Rare Shocks, Great Recessions

Vasco Cúrdia
Marco Del Negro
Daniel L. Greenwald

FRBNY
Staff
REPORTS

Rare Shocks, Great Recessions
Vasco Cúrdia, Marco Del Negro, and Daniel L. Greenwald
Federal Reserve Bank of New York Staff Reports, no. 585
December 2012; revised June 2013
JEL classification: C32, E32

Abstract

We estimate a DSGE 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 (2007)
model estimated on the usual set of macroeconomic time series over the 1964-2011 period
indicate that 1) 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 2) the estimated degrees of freedom
are quite low for several shocks that drive U.S. 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.

Key words: Bayesian analysis, DSGE models, fat tails, stochastic volatility, Great Recession

_____

1   Introduction

Great Recessions do not happen every decade — this is why they are dubbed "great" in the first place. To the extent that 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 (1993)'s Gibbs sampler for a linear model to the DSGE framework. The paper is closely related to Chib and Ramamurthy (2011) 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 Chib and Ramamurthy (2011)'s 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] 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 data set starts in 1964Q4 and ends in 2011Q1, but we also consider a sub-sample ending in 2004Q4 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

---

[1]Another difference between this paper and Chib and Ramamurthy (2011) is that Chib and Ramamurthy use a simple three equation New Keynesian model, while we use the full Smets and Wouters (2007) model.

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 deviations 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 (e.g., 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, that is, 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 70s and 80s 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

---

[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 (forthcoming)). 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.

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 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 non-linear models. First, if shocks have fat tails, linearization may simply produce a poor approximation of the full model. Second, non-linearities may explain away the fat tails: what we capture as large rare shocks may in fact be Gaussian shocks whose effect is amplified through a non-linear propagation mechanism. In fact, the extent to which non-linearities 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 non-linearly a model with capital-constrained financial intermediaries and provide very interesting evidence suggesting that the non-linear propagation of shocks may induce outcomes that resemble those due to fat-tailed shocks in a linear model.

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 the 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 the 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 reevaluate 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)'s for a linear model with Student's t distributed shocks (see also Geweke (1994), and Geweke (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 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 Jacquier et al. (1994) (henceforth, JPR).

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 $t = 1, .., T$, 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 $\underline{\nu}$:

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

where $\underline{\lambda}$ is the mean and $\underline{\nu}$ is the number of degrees of freedom (Geweke (2005) assumes a Gamma 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

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

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:

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

The prior distribution for ω_q^2 is an inverse Gamma $\mathcal{IG}(\nu_\omega/2, \nu_\omega \underline{\omega}^2/2)$, that is:

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

We consider two types of priors for ρ_q:

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

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 (3): $\sigma_{q,t}$ allows for slow-moving trends in volatility, while $\tilde{h}_{q,t}$ 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 ($\tilde{\sigma}_{q,t}$) and one that is high frequency — in our case, i.i.d. ($\tilde{h}_{q,t}$). The literature so far has focused on the former and largely ignored the latter. Of course the i.i.d. assumption for the high frequency movements in volatility is very stark.[4] In future research it will be worth trying to relax this assumption. Considering two overlapping

---

[3]We do not have measurement errors at all in this model, let alone Student's t ones. From looking at the plots of the shocks in Figure 1 (and the posterior estimates of $\tilde{h}_t$ in Figure B.5) it appears that the many large shocks coincide with macro events, and are therefore unlikely to just represent measurement error.

[4]There is also an ongoing debate in empirical macro on whether the low frequency movements in volatility should be modeled using autoregressive processes or regime-switching models (e.g., Liu et al. (2011)).

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 $\tilde{\sigma}_{q,0}$:

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

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 $\{x_{1,1}, .., x_{1,T}, .., x_{\bar{q},T}\}$ 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 $\{\tilde{h}_{1,1}, .., \tilde{h}_{1,T}, .., \tilde{h}_{\bar{q},T}\}$.

2.1 The Gibbs-Sampler

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

$$p(y_{1:T}|s_{1:T}, \theta)p(s_{1:T}|\varepsilon_{1:T}, \theta)p(\varepsilon_{1:T}|\tilde{h}_{1:T}, \tilde{\sigma}_{1:T}, \theta)p(\tilde{h}_{1:T}|\lambda_{1:\bar{q}})$$

$$p(\tilde{\sigma}_{1:T}|\rho_{1:\bar{q}}, \omega_{1:\bar{q}}^2)p(\lambda_{1:\bar{q}})p(\rho_{1:\bar{q}}|\omega_{1:\bar{q}}^2)p(\omega_{1:\bar{q}}^2)p(\theta), \quad (12)$$

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

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

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

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

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

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

where

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

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

2.1.1 The KSC Version

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

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

In this section we describe the sampler under the KSC approach, and in the next section we consider the JPR approach. The key insight of KSC is that if $p(\varepsilon_{1:T}|\tilde{h}_{1:T}, \tilde{\sigma}_{1:T}, \theta)$ in (17) were linear in $\tilde{\sigma}_{1:T}$ and Gaussian, one could use standard state-space methods for drawing $\tilde{\sigma}_{1:T}$. In fact, taking squares and then logs of (3) one can see that $\varepsilon_{q,t}^* = \log(\sigma_q^{-2}\tilde{h}_{q,t}\varepsilon_{q,t}^2 + c)$ (where $c = .001$ is an offset constant) is linear in $\sigma_{q,t}$:

$$\varepsilon_{q,t}^* = 2\tilde{\sigma}_{q,t} + \eta_{q,t}^*, \;\; \eta_{q,t}^* \simeq \log(\eta_{q,t}^2), \tag{18}$$

but is not Gaussian, since $\eta_{q,t}^* \sim \log(\chi_1^2)$. KSC suggest approximating the distribution of $\eta_{q,t}^*$ using a mixture of normals:[5]

$$p(\eta_{q,t}^*) \simeq \sum_{k=1}^{K} \pi_k^* \mathcal{N}(m_k^*, \nu_k^{*\,2}) \tag{19}$$

, or equivalently, $\eta_{q,t}^* \,|\, \varsigma_{q,t} = k \sim \mathcal{N}(m_k^* - 1.2704, \nu_k^{*\,2}), \;\; Pr(\varsigma_{q,t} = k) = \pi_k^*$.

Call $\vartheta = \{\theta, s_{1:T}, \tilde{h}_{1:T}, \lambda_{1:\bar{q}}, \rho_{1:\bar{q}}, \omega_{1:\bar{q}}^2, \varepsilon_{1:T}\}$ all unobservables other than $\varsigma_{1:T}$ and $\tilde{\sigma}_{1:T}$. Del Negro and Primiceri (2013) recognize that in standard macro models it is often difficult to draw from $p(\vartheta|\tilde{\sigma}_{1:T}, \varsigma_{1:T}, y_{1:T})$, and that therefore a Gibbs sampler such as the following one would not work: i) draw ϑ from $\vartheta|\tilde{\sigma}_{1:T}, \varsigma_{1:T}, y_{1:T}$, ii) draw $\tilde{\sigma}_{1:T}$ from $\tilde{\sigma}_{1:T}|\vartheta, \varsigma_{1:T}, y_{1:T}$, iii) draw $\varsigma_{1:T}$ from $\varsigma_{1:T}|\tilde{\sigma}_{1:T}, \vartheta, y_{1:T}$.[6] They suggest the following sampler instead:

(1) draw $\tilde{\sigma}_{1:T}$ from $\tilde{\sigma}_{1:T}|\vartheta, \varsigma_{1:T}, y_{1:T}$;

(2) draw $\vartheta, \varsigma_{1:T}$ from $\vartheta, \varsigma_{1:T}|\tilde{\sigma}_{1:T}, y_{1:T}$, which can be divided into two substeps:

    (2-1) draw ϑ from the marginal $\vartheta|\tilde{\sigma}_{1:T}, y_{1:T}$;

    (2-2) draw $\varsigma_{1:T}$ from the conditional $\varsigma_{1:T}|\vartheta, \tilde{\sigma}_{1:T}, y_{1:T}$.

---

[5] We follow Omori et al. (2007) in using a 10-mixture approximation, as opposed to the 7-mixture approximation adopted in Kim et al. (1998). The parameters that optimize this approximation, namely $\{\pi_k^*, m_k^*, \nu_k^*\}_{k=1}^{K}$, are given in Omori et al. (2007). Note that these parameters are independent of the specific application.

[6] In several macro applications, including previous drafts of this paper, the sampling procedure is described in this way except that the step $\vartheta|\tilde{\sigma}_{1:T}, \varsigma_{1:T}, y_{1:T}$ is mistakenly replaced with $\vartheta|\tilde{\sigma}_{1:T}, y_{1:T}$.

2.1.2 The JPR Version

Under the JPR approach the volatilities can be drawn directly from (17). The Gibbs sampler is therefore simply:

(1) draw $\tilde{\sigma}_{1:T}$ from $\tilde{\sigma}_{1:T}|\vartheta, y_{1:T}$;

(2) draw ϑ from $\vartheta|\tilde{\sigma}_{1:T}, y_{1:T}$.

For step (1), $\tilde{\sigma}_{1:T}$ is drawn in an additional Gibbs step, in which each $\tilde{\sigma}_t$ is drawn conditional on $(\theta, y_{1:T}, \tilde{\sigma}_{-t})$, where $\tilde{\sigma}_{-t}$ contains all elements of $\tilde{\sigma}_{1:T}$ except for $\tilde{\sigma}_t$. Each $\tilde{\sigma}_t$ is drawn from an inverse gamma proposal distribution, and is then accepted or rejected in a Metropolis Hastings step.[7]

Note that step (2-2) and (2) in the KSC and JPR samplers, respectively, are identical. The next section describes this step in detail.

2.1.3 Drawing from $\vartheta|\tilde{\sigma}_{1:T}, y_{1:T}$

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

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

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

$$
\begin{aligned}
p(\theta|&\tilde{h}_{1:T}, \tilde{\sigma}_{1:T}, \lambda_{1:\bar{q}}, \rho_{1:\bar{q}}, \omega^2_{1:\bar{q}}, y_{1:T}) \\
&\propto \left[\int p(y_{1:T}|s_{1:T}, \theta)p(s_{1:T}|\varepsilon_{1:T}, \theta)p(\varepsilon_{1:T}|\tilde{h}_{1:T}, \tilde{\sigma}_{1:T}, \theta) \cdot d(s_{1:T}, \varepsilon_{1:T}) \right] p(\theta) \\
&= p(y_{1:T}|\tilde{h}_{1:T}, \tilde{\sigma}_{1:T}, \theta)p(\theta)
\end{aligned}
$$

(20)

where

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

---

[7]As noted in Jacquier et al. (1994), the posterior $\tilde{\sigma}_t|\tilde{\sigma}_{-t}, \vartheta, y_{1:T}$ can also be easily sampled using a lognormal density and then applying a Metropolis Hastings step. However, JPR warn that the tails of the lognormal distribution may not be thick enough for good sampling. In line with these results, our experiments with a log-normal proposal produced largely similar results, but substantially worse convergence properties, using the criteria of Section A.4.

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

$$\varepsilon_t | \tilde{h}_{1:T}, \tilde{\sigma}_{1:T} \sim \mathcal{N}(0, \Delta_t), \tag{21}$$

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.[8]

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

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

$$p(\varepsilon_{1:T} | \tilde{h}_{1:T}, \tilde{\sigma}_{1:T}, \theta) p(\tilde{h}_{1:T} | \lambda_{1:\bar{q}})$$
$$\propto \prod_{q=1}^{\bar{q}} \prod_{t=1}^{T} \tilde{h}_{q,t}^{(\lambda_q - 1)/2} \exp(-\left[\lambda_q + \varepsilon_{q,t}^2 / \sigma_{q,t}^2\right] \tilde{h}_{q,t}/2), \tag{22}$$

which implies

$$\left[\lambda_q + \varepsilon_{q,t}^2 / \sigma_{q,t}^2\right] \tilde{h}_{q,t} | \theta, \varepsilon_{1:T}, \tilde{\sigma}_{1:T}, \lambda_q \sim \chi^2(\lambda_q + 1). \tag{23}$$

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

$$p(\tilde{h}_{1:T} | \lambda_{1:\bar{q}}) p(\lambda_{1:\bar{q}}) \propto \prod_{q=1}^{\bar{q}} \left((\underline{\lambda}/\underline{\nu})^{\underline{\nu}} \Gamma(\underline{\nu}) \right)^{-1} \left[2^{\lambda_q/2} \Gamma(\lambda_q/2) \right]^{-T} \lambda_q^{T\lambda_q/2 + \underline{\nu} - 1}$$
$$\left(\prod_{t=1}^{T} \tilde{h}_{q,t}^{(\lambda_q - 2)/2} \right) \exp\left[-\left(\frac{\underline{\nu}}{\underline{\lambda}} + \frac{1}{2} \sum_{t=1}^{T} \tilde{h}_{q,t} \right) \lambda_q \right]. \tag{24}$$

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

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

$$p(\tilde{\sigma}_{1:T} | \rho_{1:\bar{q}}, \omega_{1:\bar{q}}^2) p(\omega_{1:\bar{q}}^2) p(\rho_{1:\bar{q}} | \omega_{1:\bar{q}}^2) \propto \prod_{q=1}^{\bar{q}} (\omega_q^2)^{-\frac{\nu + T - 1}{2} - 1}$$
$$\exp\left[-\frac{\nu \underline{\omega}^2 + \sum_{t=2}^{T} (\tilde{\sigma}_{q,t} - \rho_q \tilde{\sigma}_{q,t-1})^2}{2\omega_q^2} \right] p(\tilde{\sigma}_{q,1} | \rho_q, \omega_q^2) p(\rho_q | \omega_q^2), \tag{25}$$

---

[8]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.

