:\bar{q}} \left| \tilde{h}_{1:T}, \theta, s_{1:T}, \varepsilon_{1:T}, \rho_{1:\bar{q}}, \omega_{1:\bar{q}}^2, y_{1:T} \right.\right)$. This is accomplished by drawing from

$$\begin{aligned} p\left(\tilde{h}_{1:T} \left| \lambda_{1:\bar{q}} \right.\right) p\left(\lambda_{1:\bar{q}}\right) \propto & \prod_{q=1}^{\bar{q}} \left(\left(\underline{\lambda}/\underline{\nu}\right)^{\underline{\nu}} \Gamma\left(\underline{\nu}\right)\right)^{-1} \left[2^{\lambda_q/2} \Gamma(\lambda_q/2)\right]^{-T} \lambda_q^{T\lambda_q/2 + \underline{\nu} - 1} \\ & \left(\prod_{t=1}^{T} \tilde{h}_{q,t}^{(\lambda_q - 2)/2}\right) \exp\left[-\left(\frac{\underline{\nu}}{\underline{\lambda}} + \frac{1}{2}\sum_{t=1}^{T} \tilde{h}_{q,t}\right) \lambda_q\right] \end{aligned} \tag{22}$$

This is a non-standard distribution; hence the draw is obtained from a Metropolis–Hastings step. Following Geweke (2005), we use a log-normal proposal.

(iv) Draw from $p\left(\omega_{1:\bar{q}}^2, \rho_{1:\bar{q}} \left| \tilde{\sigma}_{1:T}, \theta, s_{1:T}, \varepsilon_{1:T}, \tilde{h}_{1:T}, \lambda_{1:\bar{q}}, y_{1:T} \right.\right)$ using

$$\begin{aligned} & p\left(\tilde{\sigma}_{1:T} \left| \rho_{1:\bar{q}}, \omega_{1:\bar{q}}^2 \right.\right) p\left(\omega_{1:\bar{q}}^2\right) p\left(\rho_{1:\bar{q}} \left| \omega_{1:\bar{q}}^2 \right.\right) \propto \prod_{q=1}^{\bar{q}} \left(\omega_q^2\right)^{-\frac{\nu+T-1}{2}-1} \\ & \exp\left[-\frac{\nu\underline{\omega}^2 + \sum_{t=2}^{T}\left(\tilde{\sigma}_{q,t} - \rho_q \tilde{\sigma}_{q,t-1}\right)^2}{2\omega_q^2}\right] p\left(\tilde{\sigma}_{q,1} \left| \rho_q, \omega_q^2 \right.\right) p\left(\rho_q \left| \omega_q^2 \right.\right) \end{aligned} \tag{23}$$

where $p\left(\tilde{\sigma}_{q,1} \left| \rho_q, \omega_q^2 \right.\right)$ is given by equation (16). In the SV-UR case ρ_q is fixed to 1, and we can draw ω_q^2 (i.i.d. across q) from

$$\omega_q^2 \left| \tilde{\sigma}_{1:T}, \cdots \sim \mathcal{IG}\left(\frac{\nu + T}{2}, \frac{1}{2}\left(\nu\underline{\omega}^2 + \sum_{t=2}^{T}\left(\tilde{\sigma}_{q,t} - \tilde{\sigma}_{q,t-1}\right)^2 + \tilde{\sigma}_{q,1}^2\right)\right) \right. \tag{24}$$

[6] In keeping with the spirit of the paper, we use a Student's t-distribution with 10 degrees of freedom, as opposed to the customary Gaussian, as our proposal density.

DOI: 10.1002/jae

In the SV-S case the joint posterior of ρ_q, ω_q^2 is non-standard because of the likelihood of the first observation $p\left(\tilde{\sigma}_1 \left| \rho_q, \omega_q^2\right.\right)$. We therefore use the Metropolis–Hastings step proposed by Chib and Greenberg (1994). Specifically, we use as proposal density the usual normal-inverse Gamma distribution, i.e.

$$\begin{aligned} \omega_q^2 \left|\tilde{\sigma}_{1:T}, \ldots \right. &\sim \mathcal{IG}\left(\frac{\nu + T - 1}{2}, \frac{1}{2}\left(\nu\underline{\omega}^2 + \sum_{t=2}^{T} \tilde{\sigma}_{q,t}^2 + \bar{v}_\rho^{-1}\bar{\rho}^2 - \hat{V}_q^{-1}\hat{\rho}_q^2\right)\right), \\ \rho_q \left|\omega_q^2, \tilde{\sigma}_{1:T}, \ldots \right. &\sim \mathcal{N}\left(\hat{\rho}_q, \omega_q^2\hat{V}_q\right), \text{ i.i.d. across } q \end{aligned} \tag{25}$$

where $\hat{\rho}_q = \hat{V}_q\left(\bar{v}_\rho^{-1}\bar{\rho} + \sum_{t=2}^{T}\tilde{\sigma}_{q,t}\tilde{\sigma}_{q,t-1}\right), \hat{V}_q = \left(\bar{v}_\rho^{-1} + \sum_{t=2}^{T}\tilde{\sigma}_{q,t-1}^2\right)^{-1}$. We then accept/reject this draw using the proposal density and the acceptance ratio

$$\frac{p\left(\tilde{\sigma}_1, \rho_q^{(*)}, \omega_q^{2\,(*)}\right)\mathcal{I}\left(\rho_q^{(*)}\right)}{p\left(\tilde{\sigma}_1, \rho_q^{(j-1)}, \omega_q^{2\,(j-1)}\right)\mathcal{I}\left(\rho_q^{(j-1)}\right)}$$

with $\left(\rho^{(j-1)}, \omega_q^{2\,(j-1)}\right)$ and $\left(\rho^{(*)}, \omega_q^{2\,(*)}\right)$ being the draw at the $(j-1)$th iteration and the proposed draw, respectively.

3. THE DSGE MODEL

The model considered is the one used in Smets and Wouters (2007), which is based on earlier work by Christiano *et al.* (2005) and Smets and Wouters (2003). This is a medium-scale DSGE model, which augments the standard neoclassical stochastic growth model with nominal price and wage rigidities as well as habit formation in consumption and investment adjustment costs. Owing to space constraints, and to the fact that the model is the same as in SW, we relegate the detailed model description to online Appendix A (supporting information); it is also available in Cúrdia *et al.* (2012), the working paper version of this paper. Since this is a paper about shocks, however, we describe what they are and how they enter the model. The model has seven exogenous processes, and the shocks (ε_t) are innovations to these processes. Government spending g enters the resource constraint—these are units of output the government demands in any period but have no productive use; the discount rate process b enters the Euler equation, and drives a wedge between the marginal utility of consumption and the riskless real return; the 'marginal efficiency of investment' (MEI) process μ affects the rate of transformation between consumption and installed capital; total factor productivity (TFP) z enters the production function; r^m is the residual in the policy interest rate rule (and innovations to this process are called 'policy shocks'); price (λ_f) and wage (λ_g) markup shocks enter the price and wage Phillips curves, respectively. All these shocks follow univariate AR(1) processes, with the exception of price and wage markup shocks, which follow ARMA(1,1) processes. Innovations to the TFP process also affect the government spending process contemporaneously.