where $p(\tilde{\sigma}_{q,1}|\rho_q, \omega_q^2)$ 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:

$$\omega_q^2|\tilde{\sigma}_{1:T}, \cdots \sim \mathcal{IG}\left(\frac{\nu + T}{2}, \frac{1}{2}\left(\nu\underline{\omega}^2 + \sum_{t=2}^{T}(\tilde{\sigma}_{q,t} - \tilde{\sigma}_{q,t-1})^2 + \tilde{\sigma}_{q,1}^2\right)\right). \qquad (26)$$

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 $p(\tilde{\sigma}_1|\rho_q, \omega_q^2)$. We therefore use the Metropolis-Hastings step proposed by Chib and Greenberg (1994). Specifically, we use as proposal density the usual Normal-Inverted Gamma distribution, that is,

$$\begin{aligned}
\omega_q^2|\tilde{\sigma}_{1:T}, \cdots &\sim \mathcal{IG}\left(\frac{\nu+T-1}{2}, \frac{1}{2}\left(\nu\underline{\omega}^2 + \sum_{t=2}^{T}\tilde{\sigma}_{q,t}^2 + \bar{v}_\rho^{-1}\bar{\rho}^2 - \hat{V}_q^{-1}\hat{\rho}_q^2\right)\right), \\
\rho_q|\omega_q^2, \tilde{\sigma}_{1:T}, \cdots &\sim \mathcal{N}\left(\hat{\rho}_q, \omega_q^2\hat{V}_q\right), \text{ i.i.d. across } q,
\end{aligned} \qquad (27)$$

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

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

with $(\rho^{(j-1)}, \omega_q^{2\ (j-1)})$ and $(\rho^{(*)}, \omega_q^{2\ (*)})$ being the draw at the $(j-1)^{\text{th}}$ 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). It 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. Due to space constraints, and to the fact that the model is the same as in SW, we relegate the detailed model description to appendix B (it is also available in Cúrdia et al. (2012), the working paper version). Since this is a paper about shocks, we nonetheless need to 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 an ARMA(1,1) process. Innovations to the TFP process will 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 for our vector of observables y_t, 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 real wage growth, hours, inflation, and interest rates, all measured in percent (the data construction is the same as in SW). Section B.2 in appendix B provides further details on the data. In our benchmark specification we use data from 1964Q4 to 2011Q1, but we also consider a variety of sub-samples. Table B.1 in appendix B shows the priors for the DSGE model parameters, which coincide with those used in SW.[9]

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 through 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 (Section 2.1.1). Section

---

[9] 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.

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 B which is available online and in the working paper version (Cúrdia et al. (2012)), and we briefly discuss them in Sections A.2 and A.3.

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.[10] 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 (5). If λ is high, this prior concentrates around one, and the likelihood of observing large

---

[10]A referee asked to see how the specifications with Student's t distributed shocks and stochastic volatility deal with these large innovations — that is, 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 B.6 and B.5 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.2, 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 70s/early 80s are due mainly to high volatility during that period (high values of $\tilde{\sigma}_t$).

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 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
∞	.54	.012	$1e^{-4}$
15	1.14	.13	.02
9	1.57	.28	.06
6	2.08	.54	.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.[11] Figure 2 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 1 shows the log-marginal likelihood — the standard measure of fit in a Bayesian framework — for models with different assumptions on the shocks distribution.[12] We consider four different combinations: i) Gaussian shocks with constant volatility (baseline), ii) Gaussian shocks with time-varying volatility, iii) Student's

---

[11]Results obtained using $\underline{\nu} = 1$ are very similar.
[12]Appendix A.1 provides details on the computation of the marginal likelihood for these models.

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).[13] Finally, our prior for the innovations in stochastic volatility is an $\mathcal{IG}$ prior with mode at $(.01)^2$ and .1 degrees of freedom.[14]

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 1, 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 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.[15]

---

[13]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 follow a stationary autoregressive process (SV-S case), but we discuss the results for the SV-S case in Section 4.5.1.

[14]Specifically, in the $\mathcal{IG}(\nu_\omega/2, \nu_\omega\underline{\omega}^2/2)$ 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 $(.01)^2$. 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 ($\nu_\omega/2 = .1$) 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 B.1 in the appendix B provides a plot of the prior distribution for ω^2.

[15]The difference in log-marginal likelihoods between the Gaussian/constant-volatility model and the SV model only 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.

Our second piece of evidence comes from the posterior distribution of the degrees of freedom λ. Table 2 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 2 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 because the posterior distribution becomes more skewed to the right (that is, 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 seventies/early eighties 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 (23):

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

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.[16]

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 panel of Figure 3 shows these counterfactual histories for output, consumption, and investment growth. For all plots the magenta solid lines are the median counterfactual paths, the magenta dashed lines represent the 90% bands, and the solid black lines represent the actual data. The right 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.[17]

The left 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

---

[16]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.

[17] 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. So one should not interpret the counterfactual rolling window standard deviations as measures of "true" underlying volatility.

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.[18] Figures B.2 and B.3 in appendix B 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 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 '70s and early '80s were more volatile than the Great Moderation period at least in part because rare shocks took place.[19] 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 reevaluate the magnitude of low-frequency changes in volatility.[20] Inference about the stochastic volatility

---

[18] Dewachter and Wouters (2012) solve non-linearly 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 non-linear model. They then compare the path of several variables under the non-linear and linear solution methods. These two paths for investment growth look strikingly similar to the actual (for the non-linear method) and counterfactual (for the linear method) paths of investment growth in Figure 3, suggesting that the non-linear propagation of shocks may induce outcomes that resemble those due to fat-tailed shocks in a linear model.

[19] 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.

[20] 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).

is conducted using state-space methods under the KSC procedure, where

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

is the measurement equation (with c a small constant), and equation (8) is the transition equation. Intuitively, the estimated time-varying volatilities will try to fit the time series $\log(\sigma^{-2}\tilde{h}_t\varepsilon_t^2 + c)$. 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_ts. 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.[21] Specifically, Figure 4 shows the 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 (that is, 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 70s and early 80s), 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.

---

[21]Figure B.6 in appendix B 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.

The middle panel of Figure 4 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 B.2 and B.3 in appendix B 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 3. 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?[22] The bottom panels of Figure 4 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

---

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

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 3 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 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: 1964Q4 (beginning of sample), 1981Q1 (peak of the high volatility period), 1994Q1 (great moderation), 2007Q1 (pre-great recession) and 2011Q1 (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.[23]

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 sub-samples. The results are shown in appendix B. 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 B.18 shows the log-marginal likelihoods under this approach, and they are markedly higher for all SV speci-

---

[23]Table B.4 in appendix B shows the variance decomposition for the other specifications (Gaussian, TD only, and SV only).

fications than those computed using the KSC approach (Table 1). Inference about time-variation in volatility is also quite different, as reported in the bottom panel of Figure B.7 for output growth, especially as far as the spike in volatility in the late 70s and early 80s 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.[24] In spite of these differences, the results are very robust as far as the message of this paper is concerned. Table B.18 shows that the model's fit improves sizably by allowing for fat tailed shocks, and that best specification is still the one with the lowest prior for the degrees of freedom ($\underline{\lambda} = 6$). Table B.19 shows that the inference about the degrees of freedom of the Student's t distribution are virtually identical to those reported for the KSC algorithm for the best specification. The top panel of Figure B.7 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 $\bar{\rho} = .6$ and $\bar{v}_\rho = (.2)^2/\underline{\omega_q}^2$.[25] Table B.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.[26]

---

[24]Section A.4 discusses the convergence results for the KSC algorithm, which are available in appendix B. Convergence results for the JPR algorithm are available upon request.

[25]We also tried a prior for ρ_q that does not depend on $\underline{\omega_q}$, i.e., $\mathcal{N}(.6, (.2)^2)$, and obtained very similar results.

[26]Table B.22 shows the posterior estimates of the degrees of freedom of the Student's t distribution; Figure B.8 is the analog of Figure B.7 for this specification; Table B.22 shows the posterior estimates of the SV persistence parameter ρ_q.

4.5.2 Sub-sample Analysis

Does evidence in favor of fat-tailed shocks depends on whether the Great Recession is included in the estimation sample? In order to address this question, we re-estimated the model for the sub-sample ending in the fourth quarter of 2004 — the end date of the sample used in JP.[27] Table B.23 shows in columns two and three the log-marginal likelihood for all the specifications considered above but estimated on the shorter sub-sample. The results for this sub-sample 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 B.24) are also mostly in line with those for the full sample. The top panel of Figure B.9 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 B.9 provide evidence regarding time-variation in volatility. It appears that for the short sample this is more muted than in the full sample, although the result that accounting for fat tails leads to less volatile estimates of the low-frequency SV process still holds.[28] Figure B.10 reproduces some of the results in Figure B.9 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.

We also investigate how much the results depend on the period of high volatility experienced in the late 1970s by focusing on two sub-samples starting in the Great Moderation period: 1984Q1 and 1991Q4. Table B.23 shows in columns four through seven the log-marginal likelihoods for these sub-samples. 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 B.26 and B.27) are mostly consistent with previous results, and the top panels of

---

[27]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.

[28]Figures B.11 and B.12 provide the posterior for $\tilde{h}_{q,t}^{-1/2}$ and $\sigma_{q,t}$, respectively, for this sub-sample. These figures broadly look consistent with those for the full sample, i.e. they appear to be the chopped up version of Figures B.5 and B.6, respectively.

Figures B.13 and B.14 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 sub-samples. 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 B.13 and B.14 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 in Figure 1 make it plain that most large shocks occur during recessions, implying that skewness may also be an salient feature of the shocks distribution. Müller (forthcoming) 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 U.S. 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 non linear models, as we discuss in the introduction.[29] 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?

---

[29]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-á-vis* fat-tailed shocks (e.g., Bos et al. (2000)).

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.

References

AN, S. AND F. SCHORFHEIDE (2007): "Bayesian Analysis of DSGE Models–Rejoinder," *Econometric Reviews*, 26, 211 – 219.

ASCARI, G., G. FAGIOLO, AND A. ROVENTINI (2012): "Fat Tails Distributions and Business Cycle Models," *Economix, Department de Travail Working papers*, 2012-07.

BARNETT, W. A. AND M. CHAUVET (2008): "The End of the Great Moderation?" *mimeo, University of Kansas, University of California at Riverside.*

BERNANKE, B., M. GERTLER, AND S. GILCHRIST (1999): "The Financial Accelerator in a Quantitative Business Cycle Framework," in *Handbook of Macroeconomics*, ed. by J. B. Taylor and M. Woodford, North Holland, Amsterdam, vol. 1C.

BOS, C. S., R. J. MAHIEU, AND H. K. VAN DIJK (2000): "Daily Exchange Rate Behaviour and Hedging of Currency Risk," *Journal of Applied Econometrics*, 15, 671–696.

CHIB, S. AND E. GREENBERG (1994): "Bayes inferences in regression models with ARMA(p,q) errors," *Journal of Econometrics*, 64, 183–206.

CHIB, S., F. NARDARI, AND N. SHEPHARD (2002): "Markov chain Monte Carlo methods for stochastic volatility models," *Journal of Econometrics*, 108, 281 316.

CHIB, S. AND S. RAMAMURTHY (2011): "DSGE Models with Student-t errors," *mimeo, Washington University in St. Louis.*

CHRISTIANO, L., M. EICHENBAUM, AND C. L. EVANS (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113, 1–45.

CHRISTIANO, L., R. MOTTO, AND M. ROSTAGNO (2009): "Financial Factors in Economic Fluctuations," *Manuscript, Northwestern University and European Central Bank.*

CHRISTIANO, L. J. (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.

CLARK, T. E. (2009): "Is the Great Moderation over? an empirical analysis," *Economic Review, Federal Reserve Bank of Kansas City*, QIV, 5–42.

CÚRDIA, V., M. DEL NEGRO, AND D. L. GREENWALD (2012): "Rare Shocks, Great Recessions," *Federal Reserve Bank of New York Staff Reports*, 585.

DEL NEGRO, M. AND G. PRIMICERI (2013): "Time-Varying Structural Vector Autoregressions and Monetary Policy: A Corrigendum," *Federal Reserve Bank of New York Staff Reports*, 619.

DEL NEGRO, M. AND F. SCHORFHEIDE (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. AND F. SCHORFHEIDE (forthcoming): "DSGE Model-Based Forecasting," in *Handbook of Economic Forecasting, Volume 2*, ed. by G. Elliott and A. Timmermann, Elsevier.

DEWACHTER, H. AND R. WOUTERS (2012): "Endogenous risk in a DSGE model with capital-constrained Financial intermediaries," *National Bank of Belgium Working Paper*, 235.

DURBIN, J. AND S. J. KOOPMAN (2002): "A Simple and Efficient Simulation Smoother for State Space Time Series Analysis," *Biometrika*, 89, 603–616.

EDGE, R. AND R. GÜRKAYNAK (2010): "How Useful Are Estimated DSGE Model Forecasts for Central Bankers," *Brookings Papers of Economic Activity*, forthcoming.

FERNÁNDEZ-VILLAVERDE, J. AND J. F. RUBIO-RAMÍREZ (2007): "Estimating Macroeconomic Models: A Likelihood Approach," *Review of Economic Studies*, 74, 1059–1087.

GELMAN, A., J. B. CARLIN, H. S. STERN, AND D. B. RUBIN (2004): *Bayesian data analysis*, Chapman & Hall/CRC.

GELMAN, A. AND D. B. RUBIN (1992): "Inference from iterative simulation using multiple sequences," *Statistical science*, 457–472.

GEWEKE, J. (1993): "Bayesian Treatment of the Independent Student-t Linear Model," *Journal of Applied Econometrics,*, 8, S19–S40.

——— (1994): "Priors for Macroeconomic Time Series and Their Application," *Econometric Theory*, 10, 609–632.

——— (1999): "Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication," *Econometric Reviews*, 18, 1–126.