We collect all the DSGE model parameters in the vector θ, stack the structural shocks in the vector ε_t, and derive a state-space representation which is composed of the transition equation (2), which summarizes the evolution of the states s_t, and the measurement equation (1), which maps the states onto the vector of observables y_t, and where $D(\theta)$ represents the vector of steady-state values for these observables. Specifically, the model is estimated based on seven quarterly macroeconomic time series: real output, consumption, investment, and wage growth, hours, inflation and interest rates, all measured in percent (the data construction is the same as in SW). Section A.2 in online Appendix A

DOI: 10.1002/jae

(supporting information) provides further details on the data. In our benchmark specification we use data from 1964:Q4 to 2011:Q1, but we also consider a variety of subsamples. Table A.1 in online Appendix A shows the priors for the DSGE model parameters, which coincide with those used in SW.[7]

4. RESULTS

This section describes our findings. First, we present the evidence in favor of Student's t-distributed shocks. Second, we quantify the impact of rare shocks on the macroeconomy. Third, we show the extent to which allowing for Student's t shocks affects the inference about time variation in volatility. Finally, we discuss the variance decomposition for real GDP growth. The results in Sections 4.1–4.4 are obtained under a specification using (i) the full sample, (ii) the SV-UR assumption for the law of motion of the stochastic volatilities and (iii) the KSC procedure for drawing the volatilities (see Section A.3.1). Section 4.5 studies the robustness of the results using different assumptions and estimation approaches for the SV component, and various subsamples. There are many more objects of interest than we have space to show and discuss here, such as the posterior estimates of the DSGE parameters, and the time series for the time-varying volatilities $\sigma_{q,t}$, the Student's t components $\tilde{h}_{q,t}$ and the standardized shocks $\eta_{q,t}$. We show these results in Appendix A, which is available online as supporting information, as well as in the working paper version (Cúrdia *et al.*, 2012), and we briefly discuss them in Sections A.6 and A.7.

4.1. Evidence of Fat Tails

In the Introduction we showed that shocks extracted from standard Gaussian estimation are sometimes quite large—four standard deviations or more in size. In this section we consider more formal evidence against Gaussianity.[8] Specifically, this section addresses two questions. First, do we still need fat-tailed shocks once we allow for low-frequency movements in volatility? Second, which shocks are fat-tailed, and how fat are the tails? We address these questions by: (i) assessing the improvement in fit obtained by allowing for Student's t-distributed shocks relative to both the standard model as well as models with low-frequency movements in the volatility; (ii) presenting the posterior distribution of the degrees of freedom for each shock.

Before we delve into the results, we provide an intuitive description of the relationship between the degrees of freedom of the Student's t-distribution and the likelihood of observing large shocks. Recall from equation (3) that a Student's t-distributed shock ε_t can be decomposed as $\varepsilon_t = \sigma_t \tilde{h}_t^{-1/2} \eta_t$, where η_t is drawn from a standard Gaussian distribution (we omit the q subscript for ease of exposition). Therefore, given σ_t, the chances of observing a very large ε_t depend on the chances of $\tilde{h}_t$ being small. The prior for $\tilde{h}_t$ is given by equation (5). If λ is high, this prior concentrates around one, and the likelihood of observing large shocks is slim (the $\lambda \to \infty$ limit represents Gaussianity). As λ drops, the distribution of $\tilde{h}_t$ spreads out and the chances of observing a low $\tilde{h}_t$ increase. The following table provides a quantitative feel for what different λs imply in terms of the model's ability

[7] We follow SW (but not necessarily best practice in DSGE model estimation, e.g. Del Negro and Schorfheide, 2008) in assuming independent priors for the DSGE parameters.

[8] A referee asked to see how the specifications with Student's t-distributed shocks and stochastic volatility deal with these large innovations—i.e. whether the large $|\varepsilon_t|$ is rationalized with a large σ_t or a large Student's t component $\tilde{h}_t^{-1/2}$. Figures A.6 and A.7 in the online Appendix show the time series of σ_t and $\tilde{h}_t^{-1/2}$ and therefore address this question. Note that since the posterior estimates of the DSGE model parameters are quite similar across specifications, as discussed in Section A.6, the innovations ε_t are also very similar, and provide a good basis for comparison. The figures show that the large discount rate and MEI shocks are explained overwhelmingly as rare shocks (high values of $\tilde{h}_t^{-1/2}$ while, not surprisingly, the large policy shocks in the late 1970s/early 1980s are due mainly to high volatility during that period (high values of $\tilde{\sigma}_t$).

DOI: 10.1002/jae

to generate fat-tailed shocks. Specifically, the table shows the number of shocks larger (in absolute value) than x standard deviations per 200 periods, which is the size of our sample. The table shows that with 9 degrees of freedom the chances of seeing even a single shock as large as those shown in Figure 1 are not high (.28 for shocks larger than 4 standard deviations), and become negligible for 15 or more degrees of freedom.

λ, x:	3	4	5
∞	0.54	0.012	$1e^{-4}$
15	1.14	0.13	0.02
9	1.57	0.28	0.06
6	2.08	0.54	0.17

In what follows, we consider three different Gamma priors of the form (6) for the degrees of freedom parameter λ, which capture different a priori views on the importance of fat tails. The first prior, $\underline{\lambda} = 15$, captures the view that the world is not quite Gaussian, but not too far from Gaussianity either. The second prior, $\underline{\lambda} = 9$, embodies the idea that the world is quite far from Gaussian, yet not too extreme. The last prior, $\underline{\lambda} = 6$, implies prior belief in a model with fairly heavy tails.

The tightness of these priors depends on the degrees of freedom parameter $\underline{\nu}$ in equation (6), which we set equal to 4.[9] Figure A.1 in the online Appendix shows the three priors for $\underline{\lambda} = 6$, 9 and 15. Since the variance of the prior is $\underline{\lambda}^2/\underline{\nu}$, lower values of $\underline{\lambda}$ correspond to a tighter prior. Note, however, that the prior with higher variance ($\underline{\lambda} = 15$) puts most of the mass on high values of the degrees of freedom. For instance, the prior probability put on the regions $\{\lambda < 4\}, \{\lambda < 6\}$ and $\{\lambda < 9\}$ by the $\underline{\lambda} = 15$ prior is roughly 2.5%, 8% and 22%, respectively.

With the description of the prior in hand, we are now ready to discuss our evidence on the importance of fat tails. Table I shows the log-marginal likelihood—the standard measure of fit in a Bayesian framework—for models with different assumptions on the shocks distribution.[10] We consider four different combinations: (i) Gaussian shocks with constant volatility (baseline); (ii) Gaussian shocks with time-varying volatility; (iii) Student's t-distributed shocks with constant volatility; and (iv) both Student's t shocks and time variation in volatility (in the remainder of the paper we will sometimes refer to specifications (ii), (iii) and (iv) as TD, SV and SVTD, respectively).[11] Finally, our prior for the innovations in stochastic volatility is an $\mathcal{IG}$ prior with mode at $(0.01)^2$ and 0.1 degrees of freedom.[12]

The Gaussian/constant-volatility model is clearly rejected by the data. This is not surprising in light of the evidence contained in JP, Fernández-Villaverde and Rubio-Ramírez (2007) and Liu *et al.* (2011). Both the specification with SV only and those with Student's t shocks only perform substantially better than the Gaussian/constant-volatility specification. Also, the log-marginal likelihoods indicate that even after accounting for fat-tailed shocks, allowing for time variation in the volatility improves fit: for any row in Table I, the log-marginal likelihood increases moving from the left (no SV) to the right (SV) column. From the perspective of this paper, however, the main finding is provided by the fact that the fit improves when Student's t shocks are included, and continues to improve as the prior puts more weight on fat tails, regardless of whether we include SV. In summary, the data strongly

[9] Results obtained using $\underline{\nu} = 1$ are very similar.