——— (2005): *Contemporary Bayesian Econometrics and Statistics*, Wiley.

GEWEKE, J., J. BERNARDO, J. BERGER, A. DAWID, AND A. SMITH (1992): "Bayesian statistics," *Bayesian statistics*.

GREENWOOD, J., Z. HERCOVITZ, AND P. KRUSELL (1998): "Long-Run Implications of Investment-Specific Technological Change," *American Economic Review*, 87, 342–36.

JACQUIER, E., N. G. POLSON, AND P. E. ROSSI (1994): "Bayesian Analysis of Stochastic Volatility Models," *Journal of Business & Economic Statistics*, 12, 371–389.

——— (2004): "Bayesian analysis of stochastic volatility models with fat-tails and correlated errors," *Journal of Econometrics*, 122, 185 212.

JUSTINIANO, A. AND G. PRIMICERI (2008): "The Time-Varying Volatility of Macroeconomic Fluctuations," *American Economic Review*, 98, 604 – 641.

JUSTINIANO, A., G. E. PRIMICERI, AND A. TAMBALOTTI (2009): "Investment Shocks and Business Cycles," *NBER Working Paper*, 15570.

KIM, S., N. SHEPHARD, AND S. CHIB (1998): "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *Review of Economic Studies*, 65, 361–393.

LIU, Z., D. F. WAGGONER, AND T. ZHA (2011): "Sources of macroeconomic fluctuations: A regime-switching DSGE approach," *Quantitative Economics*, 2, 251–301.

MÜLLER, U. K. (forthcoming): "Risk of Bayesian Inference in Misspecified Models, and the Sandwich Covariance Matrix," *Econometrica*.

OMORI, Y., S. CHIB, N. SHEPHARD, AND J. NAKAJIMA (2007): "Stochastic volatility with leverage: Fast and efficient likelihood inference," *Journal of Econometrics*, 140, 425449.

Sims, C. A. (2002): "Solving Linear Rational Expectations Models," *Computational Economics*, 20(1-2), 1–20.

Smets, F. and R. Wouters (2003): "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, 1, 1123 – 1175.

——— (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97, 586 – 606.

A   Appendix

A.1   Marginal likelihood

The marginal likelihood is the marginal probability of the observed data, and is computed as the integral of (12) with respect to the unobserved parameters and latent variables:

$$
\begin{aligned}
p(y_{1:T}) &= \int p(y_{1:T}|s_{1:T},\theta)p(s_{1:T}|\varepsilon_{1:T},\theta)p(\varepsilon_{1:T}|\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\theta) \\
&\qquad p(\tilde{h}_{1:T}|\lambda_{1:\bar{q}})p(\tilde{\sigma}_{1:T}|\omega^2_{1:\bar{q}})p(\lambda_{1:\bar{q}})p(\omega^2_{1:\bar{q}})p(\theta) \\
&\qquad\qquad d(s_{1:T},\varepsilon_{1:T},\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\lambda_{1:\bar{q}},\rho_{1:\bar{q}},\omega^2_{1:\bar{q}},\theta), \\
&= \int p(y_{1:T}|\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\theta)p(\tilde{h}_{1:T}|\lambda_{1:\bar{q}})p(\tilde{\sigma}_{1:T}|\omega^2_{1:\bar{q}}) \\
&\qquad p(\lambda_{1:\bar{q}})p(\omega^2_{1:\bar{q}})p(\theta)d(\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\lambda_{1:\bar{q}},\rho_{1:\bar{q}},\omega^2_{1:\bar{q}},\theta)
\end{aligned}
\tag{28}
$$

where the quantity

$$
\begin{aligned}
p(y_{1:T}|\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\theta) &= \int p(y_{1:T}|s_{1:T},\theta)p(s_{1:T}|\varepsilon_{1:T},\theta) \\
&\qquad p(\varepsilon_{1:T}|\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\theta)\cdot d(s_{1:T},\varepsilon_{1:T})
\end{aligned}
$$

is computed at step 1a of the Gibb-sampler described above.

We obtain the marginal likelihood using Geweke (1999)'s modified harmonic mean method. If $f(\theta,\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\lambda_{1:\bar{q}},\rho_{1:\bar{q}},\omega^2_{1:\bar{q}})$ is any distribution with support contained in the support of the posterior density such that

$$
\int f(\theta,\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\lambda_{1:\bar{q}},\rho_{1:\bar{q}},\omega^2_{1:\bar{q}})\cdot d(\theta,\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\lambda_{1:\bar{q}},\rho_{1:\bar{q}},\omega^2_{1:\bar{q}}) = 1,
$$

it follows from the definition of the posterior density that:

$$
\begin{aligned}
\frac{1}{p(y_{1:T})} &= \int \frac{f(\theta,\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\lambda_{1:\bar{q}},\rho_{1:\bar{q}},\omega^2_{1:\bar{q}})}{p(y_{1:T}|\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\theta)p(\tilde{h}_{1:T}|\lambda_{1:\bar{q}})p(\tilde{\sigma}_{1:T}|\omega^2_{1:\bar{q}})p(\lambda_{1:\bar{q}})p(\omega^2_{1:\bar{q}})p(\theta)} \\
&\qquad p(\theta,\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\lambda_{1:\bar{q}},\rho_{1:\bar{q}},\omega^2_{1:\bar{q}}|y_{1:T})\cdot d(\theta,\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\lambda_{1:\bar{q}},\rho_{1:\bar{q}},\omega^2_{1:\bar{q}})
\end{aligned}
$$

We follow JP in choosing

$$
f(\theta,\tilde{h}_{1:T}) = f(\theta)\cdot p(\tilde{h}_{1:T}|\lambda_{1:\bar{q}})p(\tilde{\sigma}_{1:T}|\omega^2_{1:\bar{q}})p(\lambda_{1:\bar{q}})p(\omega^2_{1:\bar{q}}),
\tag{29}
$$

where $f(\theta)$ is a truncate multivariate distribution as proposed by Geweke (1999). Hence we approximate the marginal likelihood as:

$$
\hat{p}(y_{1:T}) = \left[\frac{1}{n_{sim}}\sum_{j=1}^{n_{sim}}\frac{f(\theta^j)}{p(y_{1:T}|\tilde{h}^j_{1:T},\tilde{\sigma}^j_{1:T},\theta^j)p(\theta^j)}\right]^{-1}
\tag{30}
$$

where θ^j, $\tilde{h}_{1:T}^j$, and $\tilde{\sigma}_{1:T}^j$ are draws from the posterior distribution, and n_{sim} is the total number of draws. We are aware of the problems with (29), namely that it does not ensure that the random variable

$$\frac{f(\theta, \tilde{h}_{1:T}, \tilde{\sigma}_{1:T}, \lambda_{1:\bar{q}}, \rho_{1:\bar{q}}, \omega_{1:\bar{q}}^2)}{p(y_{1:T}|\tilde{h}_{1:T}, \tilde{\sigma}_{1:T}, \theta)p(\tilde{h}_{1:T}|\lambda_{1:\bar{q}})p(\tilde{\sigma}_{1:T}|\omega_{1:\bar{q}}^2)p(\lambda_{1:\bar{q}})p(\omega_{1:\bar{q}}^2)p(\theta)}$$

has finite variance. Nonetheless, like JP we found that this method delivers very similar results across different chains.

A.2 Parameter Estimates

Does accounting for Student's t shocks and/or stochastic volatility affect the posterior distributions for the DSGE model parameter estimates? JP find that the answer is generally no as far as stochastic volatility is concerned. In our application we broadly reach similar conclusions. Table B.2 shows the posterior means for the parameters estimated in the following four cases: 1) Gaussian shocks and constant volatility (Baseline), 2) Gaussian shocks with stochastic volatility (SV), 3) Student's t shocks (St-t), and 4) Student's t shocks with stochastic volatility (St-t+SV). For reference, we also report the prior mean and standard deviation. We find that the parameter capturing investment adjustment costs (S'') is lower in the baseline specifications relative to the alternatives. Interestingly, JP also find that this parameter is sensitive to changing the specification of the shock distribution, in spite of using a slightly different model and a different sample. Unlike JP, we find that other parameters are also sensitive to the specification of the shock distribution. Namely, we also find that the labor disutility is somewhat more convex when we depart from Gaussianity. We find that the price rigidities parameter (ζ_p) has a higher posterior mean when we account for fat-tails than when we do not (it is about 0.73 in the Gaussian case and 0.85 in the case with both fat tails and stochastic volatility). Additionally, the estimates of the persistency of the shocks are also influenced by the inclusion of stochastic volatility and/or fat tails. Finally, Table B.25 compares the posterior means of the DSGE parameters under the various specifications for the full sample and the sample ending in 2004Q4. The posterior estimates appear to change with the sample for all specifications, hence it is not clear that specifications with SV or TD provide more "robust" parameters estimates with respect to changes in the sample. Admittedly, this issue deserves a more detailed assessment.

A.3 Posterior estimates of the shocks ($\varepsilon_{q,t}$), stochastic volatilities ($\tilde{\sigma}_t$), Student's t scale component ($\tilde{h}_t$), "tamed" shocks (η_t), and Stochastic Volatility Innovation Variance ω_q^2

As discussed in Section 2, our model makes a strong assumption: we assume that changes in volatility are either very persistent or i.i.d. (Student's t). It is worth looking at, and analyzing, the posterior estimates of the Student's t scale components $\tilde{h}_{q,t}^{-1/2}$ and the "tamed" shocks $\eta_{q,t}$ to assess whether these are indeed i.i.d., as they ought to be, or whether there is still residual autocorrelation. Figures B.4 and B.5 in appendix B show precisely these quantities (the $\eta_{q,t}$ are in absolute value) for the SVTD($\underline{\lambda} = 6$) specification. In addition, tables B.13 and B.14 show the posterior mean of autocorrelation of the "raw" shocks ($\varepsilon_{q,t}$) and the "tamed" shocks ($\eta_{q,t}$), respectively, across different specifications. While the autocorrelation of the "raw" shocks ($\varepsilon_{q,t}$) is non-negligible (and often higher for the Student's t case) the autocorrelation of the $\eta_{q,t}$s is always smaller in the SVTD specification, and often substantially so. In addition, the autocorrelation of the $\eta_{q,t}$s for the specification with SV only is always larger than in the SVTD specification for those shocks where the autocorrelation is non-negligible, such as the price markup (λ_f) and the policy (r^m) shocks.

Another important assumption is that both the $\eta_{q,t}$ and the $\tilde{h}_{q,t}$ are uncorrelated across different shocks q. Table B.15 shows the cross-correlation in the "tamed" shocks ($\eta_{q,t}$). The table shows that the cross-correlations are sometimes quite large for the Gaussian and SV case (e.g., up to .47 and .38, respectively, for policy r^m and discount

rate b shocks) but is generally much smaller for the SVTD specification (at most .18). One may be concerned that with Student's t shocks the auto/cross correlations have migrated to the $\tilde{h}^{-1/2}$. Tables B.16 and B.17 suggest that this is not the case: all autocorrelations and cross-correlations are very small, less than .035.

Finally, figure B.6 shows the posterior estimates of the shocks $\varepsilon_{q,t}$ (in absolute value) and the shock volatilties $\sigma_{q,t}$ in the SV and SVTD specifications. The figure shows that the time series of $\sigma_{q,t}$ broadly reflects the time variation in the volatility of the shocks, and how allowing for fat tails affects the estimates of $\sigma_{q,t}$. Table B.3 shows the posterior of the SV innovation variance ω_q^2. One should bear in mind that the effect of SV shocks on $\sigma_{q,t}$ depends on the size of the non-time varying components σ_q, which is different across specifications (see Table B.2). Therefore ω_q^2 is sometimes smaller in the SV than the SVTD specification, but this is often because the corresponding estimate of σ_q is larger, and hence movements in $\sigma_{q,t}$ may well be larger.

A.4   Computational Issues and Convergence

Our results are based on 4 chains, each beginning from a different starting point, with 220 thousand draws each, of which we discard the first 20 thousand draws. The computational cost of using time-variation in volatility is substantial but not overwhelming. In our experiments, we found that, relative to the baseline Gaussian model, the TD, SV and SVTD estimations took roughly 2-4 times as long to sample the same number of draws. The TD, SV, and SVTD estimations all took roughly the same amount of time. The reason for this is that the main computational cost is related to drawing the disturbances on each iteration of the Gibbs sampler (which is not necessary in a Gaussian model) rather than the drawing of the time-varying volatilities or the volatility parameters. For this (expensive) step of drawing the disturbances, we recommend the simulation smoother of Durbin and Koopman (2002), which we have found to be highly efficient relative to alternative methods.

We provide a formal assessment of convergence in Tables B.5 through B.12 of appendix B. We present convergence results for our main specification (SVTD($\bar{\lambda} = 6$)). The same convergence results are available for all the samples and sub-samples and all the different specifications. We use four metrics to assess convergence, aside from plots of the evolution of the MCMC draws in each chain and comparing histograms across chains for each parameter. First, the R statistic of Gelman and Rubin (1992), which compares the variance of each parameter estimate between and within chains and estimates the factor by which these could be reduced by continuing to take draws. This statistic is always larger or equal than one, and a cut-off of 1.01 is often used. Second, the number of effective draws in each chain for each parameter, which corrects for the serial correlation across draws following Geweke et al. (1992). Third, the number of effective draws in total, which combines the previous two corrections applied to the mixed simulations from the four chains (Gelman et al. (2004), page 298). Finally, we show the separated partial means test of Geweke et al. (1992) in which few rejections implies being closer to convergence.