[10] Online Appendix A.5 provides details on the computation of the marginal likelihood for these models.

[11] Here we focus on the specification where the volatilities follow a random walk (SV-UR case in equation (10)), as in JP, because this specification obtains a better fit of the data than the specification where the volatility follows a stationary autoregressive process (SV-S case), but we discuss the results for the SV-S case in Section 4.5.1.

[12] Specifically, in the $\mathcal{IG}\left(\nu_\omega/2, \nu_\omega\underline{\omega}^2/2\right)$ prior (see equation (9)) we set $\nu_\omega/2 = .1$ and choose $\underline{\omega}^2$ so that the mode, given by $\frac{\nu_\omega\underline{\omega}^2/2}{1+\nu_\omega/2}$, equals $(0.01)^2$. It is well known that when $\underline{\omega}^2$ is very low, as is the case here, designing a prior that is not too informative is challenging. In Monte Carlo experiments we found that using very low degrees of freedom $(\nu_\omega/2 = 0.1)$ led to good performance regardless of whether the true value of ω^2 was 10^{-4}, 10^{-3}, or 10^{-2}, and therefore settled on this value. Figure A.2 in online Appendix A provides a plot of the prior distribution for ω^2.

DOI: 10.1002/jae

Table I. Log-marginal likelihoods

	Constant volatility	Stochastic volatility
Gaussian shocks		
	−1117.9	−1083.7
Student's t-distributed shocks		
$\underline{\lambda} = 15$	−999.8	−994.0
$\underline{\lambda} = 9$	−990.6	−972.2
$\underline{\lambda} = 6$	−975.9	−964.2

Note: The parameter $\underline{\lambda}$ represents the prior mean for the degrees of freedom in the Student's t-distribution.

Table II. Posterior of the Student's t degrees of freedom

	Without stochastic volatility			With stochastic volatility		
	$\underline{\lambda} = 15$	$\underline{\lambda} = 9$	$\underline{\lambda} = 6$	$\underline{\lambda} = 15$	$\underline{\lambda} = 9$	$\underline{\lambda} = 6$
Government (g)	9.9 (3.2, 16.8)	7.3 (3.1, 11.5)	5.9 (2.9, 8.9)	11.4 (5.5, 24.0)	8.3 (3.6, 13.1)	7.6 (3.6, 11.5)
Discount (b)	4.4 (2.3, 6.4)	4.1 (2.3, 5.9)	3.9 (2.3, 5.5)	5.5 (2.5, 8.4)	4.6 (2.4, 6.7)	5.3 (2.6, 7.9)
MEI (μ)	9.9 (3.3, 16.5)	7.5 (3.2, 11.8)	6.0 (2.9, 9.0)	9.2 (3.2, 15.3)	7.2 (3.1, 11.2)	6.1 (3.0, 9.3)
TFP (z)	6.0 (2.1, 10.2)	4.8 (2.1, 7.5)	4.2 (2.1, 6.2)	7.5 (2.5, 12.9)	5.5 (2.3, 8.7)	5.6 (2.5, 8.6)
Price markup (λ_f)	11.1 (3.7, 18.7)	8.0 (3.4, 12.7)	6.5 (3.1, 9.9)	12.3 (4.2, 20.5)	9.4 (3.9, 14.8)	8.2 (3.9, 12.3)
Wage markup (λ_w)	9.5 (3.5, 15.4)	7.3 (3.3, 11.3)	6.1 (3.1, 9.0)	10.8 (3.9, 17.9)	8.1 (3.5, 12.7)	6.9 (3.4, 10.3)
Policy (r^m)	3.4 (1.9, 5.0)	3.2 (1.9, 4.6)	3.1 (1.8, 4.3)	8.2 (2.3, 14.6)	5.9 (2.2, 9.8)	8.1 (3.9, 12.2)

Note: Numbers shown for the posterior mean and 90% intervals of the degrees of freedom parameter.

favor Student's t-distributed shocks with a non-negligible degree of fat tails, whether or not we allow for low-frequency movements in volatility.[13]

Our second piece of evidence comes from the posterior distribution of the degrees of freedom λ. Table II shows the posterior mean and 90% bands for the degrees of freedom for each shock, in the specifications with and without stochastic volatility. Two results emerge. First, for quite a few shocks the estimated degrees of freedom are small, even when we allow for low-frequency movements in volatility. Second, allowing for low-frequency movements in volatility substantially changes the inference about the degrees of freedom for some of the shocks, implying that this feature is necessary for a proper assessment of how fat-tailed macroeconomic shocks are.

As for the first result, Table II shows that four shocks have a mean below 6 for the best-fitting prior on the degrees of freedom according to the log-marginal likelihood criterion ($\underline{\lambda} = 6$). For priors with higher $\underline{\lambda}$ the posterior mean degrees of freedom for these four shocks increases, but this is mostly

[13] The difference in log-marginal likelihoods between the Gaussian/constant-volatility model and the SV only model is much larger than the difference in log-marginal likelihoods between any of the Student's t specification and the corresponding specification (for the same prior) with SV. This finding is robust across SV specifications and estimation procedures, as discussed in Section 4.5. This suggests that there is some substitutability between the two features (SV and Student's t) in terms of fitting the data. However, the two features cannot be complete substitutes: one captures low-frequency movements in volatility, the other very high-frequency ones. Hence we should not be surprised that the data favor specifications with SV even for relatively low values of $\underline{\lambda}$. Nonetheless, as the $\underline{\lambda} = 15$ specification is closer to the Gaussian one than, say, the $\underline{\lambda} = 6$ specification, we would have expected the difference between the case without SV and the case with SV to be larger for $\underline{\lambda} = 15$ than for $\underline{\lambda} = 6$, and we find this is not the case.

DOI: 10.1002/jae

because the posterior distribution becomes more skewed to the right (i.e. it places some mass on higher values for λ). Still, for many shocks the posterior distribution puts sizable mass on the $\{\lambda < 6\}$ region even for $\underline{\lambda} = 15$, and in quite a few cases the 5th percentile of the posterior distribution barely changes as a function of $\underline{\lambda}$. In any case, the substantive results on the importance of fat tails for macroeconomic fluctuations and on the inference about time variation in volatility are, by and large, the same regardless of the choice of $\underline{\lambda}$, as we discuss below. The shocks with the fattest tails (lowest posterior degrees of freedom) are those affecting the discount rate (b), TFP (z), the marginal efficiency of investment (μ) and the wage markup (λ_w) processes. Not surprisingly, these shocks are the usual suspects as key drivers of business cycles (see SW; Justiniano *et al.*, 2009).

As for the result that allowing for low-frequency movements in volatility substantially changes the inference about the degrees of freedom, the monetary policy shock r^m is a case in point. Its estimated degrees of freedom are very low when one ignores time variations in volatility apparent in Figure 1. In the specification ignoring time variation in volatility, the model interprets the large shocks of the late 1970s/early 1980s as evidence of fat tails. Once these secular changes in volatility are taken into account, the posterior estimates of the degrees of freedom increase substantially. What is the intuition for this finding? The posterior distribution of $\tilde{h}_t$, which determines how fat the tails are, is given by equation (21):

$$\left[\lambda + \varepsilon_t^2/\sigma_t^2\right]\tilde{h}_t \,|\theta, \varepsilon_{1:T}, \tilde{\sigma}_{1:T}, \lambda \sim \chi^2(\lambda + 1)$$

For a given value of σ_t, a larger estimated shock ε_t implies a smaller posterior value of $\tilde{h}_t$. Not surprisingly, large shocks are interpreted by the model as evidence for fat tails. However, the shock is standardized by σ_t: if a large shock occurs during a period where all shocks tend to be large, it is discounted, and the posterior value of $\tilde{h}_t$ may not be particularly small.[14]

4.2. Large Shocks and Macroeconomic Fluctuations

We have shown that quite a few important shocks in the SW model have fat tails. What does this mean in terms of business cycle fluctuations? This section tries to provide a quantitative answer to this question by performing a counterfactual experiment. Recall again from equation (3) that $\varepsilon_t = \sigma_t \tilde{h}_t^{-1/2}\eta_t$. Therefore, once we compute the posterior distribution of ε_t (the smoothed shocks) and $\tilde{h}_t$, we can purge the Student's t component from ε_t using $\tilde{\varepsilon}_t = \sigma_t \eta_t$. We can then compute counterfactual histories that would have occurred had the shocks been $\tilde{\varepsilon}_t$ instead of ε_t. All these counterfactuals are computed for the best-fitting model: SVTD($\underline{\lambda} = 6$).

The left-hand panel of Figure 2 shows these counterfactual histories for output, consumption and investment growth. For all plots the magenta (grey in print) solid lines are the median counterfactual paths, the magenta (grey in print) dashed lines represent the 90% bands, and the solid black lines represent the actual data. The right-hand panel uses actual and counterfactual histories to compute a rolling window standard deviation, where each window contains the prior 20 quarters as well as the following 20 quarters, for a total of 41 quarters. These rolling window standard deviations are commonly used measures of time variation in the volatility of the series. The difference between actual and counterfactual standard deviations measures the extent to which the change in volatility is accounted for by fat-tailed shocks.[15]

[14] Consistent with the intuition given above, the posterior mean of λ increases for all shocks (for given prior) when we include time variation in volatility, with the exception of the MEI(μ) shock, for which it remains virtually the same.

[15] The distribution of $\tilde{h}_t$ is non-time varying. However, since large shocks occur rarely, they may account for changes in the rolling window volatility. Note that in the counterfactual we 'kill' the fat tails without keeping the overall volatility of the process constant—the purpose is just to showcase the importance of rare shocks. Thus one should not interpret the counterfactual rolling window standard deviations as measures of 'true' underlying volatility.

DOI: 10.1002/jae

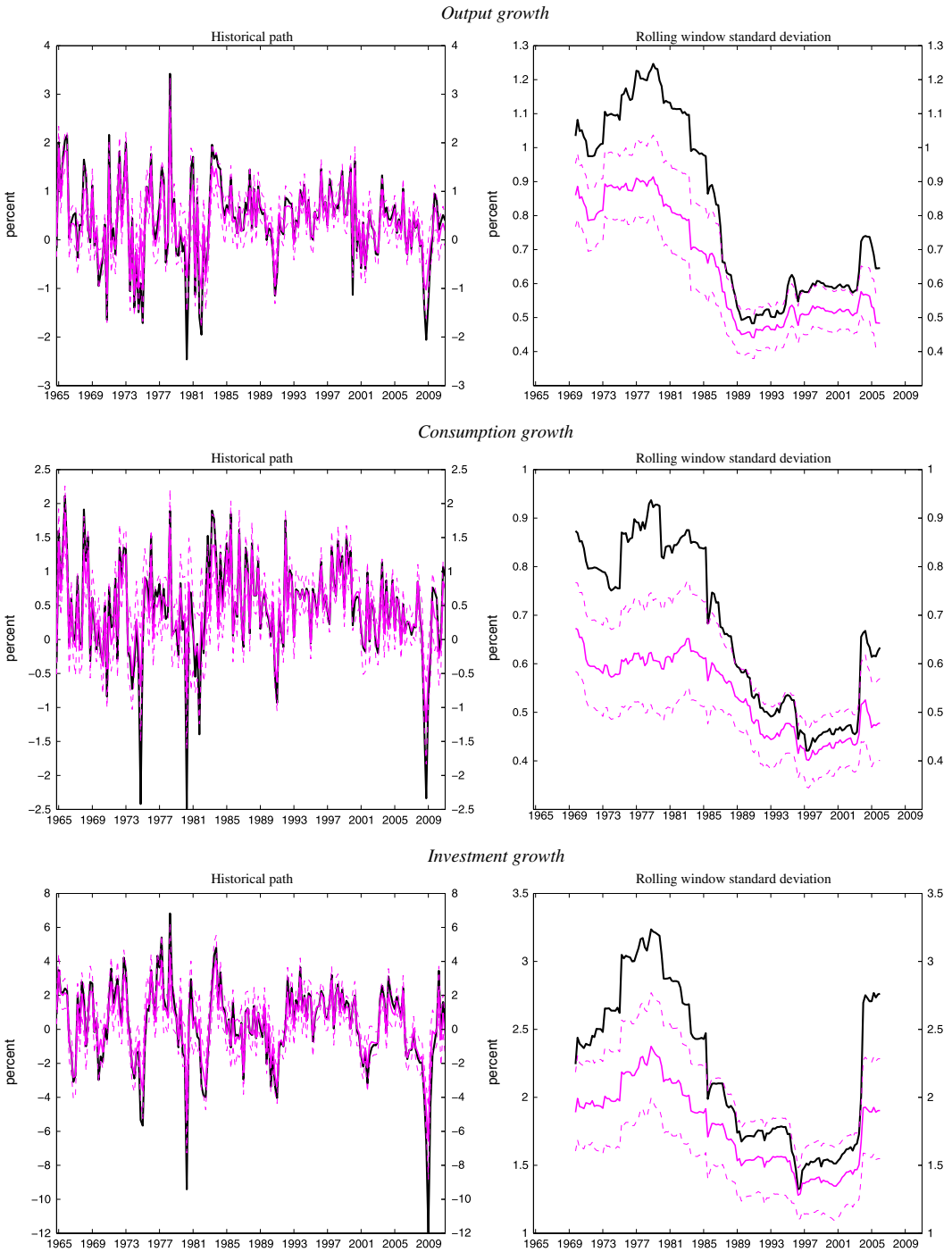

Figure 2. Counterfactual evolution of output, consumption and hours worked when the Student's t-distributed component is turned off, (estimation with Student's t-distributed shocks and stochastic volatility). Black lines are the historical evolution of the variable and magenta (grey in print) lines are the median counterfactual evolution of the same variable if we shut down the Student's t-distributed component of all shocks. The rolling window standard deviation uses 20 quarters before and 20 quarters after a given quarter

DOI: 10.1002/jae

The left-hand panels suggest that fat-tailed shocks account for a non-negligible part of fluctuations in the three variables. For output growth, the Student's t component accounted for a sizable fraction of the contraction in output growth during the Great Recession. In particular, if the fat tail component were absent the Great Recession would have been of about the same size as milder recessions, such as the 1990–91 recession. In general, without the fat-tailed component of the shocks all recessions (with the exception of the 2001 recession) would be of roughly the same magnitude in terms of output growth.[16] Figures A.3 and A.4 in online Appendix A of the paper show that these results are robust to the choice of $\underline{\lambda}$, the prior mean for the degrees of freedom of the Student's t-distribution.