We focus on showing convergence for the objects of interest in this paper: namely the value of the posterior, estimates of the degrees of freedom λ of the Student's t distribution, the variance of SV innovations, and the ratios of pre/post Great Moderation volatility. Overall, we were very satisfied with convergence for our most important specifications. As shown in the tables of Section B.3.2, the R-squared statistic for the posterior and for the parameters governing the SV and TD components all exhibit low R statistics, high numbers of effective observations, and few rejections of the separated partial means test. Convergence for the DSGE parameters θ was also generally quite good, although some parameters have only a few hundred effective draws. We found the convergence properties of our alternative specifications to be satisfactory. All convergence results are available upon request. We found that the convergence speed for our main SVTD specification and the baseline Gaussian specification were similar, as measured by the number of effective draws out of samples of the same size.

We were frequently able to improve the convergence properties of our samples by re-running the estimation. In particular, using the previous run's realized covariance matrix of the θ parameters as the covariance of our proposal density for θ often yielded much better convergence properties.

Table 1: 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

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

Table 2: Posterior of the Student's t Degrees of Freedom

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

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

Table 3: 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

Notes: 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: 1964 (beginning of sample), 1981 (peak of the high volatility period), 1994 (great moderation), 2007 (pre-great recession) and 2011 (end of sample).

Figure 1: Smoothed Shocks under Gaussianity (Absolute Value, Standardized)
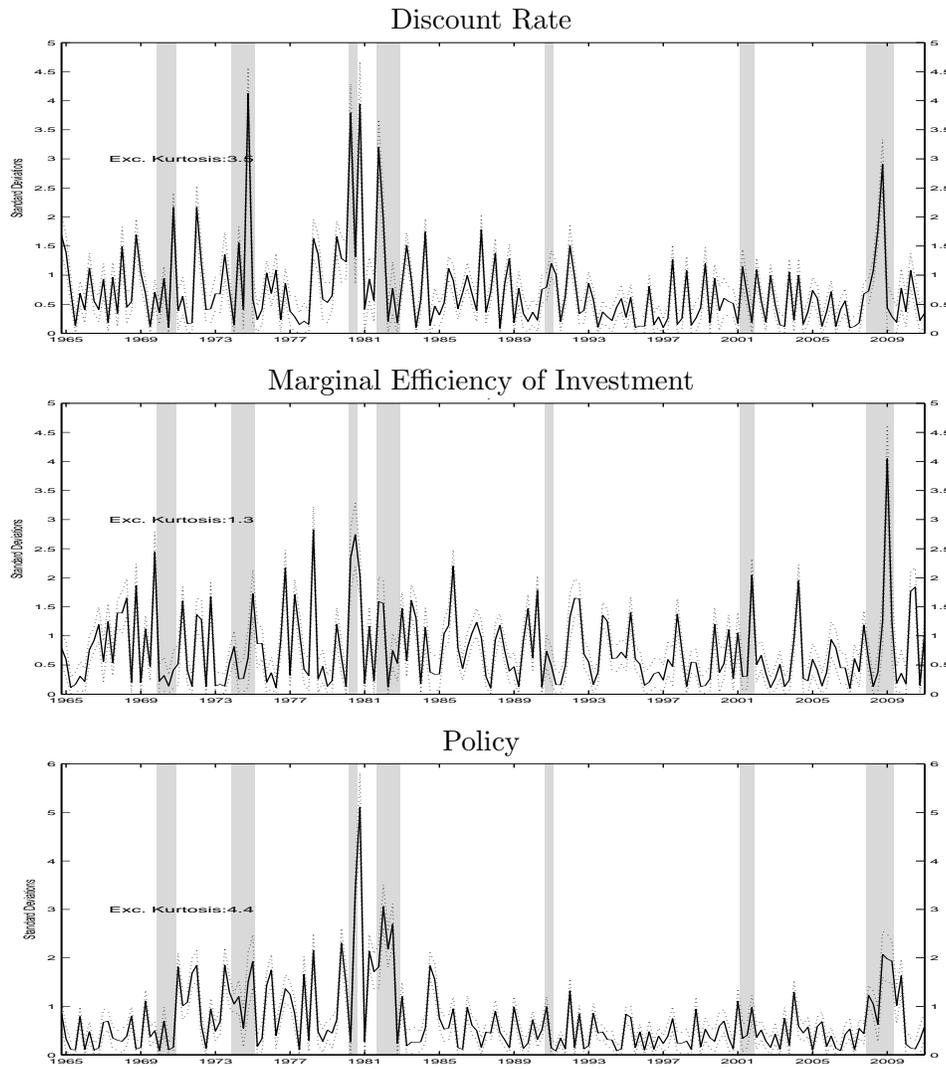
Notes: The solid line is the median, and the dashed lines are the posterior 90% bands. The vertical shaded regions identify NBER recession dates.

Figure 2: Priors on degrees of freedom of Student's t distribution
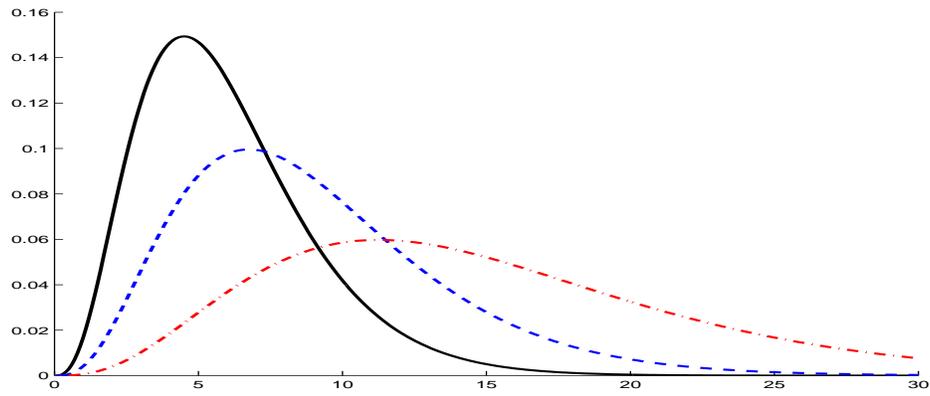
Notes: Prior density for $\underline{\lambda} = 6$ (solid), 9 (dashed), and 15 (dash-and-dotted). All priors have $\underline{\nu} = 4$ degrees of freedom.

Figure 3: 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.

Output Growth

Consumption Growth

Investment Growth
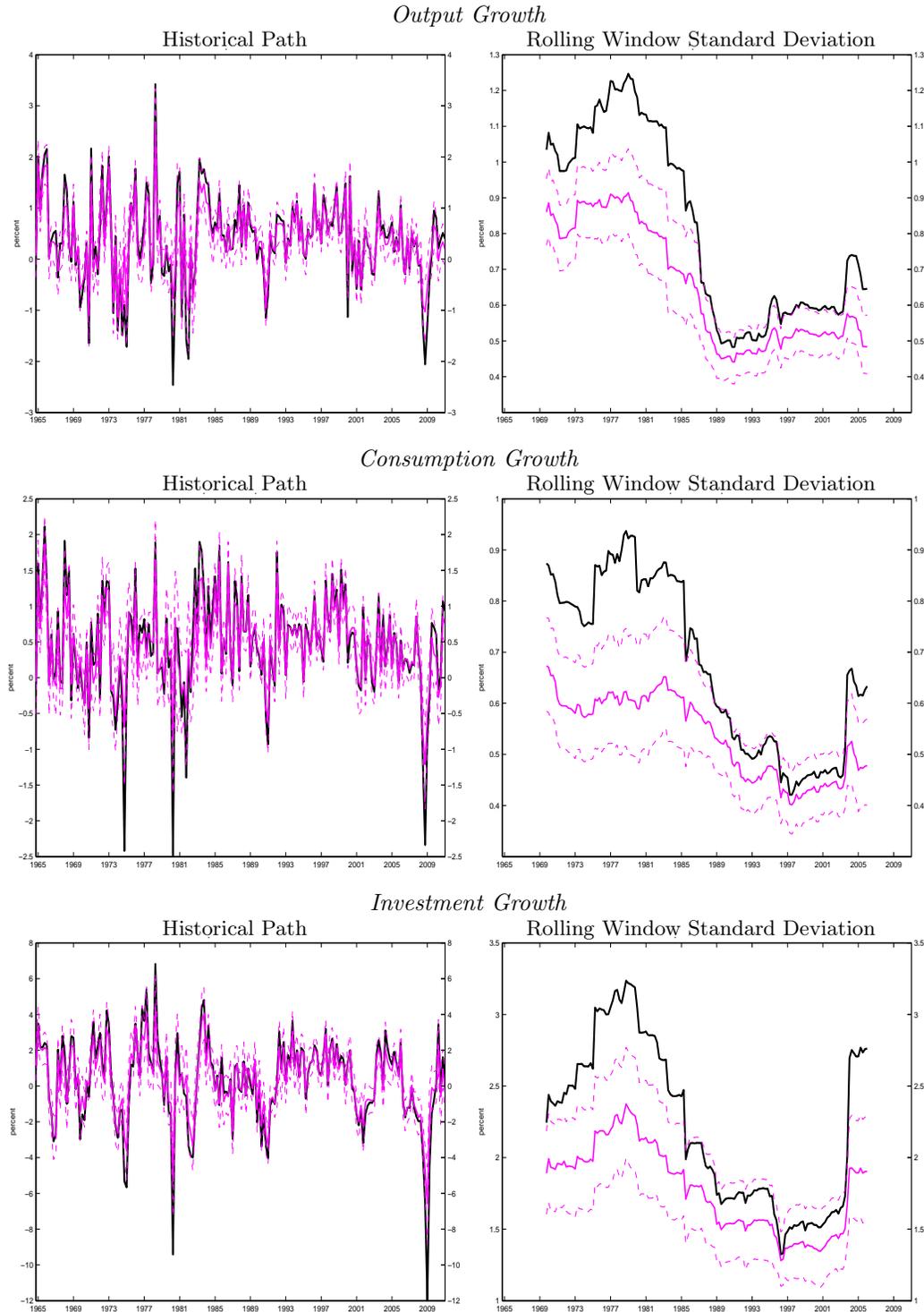
Notes: Black lines are the historical evolution of the variable, and magenta lines are the median counterfactual evolution of the same variable if we shut down the Student's t distributed component of all shocks. The rolling window standard deviation uses 20 quarters before and 20 quarters after a given quarter.

Figure 4: Time-Variation in the unconditional standard deviation of output and consumption; models estimated with and without the Student's t distributed component.

Output Growth *Consumption Growth*

Unconditional variance over time

Ratio of variance in 1981 relative to 1994

Ratio of variance in 2011 relative to 2005
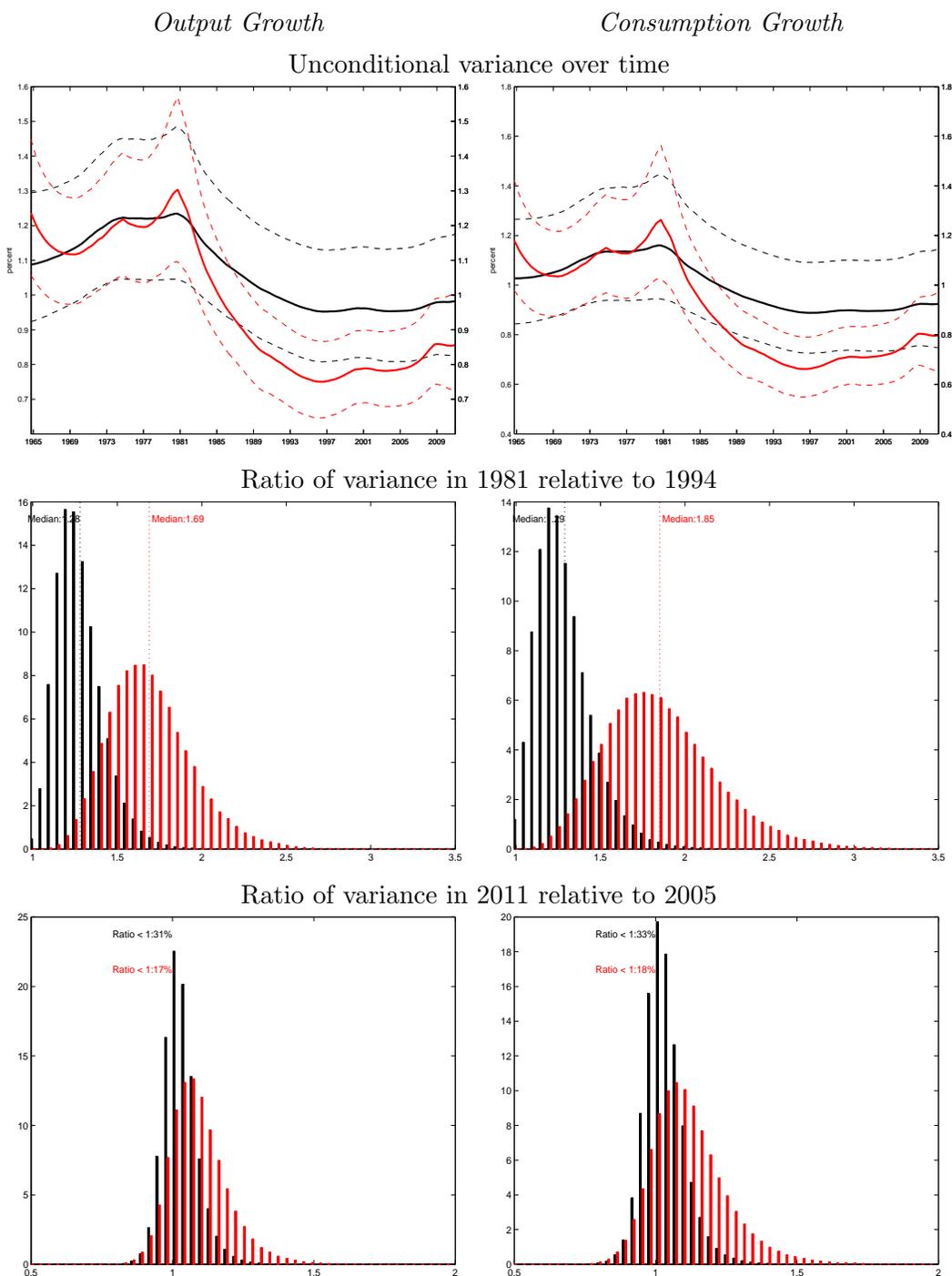
Notes: Black line in the top panel is the unconditional standard deviation in the estimation with both stochastic volatility and Student's t components, while the red line is the unconditional variance in the estimation with stochastic volatility component only. On the middle panel the black bars correspond to the posterior histogram of the ratio of volatility in 1981 over the variance in 1994 for the estimation with both stochastic volatility and Student's t components, while the red bars are for the estimation with with stochastic volatility component only. The lower panel replicates the same analysis as in the middle panel but for the ratio of volatility in 2011 over the variance in 2005.