Further, the rolling window standard deviation shown in the right-hand panel shows that the Student's t component explains a non-negligible part of changes in the realized volatility in the data. One can interpret this evidence as saying that the 1970s and early 1980s were more volatile than the Great Moderation period, at least in part because rare shocks took place.[17] Similar conclusions apply to the decomposition of consumption and investment growth (middle and bottom panels, respectively). It is notable that in the case of investment, and to a lesser extent of consumption, the rolling-window volatility computed from the data has spiked up recently to levels near those prior to the Great Moderation. When we take the Student's t component into account, however, the recent increase in the rolling-window volatility appears much milder.

4.3. Student's t Shocks and Inference about Time Variation in Volatility

This section discusses the extent to which accounting for fat tails makes us re-evaluate the magnitude of low-frequency changes in volatility. The point that inference about stochastic volatility changes when one accounts for rare large shocks is not new, and is made quite eloquently in Jacquier *et al.* (2004). Inference about the stochastic volatility is conducted using state-space methods under the KSC procedure, where

$$\log\left(\sigma^{-2}\tilde{h}_t\varepsilon_t^2 + c\right) = 2\tilde{\sigma}_t + \eta_t^*$$

is the measurement equation (with c a small constant), and equation (8) is the transition equation. Intuitively, the estimated time-varying log volatilities will try to fit the time series $\log\left(\sigma^{-2}\tilde{h}_t\varepsilon_t^2 + c\right)$. Since this quantity depends on ε_t^2, the model will interpret changes over time in the size of the squared shocks ε_t^2 as evidence of time variation in the volatilities $\tilde{\sigma}_t$. In a world with fat tails, ε_t^2 will vary over time simply because $\tilde{h}_t$ changes. If one ignores variations in $\tilde{h}_t$ by assuming Gaussian shocks, one may obtain the wrong inference about the time variation in the σ_t series. For instance, one may conclude that the Great Recession signals a permanent change in the level of macroeconomic volatility, when in fact it may be (at least in part) the result of a particularly large realization of the shocks.

Does all of this matter in practice? An implication of stochastic volatility is that the model-implied variance of the endogenous variables changes over time. Therefore, rather than looking at the posterior estimates of the stochastic volatility component for individual shocks, we focus here on the time variation of the standard deviation of output and consumption.[18] Specifically, Figure 3 shows the

[16] Dewachter and Wouters (2012) solve nonlinearly a model with capital-constrained financial intermediaries. They back out from the data a process for total factor productivity and feed this process into their nonlinear model. They then compare the path of several variables under the nonlinear and linear solution methods. These two paths for investment growth look strikingly similar to the actual (for the nonlinear method) and counterfactual (for the linear method) paths of investment growth in Figure 2, suggesting that the nonlinear propagation of shocks may induce outcomes that resemble those due to fat-tailed shocks in a linear model.

[17] For example, at the peak of the volatility in 1978, the rolling window standard deviation of output growth is about 1.25 in the data, but once we shut down the Student's t component it drops to 0.90, which is a reduction of about 28% in volatility.

[18] Figure A.7 in online Appendix A shows the posterior estimates of the shocks $\varepsilon_{q,t}$ (in absolute value) and the shock volatilities $\sigma_{q,t}$ in the SV-only and SVTD specifications.