B   Appendix — Additional Information/Results

B.1   The Smets-Wouters Model

We begin by briefly describing the log-linearized equilibrium conditions of the Smets and Wouters (2007) model. We follow Del Negro and Schorfheide (forthcoming) and detrend the non-stationary model variables by a stochastic rather than a deterministic trend.[30] Let $\tilde{z}_t$ be the linearly detrended log productivity process which follows the autoregressive law of motion

$$\tilde{z}_t = \rho_z \tilde{z}_{t-1} + \sigma_z \varepsilon_{z,t}. \tag{31}$$

We detrend all non stationary variables by $Z_t = e^{\gamma t + \frac{1}{1-\alpha} \tilde{z}_t}$, where γ is the steady state growth rate of the economy. The growth rate of Z_t in deviations from γ, denoted by z_t, follows the process:

$$z_t = \ln(Z_t/Z_{t-1}) - \gamma = \frac{1}{1-\alpha}(\rho_z - 1)\tilde{z}_{t-1} + \frac{1}{1-\alpha}\sigma_z \epsilon_{z,t}. \tag{32}$$

All variables in the following equations are expressed in log deviations from their non-stochastic steady state. Steady state values are denoted by $*$-subscripts and steady state formulas are provided in the technical appendix of Del Negro and Schorfheide (forthcoming).[31] The consumption Euler equation is given by:

$$
\begin{aligned}
c_t = &-\frac{(1 - he^{-\gamma})}{\sigma_c(1 + he^{-\gamma})}\left(R_t - I\!\!E_t[\pi_{t+1}] + b_t\right) + \frac{he^{-\gamma}}{(1 + he^{-\gamma})}\left(c_{t-1} - z_t\right) \\
&+ \frac{1}{(1 + he^{-\gamma})}I\!\!E_t\left[c_{t+1} + z_{t+1}\right] + \frac{(\sigma_c - 1)}{\sigma_c(1 + he^{-\gamma})}\frac{w_* L_*}{c_*}\left(L_t - I\!\!E_t[L_{t+1}]\right),
\end{aligned} \tag{33}
$$

where c_t is consumption, L_t is labor supply, R_t is the nominal interest rate, and π_t is inflation. The exogenous process b_t drives a wedge between the intertemporal ratio of the marginal utility of consumption and the riskless real return $R_t - I\!\!E_t[\pi_{t+1}]$, and follows an AR(1) process with parameters ρ_b and σ_b. The parameters σ_c and h capture the relative degree of risk aversion and the degree of habit persistence in the utility function, respectively. The following condition expresses the relationship between the value of capital in

---

[30]This approach makes it possible to express almost all equilibrium conditions in a way that encompasses both the trend-stationary total factor productivity process in Smets and Wouters (2007), as well as the case where technology follows a unit root process.

[31]Available at `http://economics.sas.upenn.edu/~schorf/research.htm`.

terms of consumption q_t^k and the level of investment i_t measured in terms of consumption goods:

$$q_t^k = S'' e^{2\gamma}(1 + \beta e^{(1-\sigma_c)\gamma})\Big(i_t - \frac{1}{1 + \beta e^{(1-\sigma_c)\gamma}}\,(i_{t-1} - z_t)$$
$$- \frac{\beta e^{(1-\sigma_c)\gamma}}{1 + \beta e^{(1-\sigma_c)\gamma}}\,\mathbb{E}_t\,[i_{t+1} + z_{t+1}] - \mu_t\Big), \quad (34)$$

which is affected by both investment adjustment cost (S'' is the second derivative of the adjustment cost function) and by μ_t, an exogenous process called the "marginal efficiency of investment" that affects the rate of transformation between consumption and installed capital (see Greenwood et al. (1998)). The exogenous process μ_t follows an AR(1) process with parameters ρ_μ and σ_μ. The parameter β captures the intertemporal discount rate in the utility function of the households.

The capital stock, $\bar{k}_t$, evolves as

$$\bar{k}_t = \left(1 - \frac{i_*}{\bar{k}_*}\right)(\bar{k}_{t-1} - z_t) + \frac{i_*}{\bar{k}_*}i_t + \frac{i_*}{\bar{k}_*}S'' e^{2\gamma}(1 + \beta e^{(1-\sigma_c)\gamma})\mu_t, \quad (35)$$

where $i_*/\bar{k}_*$ is the steady state ratio of investment to capital. The arbitrage condition between the return to capital and the riskless rate is:

$$\frac{r_*^k}{r_*^k + (1-\delta)}\mathbb{E}_t[r_{t+1}^k] + \frac{1-\delta}{r_*^k + (1-\delta)}\mathbb{E}_t[q_{t+1}^k] - q_t^k = R_t + b_t - \mathbb{E}_t[\pi_{t+1}], \quad (36)$$

where r_t^k is the rental rate of capital, r_*^k its steady state value, and δ the depreciation rate. Given that capital is subject to variable capacity utilization u_t, the relationship between $\bar{k}_t$ and the amount of capital effectively rented out to firms k_t is

$$k_t \;=\; u_t - z_t + \bar{k}_{t-1}. \quad (37)$$

The optimality condition determining the rate of utilization is given by

$$\frac{1-\psi}{\psi}r_t^k = u_t, \quad (38)$$

where ψ captures the utilization costs in terms of foregone consumption. Real marginal costs for firms are given by

$$mc_t \;=\; w_t + \alpha L_t - \alpha k_t, \quad (39)$$

where α is the income share of capital (after paying markups and fixed costs) in the production function. From the optimality conditions of goods producers it follows that all firms have the same capital-labor ratio:

$$k_t \;=\; w_t - r_t^k + L_t. \quad (40)$$

The production function is:

$$y_t = \Phi_p\left(\alpha k_t + (1-\alpha)L_t\right) + \mathcal{I}\{\rho_z < 1\}(\Phi_p - 1)\frac{1}{1-\alpha}\tilde{z}_t, \tag{41}$$

under trend stationarity. The last term $(\Phi_p - 1)\frac{1}{1-\alpha}\tilde{z}_t$ drops out if technology has a stochastic trend, because in this case one has to assume that the fixed costs are proportional to the trend. Similarly, the resource constraint is:

$$y_t = g_t + \frac{c_*}{y_*}c_t + \frac{i_*}{y_*}i_t + \frac{r_*^k k_*}{y_*}u_t - \mathcal{I}\{\rho_z < 1\}\frac{1}{1-\alpha}\tilde{z}_t, \tag{42}$$

where again the term $-\frac{1}{1-\alpha}\tilde{z}_t$ disappears if technology follows a unit root process. Government spending g_t is assumed to follow the exogenous process:

$$g_t = \rho_g g_{t-1} + \sigma_g \varepsilon_{g,t} + \eta_{gz}\sigma_z \varepsilon_{z,t}.$$

Finally, the price and wage Phillips curves are, respectively:

$$\pi_t = \frac{(1 - \zeta_p\beta e^{(1-\sigma_c)\gamma})(1-\zeta_p)}{(1 + \iota_p\beta e^{(1-\sigma_c)\gamma})\zeta_p((\Phi_p-1)\epsilon_p + 1)}mc_t$$
$$+ \frac{\iota_p}{1 + \iota_p\beta e^{(1-\sigma_c)\gamma}}\pi_{t-1} + \frac{\beta e^{(1-\sigma_c)\gamma}}{1 + \iota_p\beta e^{(1-\sigma_c)\gamma}}\mathbb{E}_t[\pi_{t+1}] + \lambda_{f,t}, \tag{43}$$

and

$$w_t = \frac{(1 - \zeta_w\beta e^{(1-\sigma_c)\gamma})(1-\zeta_w)}{(1 + \beta e^{(1-\sigma_c)\gamma})\zeta_w((\lambda_w-1)\epsilon_w + 1)}\left(w_t^h - w_t\right)$$
$$- \frac{1 + \iota_w\beta e^{(1-\sigma_c)\gamma}}{1 + \beta e^{(1-\sigma_c)\gamma}}\pi_t + \frac{1}{1 + \beta e^{(1-\sigma_c)\gamma}}\left(w_{t-1} - z_t - \iota_w\pi_{t-1}\right)$$
$$+ \frac{\beta e^{(1-\sigma_c)\gamma}}{1 + \beta e^{(1-\sigma_c)\gamma}}\mathbb{E}_t\left[w_{t+1} + z_{t+1} + \pi_{t+1}\right] + \lambda_{w,t}, \tag{44}$$

where ζ_p, ι_p, and ϵ_p are the Calvo parameter, the degree of indexation, and the curvature parameters in the Kimball aggregator for prices, and ζ_w, ι_w, and ϵ_w are the corresponding parameters for wages. w_t^h measures the household's marginal rate of substitution between consumption and labor, and is given by:

$$w_t^h = \frac{1}{1 - he^{-\gamma}}\left(c_t - he^{-\gamma}c_{t-1} + he^{-\gamma}z_t\right) + \nu_l L_t, \tag{45}$$

where ν_l characterizes the curvature of the disutility of labor (and would equal the inverse of the Frisch elasticity in absence of wage rigidities). The mark-ups $\lambda_{f,t}$ and $\lambda_{w,t}$ follow exogenous ARMA(1,1) processes

$$\lambda_{f,t} = \rho_{\lambda_f}\lambda_{f,t-1} + \sigma_{\lambda_f}\varepsilon_{\lambda_f,t} + \eta_{\lambda_f}\sigma_{\lambda_f}\varepsilon_{\lambda_f,t-1}, \text{ and}$$

$$\lambda_{w,t} = \rho_{\lambda_w}\lambda_{w,t-1} + \sigma_{\lambda_w}\varepsilon_{\lambda_w,t} + \eta_{\lambda_w}\sigma_{\lambda_w}\varepsilon_{\lambda_w,t-1},$$

respectively. Finally, the monetary authority follows a generalized feedback rule:

$$
\begin{aligned}
R_t &= \rho_R R_{t-1} + (1 - \rho_R)\left(\psi_1\pi_t + \psi_2(y_t - y_t^f)\right) \\
&\quad + \psi_3\left((y_t - y_t^f) - (y_{t-1} - y_{t-1}^f)\right) + r_t^m,
\end{aligned}
\tag{46}
$$

where the flexible price/wage output y_t^f is obtained from solving the version of the model without nominal rigidities (that is, Equations (33) through (42) and (45)), and the residual r_t^m follows an AR(1) process with parameters ρ_{r^m} and σ_{r^m}. We use the method in Sims (2002) to solve the log-linear approximation of the DSGE model.

The measurement equations (equation) for real output, consumption, investment, and real wage growth, hours, inflation, and interest rates are given by:

$$
\begin{aligned}
\textit{Output growth} &= \gamma + 100\left(y_t - y_{t-1} + z_t\right) \\
\textit{Consumption growth} &= \gamma + 100\left(c_t - c_{t-1} + z_t\right) \\
\textit{Investment growth} &= \gamma + 100\left(i_t - i_{t-1} + z_t\right) \\
\textit{Real Wage growth} &= \gamma + 100\left(w_t - w_{t-1} + z_t\right), \\
\textit{Hours} &= \bar{l} + 100 l_t \\
\textit{Inflation} &= \pi_* + 100\pi_t \\
\textit{FFR} &= R_* + 100 R_t
\end{aligned}
\tag{47}
$$

where all variables are measured in percent, and the parameters π_* and R_* measure the steady state level of net inflation and short term nominal interest rates, respectively and where $\bar{l}$ captures the mean of hours (this variable is measured as an index).

B.2 Data

The data set is obtained from Haver Analytics (Haver mnemonics are in italics). We compile observations for the variables that appear in the measurement equation (47). Real GDP (GDPC), the GDP price deflator (GDPDEF), nominal personal consumption expenditures (PCEC), and nominal fixed private investment (FPI) are constructed at a quarterly frequency by the Bureau of Economic Analysis (BEA), and are included in the National Income and Product Accounts (NIPA).

Average weekly hours of production and non-supervisory employees for total private industries (PRS85006023), civilian employment (CE16OV), and civilian noninstitutional

population (LNSINDEX) are produced by the Bureau of Labor Statistics (BLS) at the monthly frequency. The first of these series is obtained from the Establishment Survey, and the remaining from the Household Survey. Both surveys are released in the BLS Employment Situation Summary (ESS). Since our models are estimated on quarterly data, we take averages of the monthly data. Compensation per hour for the non-farm business sector (PRS85006103) is obtained from the Labor Productivity and Costs (LPC) release, and produced by the BLS at the quarterly frequency. Last, the federal funds rate is obtained from the Federal Reserve Board's H.15 release at the business day frequency, and is not revised. We take quarterly averages of the annualized daily data.