DOI: 10.1002/jae

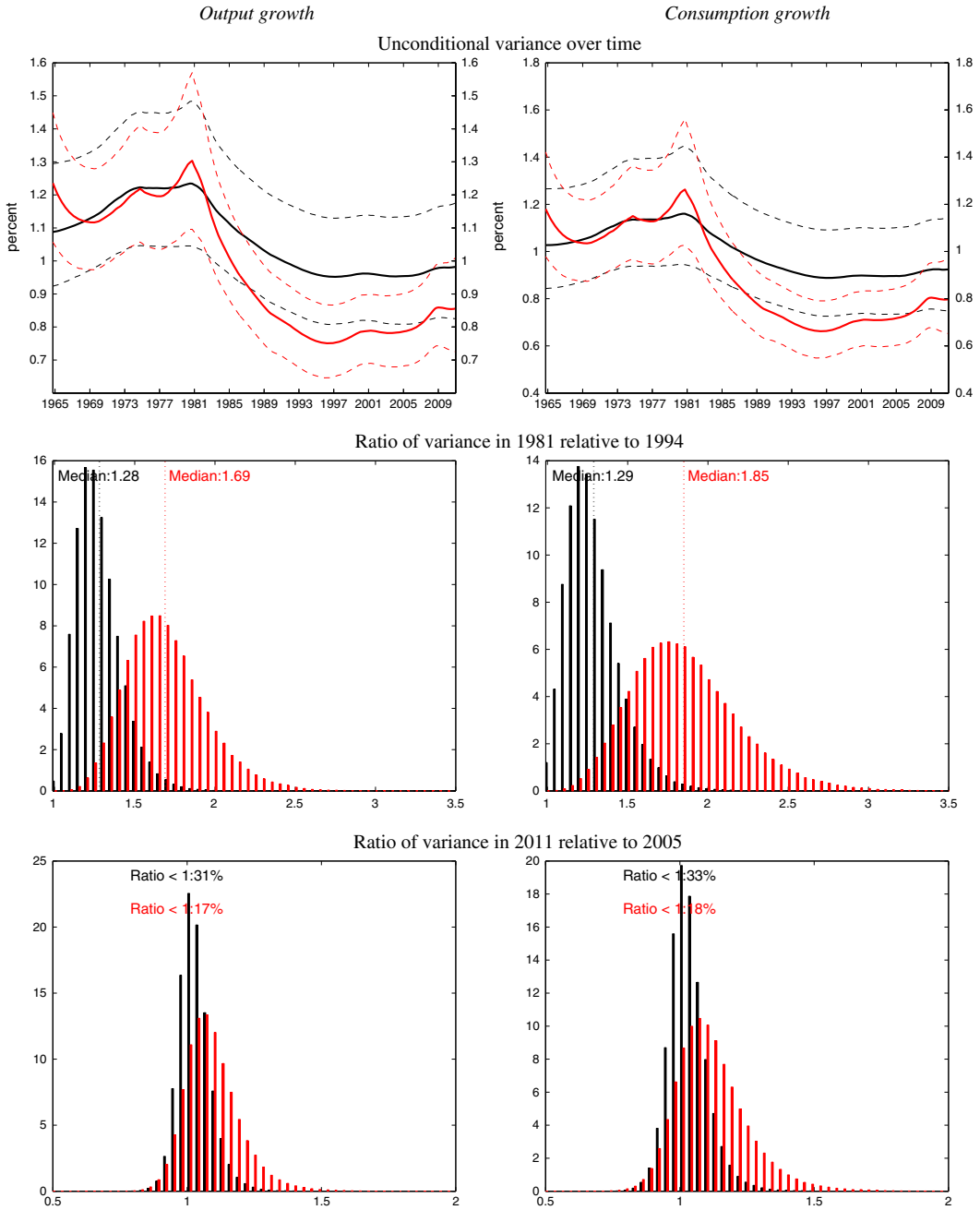

Figure 3. Time variation in the unconditional standard deviation of output and consumption; models estimated with and without the Student's t-distributed component. The black line in the top panel is the unconditional standard deviation in the estimation with both stochastic volatility and Student's t components, while the red (grey in print) line is the unconditional variance in the estimation with the stochastic volatility component only. In the middle panel the black bars correspond to the posterior histogram of the ratio of the variance in 1981 over the variance in 1994 for the estimation with both stochastic volatility and Student's t components, while the red (grey in print) bars are for the estimation with stochastic volatility component only. The lower panel replicates the same analysis as in the middle panel but for the ratio of the variance in 2011 over the variance in 2005

DOI: 10.1002/jae

model-implied volatility of output and consumption growth, as measured by the unconditional standard deviation of the series computed at each point in time t assuming that the standard deviations of the shocks is going to remain equal to the estimated value of $\sigma_{q,t} = \sigma_q e^{\tilde{\sigma}_{q,t}}$ forever after (i.e. abstracting from the fact that $\tilde{\sigma}_{q,t+s}$, for $s > 0$, will be affected by future volatility shocks according to equation (8); this is the same object computed in Figure 5 of JP).

In the top panel, the red line shows this measure for the SV-only specification, while the black lines show this volatility for the SVTD($\underline{\lambda} = 6$) specification. As in the other plots, the solid line is the posterior median and the dashed lines correspond to the 90% bands around the median. For both variables the model-implied volatility is generally higher when we account for fat-tailed shocks, which is intuitive. However, the difference between the models is not constant over time. At the peak of the high-volatility period (late 1970s and early 1980s), the two models agree. However, during the Great Moderation the model that does not allow for fat-tailed shocks seems to overestimate the decline in macroeconomic volatility.

The middle panel of Figure 3 hones in on this finding. This panel shows the posterior distribution of the ratio of the volatility in 1981 (roughly, the peak of the volatility series) relative to the volatility in 1994 (roughly, the bottom) for output and consumption growth, respectively. The red and black bars are for the SV-only and SVTD($\underline{\lambda} = 6$) specifications, respectively. Numbers greater than one indicate that volatility was higher in 1981 relative to 1994. There is no doubt that this is the case, as both histograms are well to the right of one. However, the magnitude of the decrease in volatility depends on whether or not we allow for fat-tailed shocks. The posterior distribution for the ratio of the output growth volatility in 1981 relative to the volatility in 1994 in the estimation with Student's t shocks is smaller relative to the case with Gaussian shocks. The median is 1.3 in the former, compared to 1.7 in the latter. For the red bars, most of the mass is to the right of 1.5, implying that volatility dropped by more than one-third between 1981 and 1994. The converse holds for the black bars, which show a smaller decline in volatility. This same pattern is also evident for the consumption growth. Figures A.3 and A.4 in online Appendix A show that these results are virtually identical for different choices of $\underline{\lambda}$, the prior mean for the degrees of freedom of the Student's t distribution.

As a result of the Great Recession there has been an increase in volatility in many macroeconomic variables since 2008, as measured by the rolling window standard deviations shown in Figure 2. To what extent does this increase reflect a permanent increase in the volatility of the underlying shocks, and, potentially, the end of the Great Moderation?[19] The bottom panels of Figure 3 show the ratio of the volatility in 2011 (end of the sample) relative to the volatility in 2005 (pre-Great Recession) for output and consumption growth, respectively, with numbers greater than one indicating a permanent rise in volatility. Under the model with time-varying volatility and Gaussian shocks, the probability that volatility in both output and consumption has increased after the Great Recession is quite high. The probability of the ratio being below one is 17% and 18% for output and consumption growth, respectively. The model that has Student's t shocks in addition to time-varying volatility is less confident: the probability of the ratio being below one increases to 31% in the case of output and 33% in the case of consumption growth. Moreover, this model implies that if such an increase took place it was fairly modest, with most of the mass below 1.25.

4.4. Variance Decomposition

Table III shows the relative contribution of the different shocks to the unconditional variance of real GDP growth, for the specification with both stochastic volatility and Student's t components (with $\underline{\lambda} = 6$). Since volatility is time varying, we compute the unconditional variance at each point in time t

[19] Clark (2009) also investigates the end of the Great Moderation using a univariate model with time-varying coefficients and innovation volatilities, and finds that the implied unconditional volatility of several macroeconomic variables increased notably during the Great Recession. Barnett and Chauvet (2008) also discuss the possibility that the Great Moderation period is over.

DOI: 10.1002/jae

Table III. Variance decomposition for real GDP growth

	g	b	μ	z	λ_f	λ_w	r^m
σ_{1964}	0.179	0.439	0.087	0.163	0.037	0.017	0.079
σ_{1981}	0.204	0.348	0.074	0.101	0.032	0.022	0.220
σ_{1994}	0.179	0.394	0.108	0.142	0.039	0.046	0.092
σ_{2007}	0.167	0.370	0.104	0.149	0.046	0.056	0.109
σ_{2011}	0.167	0.360	0.102	0.148	0.048	0.051	0.123

Note: The table shows the relative contribution of the different shocks to the unconditional variance of real GDP, for the specification with stochastic volatility and Student's t-distributed shocks. Since volatility is time varying we evaluate this contribution at different points in time assuming that volatility will be fixed at that period's level from then on: 1964 (beginning of sample), 1981 (peak of the high volatility period), 1994 (great moderation), 2007 (pre-great recession) and 2011 (end of sample).

assuming that the time-varying component of the standard deviations of the shocks is going to remain equal to the estimated value of σ_t forever after, as we did in the previous section. The table shows the relative contribution of each structural shock to this unconditional variance at five different points in time: 1964:Q4 (beginning of sample), 1981:Q1 (peak of the high volatility period), 1994:Q1 (great moderation), 2007:Q1 (pre-great recession) and 2011:Q1 (end of sample). In sum, we find that the shocks that are most important for explaining real GDP growth are those that exhibit fat tails, like the discount rate (b), marginal efficiency of investment (μ) and TFP (z) shocks.[20]

4.5. Robustness

This section discusses the robustness of our results to different assumptions and estimation approaches for the stochastic volatility component, and to different subsamples. The results are shown in online Appendix A. The lesson from these robustness exercises is that the two main results of the paper, namely (i) there is strong evidence in favor of Student's t-distributed shocks and (ii) accounting for fat tails changes the inference about low frequency movements in volatility, are very robust.