All data are transformed following Smets and Wouters (2007). Specifically:

$$
\begin{aligned}
Output\ growth &= LN((GDPC)/LNSINDEX) * 100 \\
Consumption\ growth &= LN((PCEC/GDPDEF)/LNSINDEX) * 100 \\
Investment\ growth &= LN((FPI/GDPDEF)/LNSINDEX) * 100 \\
Real\ Wage\ growth &= LN(PRS85006103/GDPDEF) * 100 \\
Hours &= LN((PRS85006023 * CE16OV/100)/LNSINDEX) * 100 \\
Inflation &= LN(GDPDEF/GDPDEF(-1)) * 100 \\
FFR &= FEDERAL\ FUNDS\ RATE/4
\end{aligned}
$$

B.3   Additional Results for Baseline Estimation

B.3.1   Parameters and Variance Decomposition

Table B.1: Priors for the Medium-Scale Model

	Density	Mean	St. Dev.		Density	Mean	St. Dev.
Policy Parameters							
ψ_1	Normal	1.50	0.25	ρ_R	Beta	0.75	0.10
ψ_2	Normal	0.12	0.05	ρ_{r^m}	Beta	0.50	0.20
ψ_3	Normal	0.12	0.05	σ_{r^m}	InvG	0.10	2.00
Nominal Rigidities Parameters							
ζ_p	Beta	0.50	0.10	ζ_w	Beta	0.50	0.10
Other "Endogenous Propagation and Steady State" Parameters							
α	Normal	0.30	0.05	π^*	Gamma	0.62	0.10
Φ	Normal	1.25	0.12	γ	Normal	0.40	0.10
h	Beta	0.70	0.10	S''	Normal	4.00	1.50
ν_l	Normal	2.00	0.75	σ_c	Normal	1.50	0.37
ι_p	Beta	0.50	0.15	ι_w	Beta	0.50	0.15
r_*	Gamma	0.25	0.10	ψ	Beta	0.50	0.15
ρs, σs, and ηs							
ρ_z	Beta	0.50	0.20	σ_z	InvG	0.10	2.00
ρ_b	Beta	0.50	0.20	σ_b	InvG	0.10	2.00
ρ_{λ_f}	Beta	0.50	0.20	σ_{λ_f}	InvG	0.10	2.00
ρ_{λ_w}	Beta	0.50	0.20	σ_{λ_w}	InvG	0.10	2.00
ρ_μ	Beta	0.50	0.20	σ_μ	InvG	0.10	2.00
ρ_g	Beta	0.50	0.20	σ_g	InvG	0.10	2.00
η_{λ_f}	Beta	0.50	0.20	η_{λ_w}	Beta	0.50	0.20
η_{gz}	Beta	0.50	0.20				

Notes: Note that $\beta = (1/(1 + r_*/100))$. The following parameters are fixed in Smets and Wouters (2007): $\delta = 0.025$, $g_* = 0.18$, $\lambda_w = 1.50$, $\varepsilon_w = 10.0$, and $\varepsilon_p = 10$. The columns "Mean" and "St. Dev." list the means and the standard deviations for Beta, Gamma, and Normal distributions, and the values s and ν for the Inverse Gamma (InvG) distribution, where $p_{\mathcal{IG}}(\sigma|\nu, s) \propto \sigma^{-\nu-1} e^{-\nu s^2/2\sigma^2}$. The effective prior is truncated at the boundary of the determinacy region. The prior for $\bar{l}$ is $\mathcal{N}(-45, 5^2)$.

Table B.2: Posterior Means of the DSGE Model Parameters

	Prior Mean	Prior SD	Baseline	SV	St-t	St-t+SV
α	0.300	0.050	0.150	0.134	0.150	0.135
ζ_p	0.500	0.100	0.734	0.780	0.808	0.846
ι_p	0.500	0.150	0.315	0.344	0.383	0.286
Φ	1.250	0.120	1.580	1.518	1.575	1.551
S''	4.000	1.500	4.686	5.013	5.070	5.651
h	0.700	0.100	0.611	0.609	0.582	0.571
ψ	0.500	0.150	0.714	0.734	0.670	0.666
ν_l	2.000	0.750	2.088	2.212	2.300	2.476
ζ_w	0.500	0.100	0.803	0.826	0.830	0.843
ι_w	0.500	0.150	0.541	0.547	0.495	0.511
β	0.250	0.100	0.206	0.184	0.202	0.175
ψ_1	1.500	0.250	1.953	1.866	1.820	1.884
ψ_2	0.120	0.050	0.083	0.073	0.115	0.116
ψ_3	0.120	0.050	0.245	0.217	0.213	0.184
π^*	0.620	0.100	0.683	0.719	0.706	0.808
σ_c	1.500	0.370	1.236	1.109	1.248	1.274
ρ	0.750	0.100	0.835	0.854	0.875	0.875
γ	0.400	0.100	0.306	0.321	0.356	0.389
$\bar{L}$	-45.00	5.000	-44.17	-46.67	-43.38	-44.73
ρ_g	0.500	0.200	0.977	0.977	0.982	0.988
ρ_b	0.500	0.200	0.758	0.845	0.844	0.852
ρ_μ	0.500	0.200	0.748	0.753	0.791	0.806
ρ_z	0.500	0.200	0.994	0.991	0.987	0.981
ρ_{λ_f}	0.500	0.200	0.791	0.797	0.811	0.830
ρ_{λ_w}	0.500	0.200	0.981	0.952	0.962	0.923
ρ_{rm}	0.500	0.200	0.154	0.219	0.219	0.227
σ_g	0.100	2.000	2.892	3.169	2.387	2.665
σ_b	0.100	2.000	0.125	0.122	0.072	0.100
σ_μ	0.100	2.000	0.430	0.454	0.325	0.300
σ_z	0.100	2.000	0.493	0.869	0.362	0.473
σ_{λ_f}	0.100	2.000	0.164	0.191	0.163	0.127
σ_{λ_w}	0.100	2.000	0.281	0.203	0.213	0.151
σ_{rm}	0.100	2.000	0.228	0.243	0.133	0.095
η_{gz}	0.500	0.200	0.787	0.775	0.786	0.765
η_{λ_f}	0.500	0.200	0.670	0.749	0.815	0.734
η_{λ_w}	0.500	0.200	0.948	0.914	0.924	0.865

Notes: We use a prior mean of 6 degrees of freedom for the Student's t distributed component. The stochastic volatility component assumes a prior mean for the size of the shocks to volatility of $(0.01)^2$.

Figure B.1: Prior on Innovation Variance (ω^2) for Stochastic Volatility
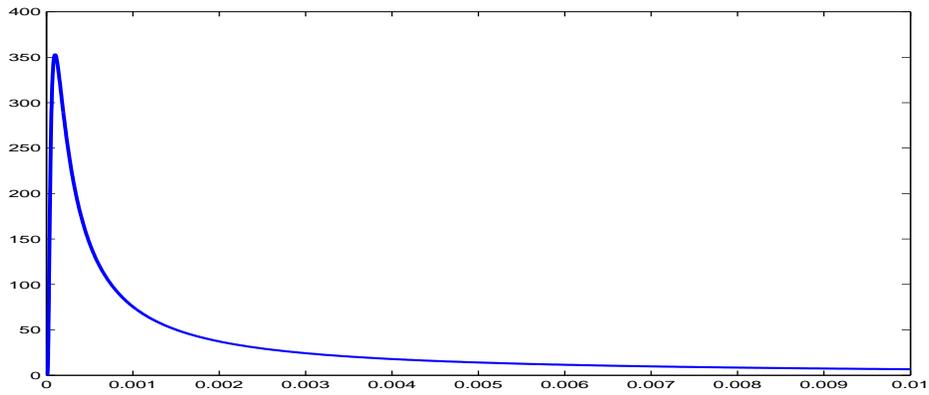
Table B.3: Posterior of the Stochastic Volatility Innovation Variance

	Without Student's t	With Student's t
g	0.001	0.007
	(0.000,0.002)	(0.000,0.015)
b	0.003	0.005
	(0.000,0.006)	(0.000,0.0012)
μ	0.000	0.002
	(0.000,0.001)	(0.000,0.005)
z	0.002	0.003
	(0.000,0.004)	(0.000,0.007)
λ_f	0.001	0.008
	(0.000,0.003)	(0.001,0.016)
λ_w	0.001	0.002
	(0.000,0.002)	(0.002,0.005)
r^m	0.006	0.022
	(0.000,0.011)	(0.005,0.039)

Notes: Numbers shown for the posterior mean and the 90% intervals of the stochastic volatility innovation variance.

Table B.4: Variance Decomposition for Real GDP Growth

Gaussian Shocks

	g	b	μ	z	λ_f	λ_w	r^m
Without Stochastic Volatility							
	0.225	0.354	0.095	0.140	0.041	0.036	0.109
With Stochastic Volatility							
σ_{1964}	0.162	0.359	0.062	0.252	0.031	0.012	0. 123
σ_{1981}	0.160	0.438	0.052	0.075	0.027	0.016	0. 233
σ_{1994}	0.240	0.306	0.120	0.136	0.052	0.057	0. 089
σ_{2007}	0.189	0.312	0.106	0.156	0.062	0.068	0. 106
σ_{2011}	0.175	0.318	0.100	0.163	0.062	0.061	0. 121

Student's t Shocks

	g	b	μ	z	λ_f	λ_w	r^m
Without Stochastic Volatility							
	0.184	0.352	0.083	0.124	0.036	0.020	0.199
With Stochastic Volatility							
σ_{1964}	0.170	0.417	0.082	0.207	0.034	0.016	0. 073
σ_{1981}	0.205	0.350	0.075	0.091	0.032	0.021	0. 225
σ_{1994}	0.178	0.394	0.109	0.142	0.038	0.045	0. 094
σ_{2007}	0.163	0.361	0.104	0.155	0.044	0.054	0. 118
σ_{2011}	0.168	0.359	0.105	0.138	0.048	0.051	0. 132

Notes: The tables show the relative contribution of the different shocks to the unconditional variance of real GDP. In the case with stochastic volatility we evaluate this contribution at different points in time: 1964 (beginning of sample), 1981 (peak of the high volatility period), 1994 (great moderation), 2007 (pre-great recession) and 2011 (end of sample).

B.3.2   Convergence Tables

Table B.5: R Statistic and Number of Effective Draws: Posterior

	R	mn^{eff}	$n^{eff}(1)$	$n^{eff}(2)$	$n^{eff}(3)$	$n^{eff}(4)$
post	1.0001	16393	4871	4897	4071	4400

Table B.6: Separated Partial Means Test: Posterior

	$SPM_4(1)$	$SPM_4(2)$	$SPM_4(3)$	$SPM_4(4)$
post	3.76	1.61	2.02	2.37

The null hypothesis of the SPM test is that the mean in two separate subsamples is the same. * indicates pvalue less than 5%. ** indicates pvalue less than 1%.

Table B.7: R Statistic and Number of Effective Draws: Student's t Degrees of Freedom (λ)

	R	mn^{eff}	$n^{eff}(1)$	$n^{eff}(2)$	$n^{eff}(3)$	$n^{eff}(4)$
g	1.0001	28090	5781	7019	5894	7309
b	1.0001	16072	4118	2268	3113	1490
μ	1.0001	32383	4703	4658	5683	4388
z	1.0002	11328	1981	2657	2859	2550
λ_f	1.0001	19064	6741	6382	4886	4382
λ_w	1.0001	16329	9489	8524	7566	3714
r^m	1.0002	7946	3956	3716	4365	3000

Table B.8: Separated Partial Means Test: Student's t Degrees of Freedom (λ)

	$SPM_4(1)$	$SPM_4(2)$	$SPM_4(3)$	$SPM_4(4)$
g	3.12	3.99	2.59	2.99
b	2.01	1.50	0.72	6.32
μ	1.55	3.48	1.07	0.48
z	7.75	4.58	0.50	8.56 *
λ_f	0.06	0.21	6.16	2.31
λ_w	0.65	0.82	3.89	3.86
r^m	3.98	2.09	1.44	2.72

The null hypothesis of the SPM test is that the mean in two separate subsamples is the same. * indicates pvalue less than 5%. ** indicates pvalue less than 1%.

Table B.9: R Statistic and Number of Effective Draws: SV Innovation Variance

	R	mn^{eff}	$n^{eff}(1)$	$n^{eff}(2)$	$n^{eff}(3)$	$n^{eff}(4)$
g	1.0014	1451	1038	1739	1428	1159
b	1.0001	32466	4481	7342	6195	4399
μ	1.0003	6324	4139	5848	6985	6569
z	1.0001	14141	2118	3256	2908	2533
λ_f	1.0013	1549	1936	1475	1780	769
λ_w	1.0018	1133	1794	2393	2429	348
r^m	1.0003	6221	2179	2089	3343	3020

Table B.10: Separated Partial Means Test: SV Innovation Variance

	$SPM_4(1)$	$SPM_4(2)$	$SPM_4(3)$	$SPM_4(4)$
g	2.75	4.77	5.37	2.03
b	0.96	1.98	4.17	1.63
μ	3.94	1.96	2.37	3.14
z	3.58	1.80	0.27	10.60 *
λ_f	0.23	1.95	1.00	4.40
λ_w	1.47	4.15	3.22	4.72
r^m	3.65	4.02	1.58	2.09

The null hypothesis of the SPM test is that the mean in two separate subsamples is the same. * indicates pvalue less than 5%. ** indicates pvalue less than 1%.

Table B.11: R Statistic and Number of Effective Draws: Ratio of 1981 to 1994 Variance

	R	mn^{eff}	$n^{eff}(1)$	$n^{eff}(2)$	$n^{eff}(3)$	$n^{eff}(4)$
Output Growth	1.0002	9654	2627	6393	3794	2444
Per Capita Consumption Growth	1.0003	6734	3421	1387	5583	4921
Per Capita Investment Growth	1.0000	57827	2318	5256	9423	8376

Table B.12: Separated Partial Means Test: Ratio of 1981 to 1994 Variance

	$SPM_4(1)$	$SPM_4(2)$	$SPM_4(3)$	$SPM_4(4)$
Output Growth	2.88	1.39	3.69	0.22
Per Capita Consumption Growth	2.68	4.76	4.43	1.31
Per Capita Investment Growth	1.93	5.78	0.97	2.70

The null hypothesis of the SPM test is that the mean in two separate subsamples is the same. * indicates pvalue less than 5%. ** indicates pvalue less than 1%.

B.3.3   Robustness to the choice of $\underline{\lambda}$

Figure B.2: Results for $\underline{\lambda} = 9$

Output Growth: Counterfactual evolution with Student's t component turned off

Historical Path   Rolling Window Standard Deviation

Output Growth: SV versus SV + Student's t

Unconditional variance over time   Ratio of variance in 1981 relative to 1994

Notes: Top panels: Black lines are the historical evolution of the variable, and magenta lines are the median counterfactual evolution of the same variable if we shut down the Student-t distributed component of all shocks. The rolling window standard deviation uses 20 quarters before and 20 quarters after a given quarter. Southwest panel: Black line is the unconditional standard deviation in the estimation with both stochastic volatility and Student-t components, while the red line is the unconditional variance in the estimation with stochastic volatility component only. Southeast panel: Black bars correspond to the posterior histogram of the ratio of volatility in 1981 over the variance in 1994 for the estimation with both stochastic volatility and Student-t components, while the red bars are for the estimation with with stochastic volatility component only.

Figure B.3: Results for $\underline{\lambda} = 15$

Output Growth: Counterfactual evolution with Student's t component turned off

Historical Path   Rolling Window Standard Deviation

Output Growth: SV versus SV + Student's t

Unconditional variance over time   Ratio of variance in 1981 relative to 1994
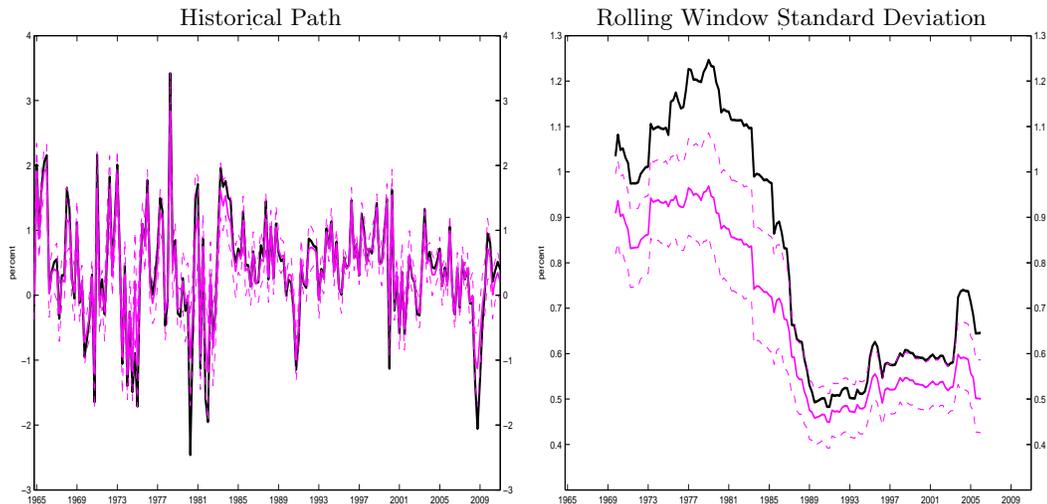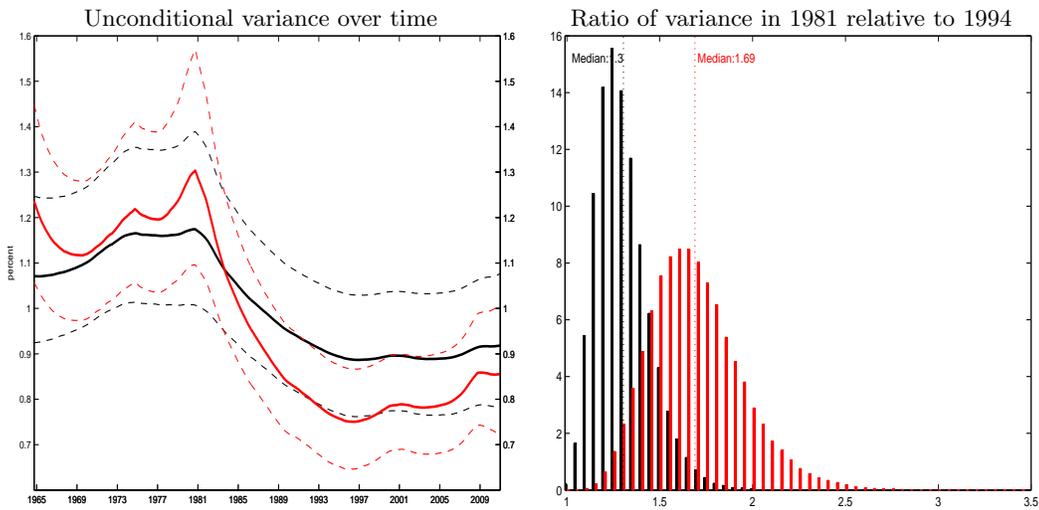
Notes: Top panels: Black lines are the historical evolution of the variable, and magenta lines are the median counterfactual evolution of the same variable if we shut down the Student-t distributed component of all shocks. The rolling window standard deviation uses 20 quarters before and 20 quarters after a given quarter. Southwest panel: Black line is the unconditional standard deviation in the estimation with both stochastic volatility and Student-t components, while the red line is the unconditional variance in the estimation with stochastic volatility component only. Southeast panel: Black bars correspond to the posterior histogram of the ratio of volatility in 1981 over the variance in 1994 for the estimation with both stochastic volatility and Student-t components, while the red bars are for the estimation with with stochastic volatility component only.

B.3.4 Posterior estimates of the shocks ($\varepsilon_{q,t}$), stochastic volatilities ($\tilde{\sigma}_t$), Student's t scale component ($\tilde{h}_t$), and "tamed" shocks (η_t)

Figure B.4: $\eta_{q,t}$

Stochastic volatility   **Stochastic volatility + Student's t**
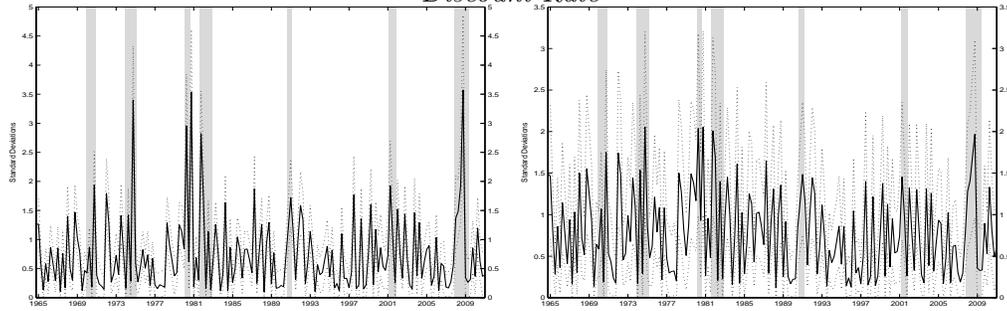
Discount Rate

Monetary Policy

MEI
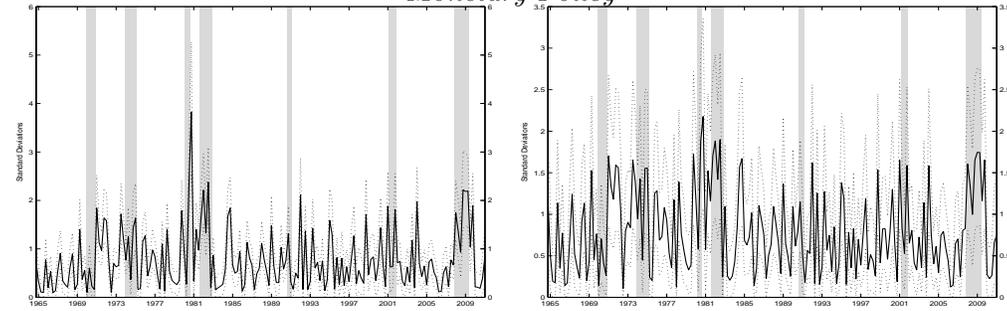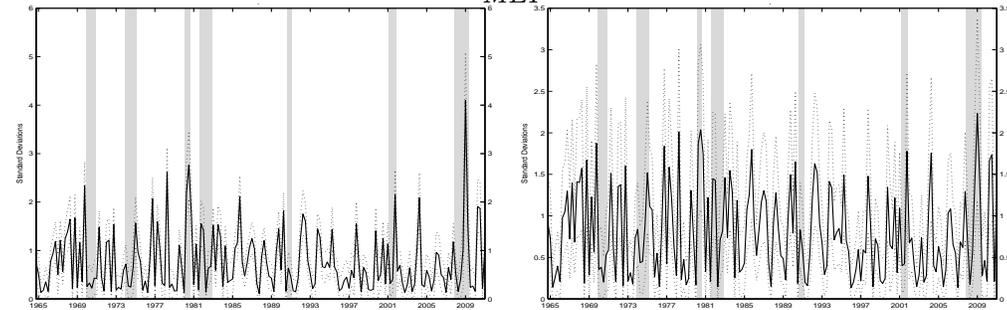
Gvmt Spending
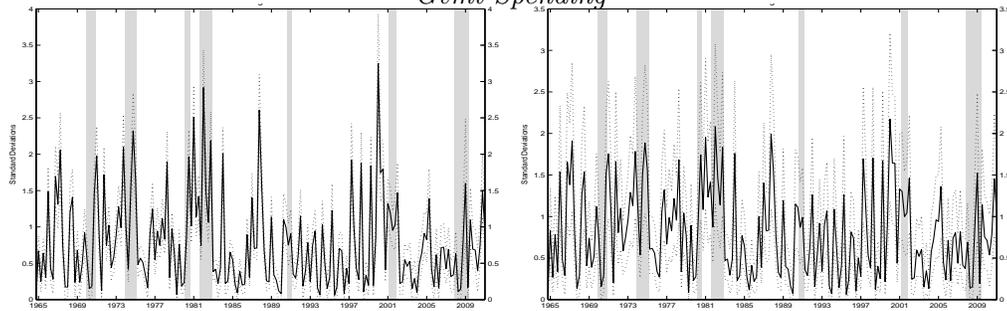
Notes: Estimation with Student's t distribution with $\underline{\lambda} = 6$. The solid line is the median, and the dashed lines are the posterior 90% bands. Black line is the absolute value of the shock, and the red line is the stochastic volatility component.

Figure B.4 — Continued
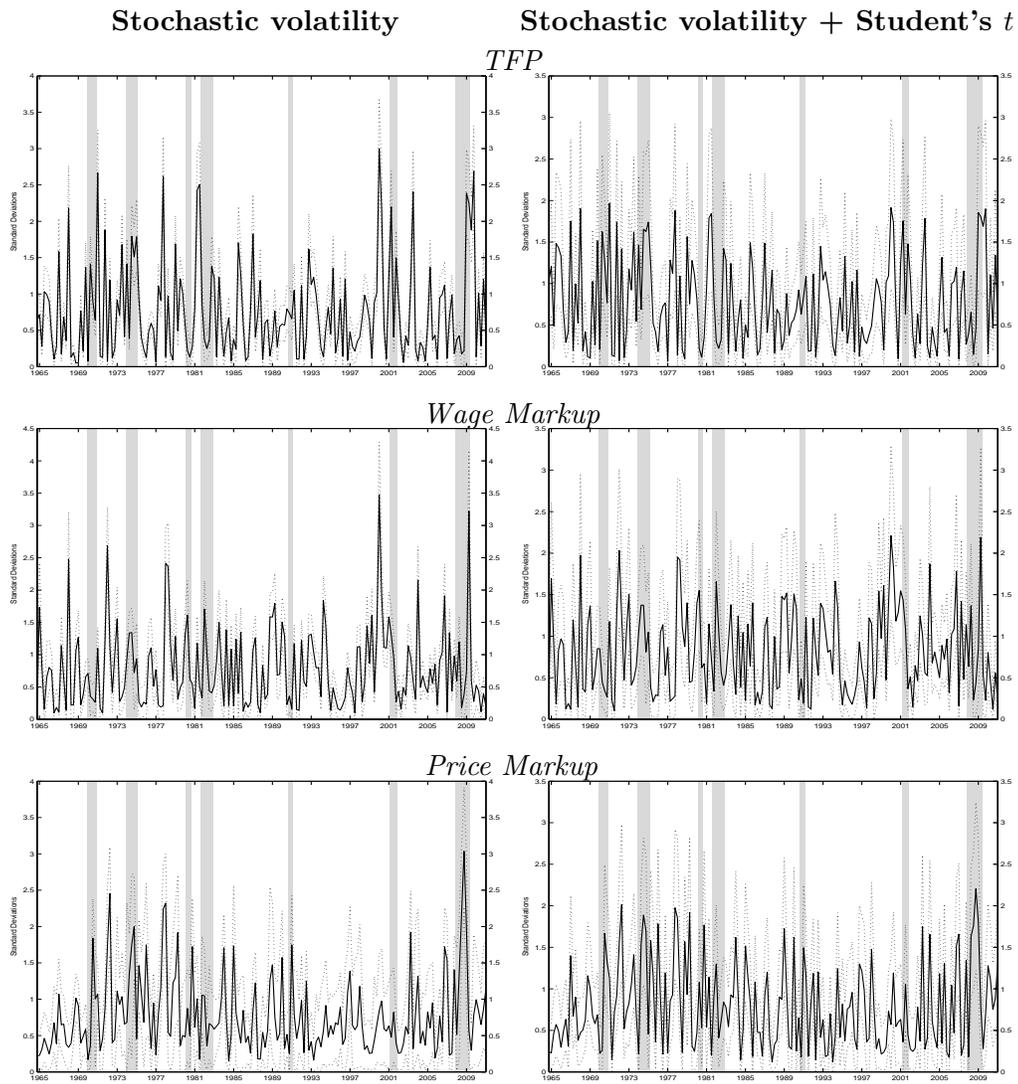
Stochastic volatility   **Stochastic volatility + Student's t**

TFP

Wage Markup

Price Markup

Notes: Estimation with Student's t distribution with $\underline{\lambda} = 6$. The solid line is the median, and the dashed lines are the posterior 90% bands. Black line is the absolute value of the shock, and the red line is the stochastic volatility component.