4.5.1. Robustness to Different Estimation Approaches and Assumptions for the Stochastic Volatility Component

Results using the JPR algorithm are quite different from those obtained with the KSC algorithm in terms of the inference about stochastic volatility. Table A.18 shows the log-marginal likelihoods under this approach, and they are markedly higher for all SV specifications than those computed using the KSC approach (Table I). Inference about time variation in volatility is also quite different, as reported in the bottom panel of Figure A.8 for output growth, especially as far as the spike in volatility in the late 1970s and early 1980s is concerned. As a result of this spike, the ratios are much larger for both the SV-only and the SVTD specifications. The difference between results under the JPR and KSC algorithm are puzzling, considering that the model/prior specification is identical, and that both estimations appeared to have converged.[21] In spite of these differences, the results are very robust as far as the message of this paper is concerned. Table A.18 shows that the model's fit improves sizably by allowing for fat-tailed shocks, and that the best specification is still the one with the lowest prior for the degrees of freedom ($\underline{\lambda} = 6$). Table A.19 shows that the inference concerning the degrees of freedom of the Student's t-distribution is virtually identical to those reported for the KSC algorithm for

[20] Table A.4 in online Appendix A shows the variance decomposition for the other specifications (Gaussian, TD only and SV only).

[21] Section A.8 discusses the convergence results for the KSC algorithm, which are available in online Appendix A. Convergence results for the JPR algorithm are available upon request.

DOI: 10.1002/jae

the best specification. The top panel of Figure A.8 shows that the importance of Student's t-distributed shocks in terms of output growth fluctuations is the same as that computed under the other algorithm. The bottom panel shows that allowing for Student's t shocks changes the inference about the size of fluctuations in volatility, perhaps even more than under the alternative algorithm. Figuring out which algorithm—JPR or KSC—proves most reliable for macro applications is arguably an important issue to sort out for applied macroeconomists, but is not the goal of this paper.

We also investigated robustness of the results to the SV-S specification, where following a referee's suggestion we use a prior for ρ_q with $\bar{\rho} = 0.6$ and $\bar{v}_\rho = (0.2)^2/\underline{\omega_q}^2$.[22] Table A.20 shows that the log-marginal likelihood results for the SV-S specification are worse than those for the SV-UR, and that allowing for fat tails improves the model's fit considerably. The results under this specification are largely in line with those of the baseline.[23]

4.5.2. Subsample Analysis

Does evidence in favor of fat-tailed shocks depend on whether the Great Recession is included in the estimation sample? In order to address this question, we re-estimated the model for the subsample ending in the fourth quarter of 2004—the end date of the sample used in JP.[24] Table A.23 shows in columns 2 and 3 the log-marginal likelihood for all the specifications considered above but estimated on the shorter subsample. The results for this subsample are in line with the results for the full sample: including Student's t-distributed shocks improves fit, regardless of whether we also consider stochastic volatility, and the lower the prior mean for the degrees of freedom, the higher the log-marginal likelihood we obtain. The posterior means of the degrees of freedom of the Student's t-distribution (Table A.24) are also mostly in line with those for the full sample. The top panel of Figure A.10 shows the importance of Student's t-distributed shocks in terms of output growth fluctuations. The results again agree with those for the full sample. The bottom panels of Figure A.10 provide evidence regarding time variation in volatility. It appears that for the short sample this is more muted than for the full sample, although the result that accounting for fat tails leads to less volatile estimates of the low-frequency SV process still holds.[25] Figure A.11 reproduces some of the results in Figure A.10 using the JPR algorithm. We find more evidence of time variation in volatility using this algorithm, as we did for the full sample, but the results concerning the importance of fat tails are unchanged.[26]

We also investigate how much the results depend on the period of high volatility experienced in the late 1970s by focusing on two subsamples starting in the Great Moderation period: 1984:Q1 and 1991:Q4. Table A.23 shows in columns 4–7 the log-marginal likelihoods for these subsamples. Again, we find that having Student's t-distributed shocks improves fit, regardless of whether we also consider stochastic volatility; and the lower the prior mean for the degrees of freedom, the higher the log-marginal likelihood we obtain. The posterior means of the degrees of freedom of the Student's t-distribution (Tables A.26 and A.27) are mostly consistent with previous results, and the top panels

[22] We also tried a prior for ρ_q that does not depend on $\underline{\omega_q}$, i.e. $\mathcal{N}(0.6, (0.2)^2)$, and obtained very similar results.

[23] Table A.22 shows the posterior estimates of the degrees of freedom of the Student's t-distribution; Figure A.9 is the analog of Figure A.8 for this specification; Table A.22 shows the posterior estimates of the SV persistence parameter ρ_q. It is worth noting that there are sizable differences across specifications in the marginal likelihood improvement due to having SV on top of Student's t shocks: this improvement is of the order of roughly 11, 47 and 5 log points for the baseline, JPR and SV-S specifications, respectively (see Tables I, A.18 and A.20).

[24] Note that our model, data and prior on ω^2 differ from JP in important ways; hence we would not expect to replicate their results.

[25] Figures A.12 and A.13 provide the posterior for $\tilde{h}_{q,t}^{-1/2}$ and $\sigma_{q,t}$, respectively, for this subsample. These figures broadly look consistent with those for the full sample, i.e. they appear to be the decomposed version of Figures A.6 and A.7, respectively.

[26] Note that under the KSC algorithm the extent to which the Great Moderation is overstated by excluding Student's t shocks is sensitive to the sample. This is simply because under this algorithm we do not find much time variation in volatility even in the SV-only case for the sample ending in 2004. However, under the JPR algorithm time variation in volatility is much more pronounced also for this shorter sample, and therefore including Student's t shocks makes more of a difference in terms of assessing the magnitude of the Great Moderation.

DOI: 10.1002/jae

of Figures A.14 and A.15 show that the impact of fat-tailed shocks on output fluctuations is in line with the results of the full sample. Adding stochastic volatility does not substantially improve the fit, if at all, in these subsamples. This suggests that, even if the Great Recession is included in the sample, most of the evidence in favor of stochastic volatility for the full sample is due to the shift in volatility from the late 1970s to the 1990s. The bottom panels of Figures A.14 and A.15 are consistent with this view, as they show little evidence of changes in volatility for output growth.

5. CONCLUSIONS

We provide strong evidence that the Gaussianity assumption in DSGE models is counterfactual, even after allowing for low-frequency changes in the volatility of shocks. It is important to point out a number of caveats regarding our analysis. First, we allow for excess kurtosis but not for skewness. The plots in Figure 1 make it plain that most large shocks occur during recessions, implying that skewness may also be a salient feature of the shock distribution. Müller (2013) describes some of the dangers associated with departures from Gaussianity when the alternative shock distribution is also misspecified. Importantly for our analysis, not allowing for skewness may lead to an underestimation of the importance of fat tails during recessions, as we only estimate the average amount of kurtosis. Second, we allow for permanent (random walk) and i.i.d. (Student's t-distribution) changes in the variance of the shocks. These assumptions are convenient, but also extreme. Our main point is that, together with low-frequency changes in the standard deviation of shocks, there are also short-run spikes in volatility. So far, the literature for the US has mainly focused on the former phenomenon; in this paper we emphasize the latter. Still, in future research it may be important to relax the assumption that these short-run spikes are identically distributed over time. Third, in order to study the full implications of fat-tailed shocks on the macroeconomy we need to use nonlinear models, as we discuss in the Introduction.[27] Finally, we have not investigated the implication of allowing for fat tails in terms of forecasting. Do DSGE models with Student's t shocks have a superior forecasting performance relative to models with Gaussian shocks in terms of root mean squared errors? Or are the gains in marginal likelihood mainly reflected in an improved precision of the forecast distribution? We leave these important extensions for future research.

ACKNOWLEDGEMENTS

We thank the editor, Herman Van Dijk, and three anonymous referees for very helpful comments, as well as Marek Jarocinski, Alejandro Justiniano, Ulrich Mueller, Andriy Norets, and seminar participants at the Bank of France, 2011 EEA, ECB, 2011 ESOBE, 2012 ESSIM, FRB Chicago, FRB St Louis Econometrics Workshop, 2012 Fall Meeting of the NBER DSGE Group, RCEF 2012, Seoul National University (conference in honor of Chris Sims), 2011 SCE, for useful comments and suggestions. The views expressed in this paper do not necessarily reflect those of the Federal Reserve Bank of New York, the Federal Reserve Bank of San Francisco or the Federal Reserve System.

REFERENCES

An S, Schorfheide S. 2007. Bayesian analysis of DSGE models: rejoinder. *Econometric Reviews* **26**: 211–219.
Ascari G, Fagiolo G, Roventini A. 2012. Fat tails distributions and business cycle models. Institute of Economics, Scuola Superiore Sant'Anna Working paper 2012-07, Pisa, Italy.

[27] The optimal policy response to rare large shocks is another important issue that deserves attention. The financial econometrics literature has studied the implication of alternative hedging strategies *vis-à-vis* fat-tailed shocks (e.g. Bos *et al.*, 2000).

DOI: 10.1002/jae

Barnett WA, Chauvet M. 2008. The end of the Great Moderation? Working paper. University of Kansas, University of California at Riverside, Kansas (Missouri) and Riverside (California).
Bernanke B, Gertler M, Gilchrist S. 1999. The financial accelerator in a quantitative business cycle framework. In *Handbook of Macroeconomics*, Taylor JB, Woodford M (eds), Vol. 1C. North-Holland: Amsterdam; 1341–1393.
Bos CS, Mahieu RJ, van Dijk HK. 2000. Daily exchange rate behaviour and hedging of currency risk. *Journal of Applied Econometrics* **15**: 671–696.
Chib S, Greenberg E. 1994. Bayes inferences in regression models with ARMA(p,q) errors. *Journal of Econometrics* **64**: 183–206.
Chib S, Ramamurthy S. 2014. DSGE Models with Student-*t* errors. *Econometric Reviews* **33**: 152–171.
Chib S, Nardari F, Shephard N. 2002. Markov chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics* **108**: 281–316.
Christiano LJ. 2007. Comment on 'On the fit of New Keynesian models' by Del Negro, Schorfheide, Smets, Wouters. *Journal of Business and Economic Statistics* **25**: 143–151.
Christiano L, Eichenbaum M, Evans CL. 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy* **113**: 1–45.
Christiano L, Motto R, Rostagno M. 2009. Financial factors in economic fluctuations. Working paper. Northwestern University and European Central Bank, Evanston (Illinois) and Frankfurt (Germany).
Clark TE. 2009. Is the Great Moderation over? An empirical analysis. *Economic Review* **QIV**: 5–42.
Cúrdia V, Del Negro M, Greenwald DL. 2012. Rare shocks, great recessions. Federal Reserve Bank of New York Staff Reports, 585.
Del Negro M, Primiceri G. 2013. Time-varying structural vector autoregressions and monetary policy: a corrigendum. Federal Reserve Bank of New York Staff Reports, 619.
Del Negro M, Schorfheide F. 2008. Forming priors for DSGE models (and how it affects the assessment of nominal rigidities). *Journal of Monetary Economics* **55**: 1191–1208.
Del Negro M, Schorfheide F. 2014. DSGE model-based forecasting. In *Handbook of Economic Forecasting*, Elliott G, Timmermann A (eds), Vol. 2. Elsevier: Amsterdam; 57–140.
Dewachter H, Wouters R. 2012. Endogenous risk in a DSGE model with capital-constrained Financial intermediaries. Working Paper 235. National Bank of Belgium, Brussels (Belgium).
Durbin J, Koopman SJ. 2002. A simple and efficient simulation smoother for state space time series analysis. *Biometrika* **89**: 603–616.
Edge R, Gürkaynak R. 2010. How useful are estimated DSGE model forecasts for central bankers? *Brookings Papers of Economic Activity* **41**: 209–259.
Fernández-Villaverde J, Rubio-Ramírez JF. 2007. Estimating macroeconomic models: a likelihood approach. *Review of Economic Studies* **74**: 1059–1087.
Gelman A, Rubin DB. 1992. Inference from iterative simulation using multiple sequences. *Statistical Science* **7**: 457–472.
Gelman A, Carlin JB, Stern HS, Rubin DB. 2004. *Bayesian Data Analysis*. Chapman & Hall/CRC: Boca Raton, FL.
Geweke J. 1992. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In *Bayesian Statistics*, Bernardo JM, Berger J, Dawid AP, Smith JFM (eds). Oxford University Press: Oxford; 169–193.
Geweke J. 1993. Bayesian treatment of the independent Student-*t* linear model. *Journal of Applied Econometrics* **8**: S19–S40.
Geweke J. 1994. Priors for macroeconomic time series and their application. *Econometric Theory* **10**: 609–632.
Geweke J. 1999. Using simulation methods for Bayesian econometric models: inference, development, and communication. *Econometric Reviews* **18**: 1–126.
Geweke J. 2005. *Contemporary Bayesian Econometrics and Statistics*. Wiley: Hoboken, NJ.
Greenwood J, Hercovitz Z, Krusell P. 1998. Long-run implications of investment-specific technological change. *American Economic Review* **87**: 342–362.
Jacquier E, Polson NG, Rossi PE. 1994. Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics* **12**: 371–389.
Jacquier E, Polson NG, Rossi PE. 2004. Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics* **122**: 185–212.
Justiniano A, Primiceri G. 2008. The time-varying volatility of macroeconomic fluctuations. *American Economic Review* **98**: 604–641.
Justiniano A, Primiceri GE, Tambalotti A. 2009. Investment shocks and business cycles. National Bureau of Economic Research Working Paper 15570, Cambridge (Massachusetts).

DOI: 10.1002/jae

Kim S, Shephard N, Chib S. 1998. Stochastic volatility: likelihood inference and comparison with ARCH models. *Review of Economic Studies* **65**: 361–393.
Liu Z, Waggoner DF, Zha T. 2011. Sources of macroeconomic fluctuations: a regime-switching DSGE approach. *Quantitative Economics* **2**: 251–301.
Müller UK. 2013. Risk of Bayesian inference in misspecified models, and the sandwich covariance matrix. *Econometrica* **81**: 1805–1849.
Omori Y, Chib S, Shephard N, Nakajima J. 2007. Stochastic volatility with leverage: fast and efficient likelihood inference. *Journal of Econometrics* **140**: 425–449.
Sims CA. 2002. Solving linear rational expectations models. *Computational Economics* **20**(1-2): 1–20.
Smets F, Wouters R. 2003. An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European Economic Association* **1**: 1123–1175.
Smets F, Wouters R. 2007. Shocks and frictions in US business cycles: a Bayesian DSGE approach. *American Economic Review* **97**: 586–606.

DOI: 10.1002/jae