Figure B.5: $\tilde{h}_{q,t}^{-1/2}$
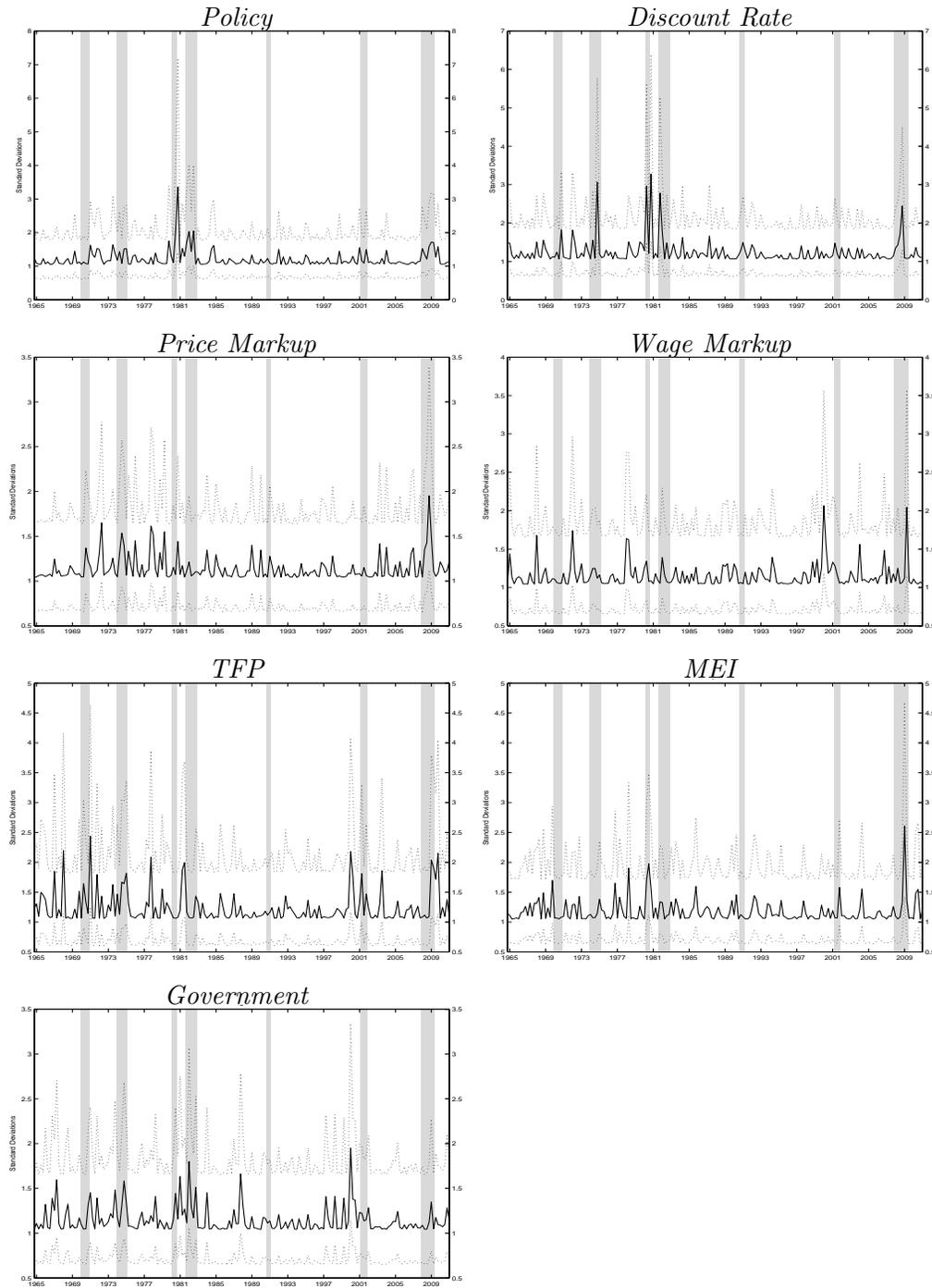
Notes: Estimation with Student's t distribution with $\underline{\lambda} = 6$. The solid line is the median, and the dashed lines are the posterior 90% bands. Black line is the absolute value of the shock, and the red line is the stochastic volatility component.

Table B.13: Autocorrelation of Squared "Raw" Shocks ($\varepsilon_{q,t}$)

Spec.	g	b	μ	z	λ_f	λ_w	r^m
Gaussian	0.191	0.054	0.095	0.101	0.282	0.158	0.350
Student t	0.216	0.004	0.131	0.106	0.306	0.199	0.263
SV	0.221	0.066	0.135	0.534	0.297	0.223	0.322
SV + t	0.303	0.061	0.222	0.256	0.382	0.213	0.290

Table B.14: Autocorrelation of Squared "Tamed" Shocks ($\eta_{q,t}$)

Spec.	g	b	μ	z	λ_f	λ_w	r^m
Gaussian	0.191	0.054	0.095	0.101	0.282	0.158	0.350
Student t	0.178	0.026	0.060	0.072	0.141	0.142	0.175
SV	0.146	0.024	0.103	0.187	0.227	0.149	0.210
SV + t	0.143	0.027	0.074	0.083	0.124	0.122	0.155

Table B.15: Cross Correlation of "Tamed" Shocks ($\eta_{q,t}$)

	g	b	μ	z	λ_f	λ_w	r^m
			Gaussian				
g	1.000	0.193	0.130	0.289	0.080	0.179	0.255
b	0.193	1.000	0.200	-0.010	0.267	0.001	0.472
μ	0.130	0.200	1.000	0.099	0.107	0.030	0.308
z	0.289	-0.010	0.099	1.000	0.111	0.251	0.106
λ_f	0.080	0.267	0.107	0.111	1.000	0.047	0.181
λ_w	0.179	0.001	0.030	0.251	0.047	1.000	0.039
r^m	0.255	0.472	0.308	0.106	0.181	0.039	1.000
			Student's t				
g	1.000	0.064	0.074	0.181	0.059	0.043	0.179
b	0.064	1.000	0.102	-0.014	0.110	0.036	0.159
μ	0.074	0.102	1.000	0.027	-0.010	0.044	0.106
z	0.181	-0.014	0.027	1.000	0.049	0.090	0.109
λ_f	0.059	0.110	-0.010	0.049	1.000	0.029	0.094
λ_w	0.043	0.036	0.044	0.090	0.029	1.000	0.040
r^m	0.179	0.159	0.106	0.109	0.094	0.040	1.000
			SV				
g	1.000	0.089	0.096	0.299	0.021	0.231	0.152
b	0.089	1.000	0.091	-0.053	0.308	0.027	0.380
μ	0.096	0.091	1.000	0.110	0.047	0.068	0.251
z	0.299	-0.053	0.110	1.000	0.052	0.297	0.153
λ_f	0.021	0.308	0.047	0.052	1.000	0.075	0.162
λ_w	0.231	0.027	0.068	0.297	0.075	1.000	0.107
r^m	0.152	0.380	0.251	0.153	0.162	0.107	1.000
			SV + Student's t				
g	1.000	0.044	0.054	0.176	0.028	0.077	0.154
b	0.044	1.000	0.114	-0.032	0.115	0.071	0.160
μ	0.054	0.114	1.000	0.029	0.026	0.058	0.130
z	0.176	-0.032	0.029	1.000	0.029	0.112	0.111
λ_f	0.028	0.115	0.026	0.029	1.000	0.073	0.089
λ_w	0.077	0.071	0.058	0.112	0.073	1.000	0.074
r^m	0.154	0.160	0.130	0.111	0.089	0.074	1.000

Table B.16: Autocorrelation of $\tilde{h}^{-1/2}$

Spec.	g	b	μ	z	λ_f	λ_w	r^m
No SV	0.033	0.005	0.011	0.019	0.023	0.024	0.061
SV	0.022	0.004	0.013	0.018	0.018	0.018	0.032

Table B.17: Cross Correlation of $\tilde{h}^{-1/2}$

	g	b	μ	z	λ_f	λ_w	r^m
			Student t				
g	1.000	0.013	0.013	0.039	0.010	0.007	0.045
b	0.013	1.000	0.022	-0.004	0.022	0.007	0.047
μ	0.013	0.022	1.000	0.005	-0.002	0.007	0.026
z	0.039	-0.004	0.005	1.000	0.010	0.019	0.033
λ_f	0.010	0.022	-0.002	0.010	1.000	0.005	0.022
λ_w	0.007	0.007	0.007	0.019	0.005	1.000	0.009
r^m	0.045	0.047	0.026	0.033	0.022	0.009	1.000
			SV + Student t				
g	1.000	0.008	0.009	0.034	0.004	0.012	0.027
b	0.008	1.000	0.023	-0.008	0.021	0.013	0.035
μ	0.009	0.023	1.000	0.005	0.004	0.010	0.025
z	0.034	-0.008	0.005	1.000	0.004	0.021	0.024
λ_f	0.004	0.021	0.004	0.004	1.000	0.011	0.015
λ_w	0.012	0.013	0.010	0.021	0.011	1.000	0.013
r^m	0.027	0.035	0.025	0.024	0.015	0.013	1.000

Figure B.6: Shocks (absolute values) and smoothed stochastic volatility component, $\sigma_{q,t}$
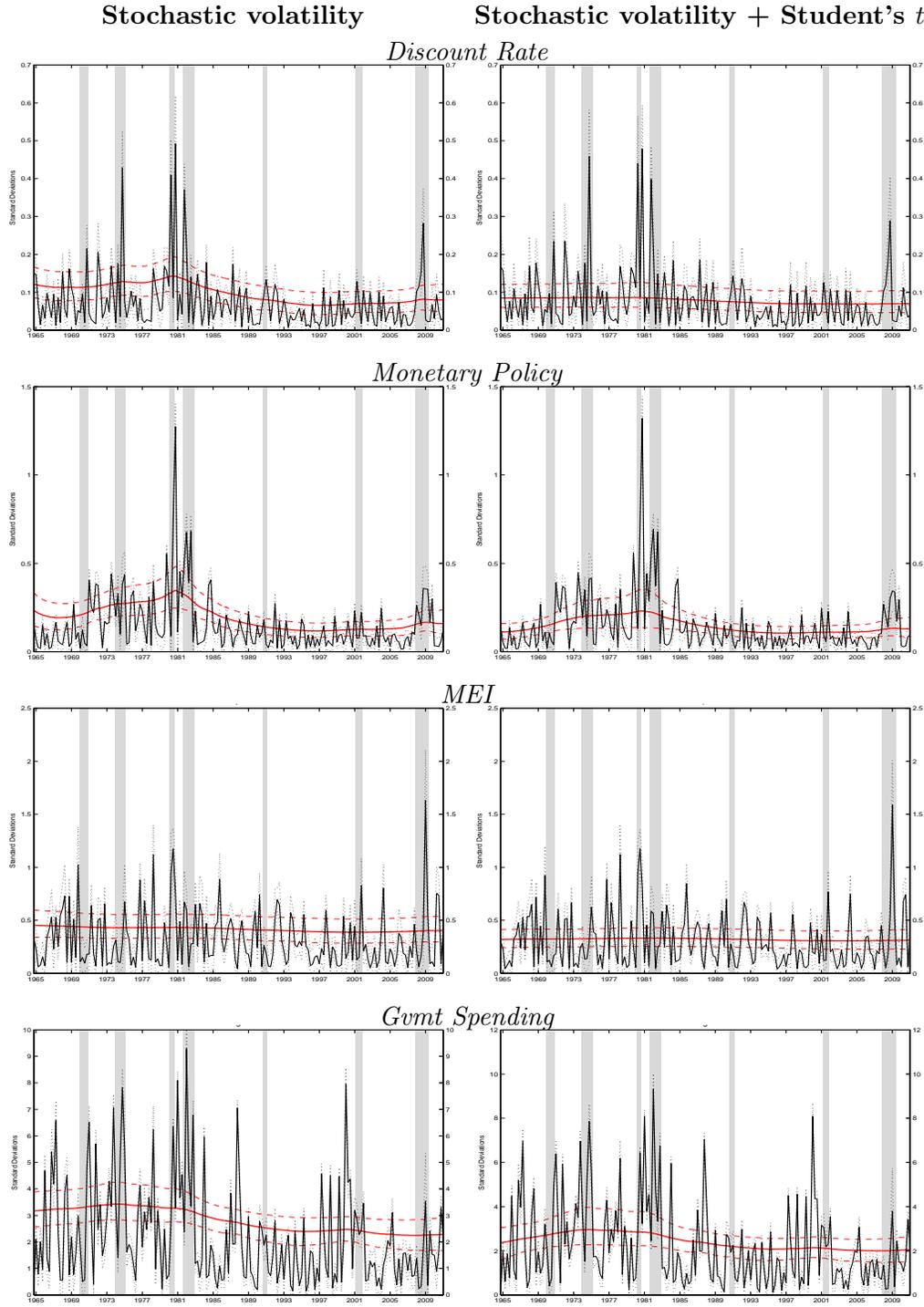
Stochastic volatility   **Stochastic volatility + Student's t**

Discount Rate

Monetary Policy

MEI

Gvmt Spending

Notes: Estimation with Student's t distribution with $\underline{\lambda} = 6$. The solid line is the median, and the dashed lines are the posterior 90% bands. Black line is the absolute value of the shock, and the red line is the stochastic volatility component.

Figure B.6 — Continued
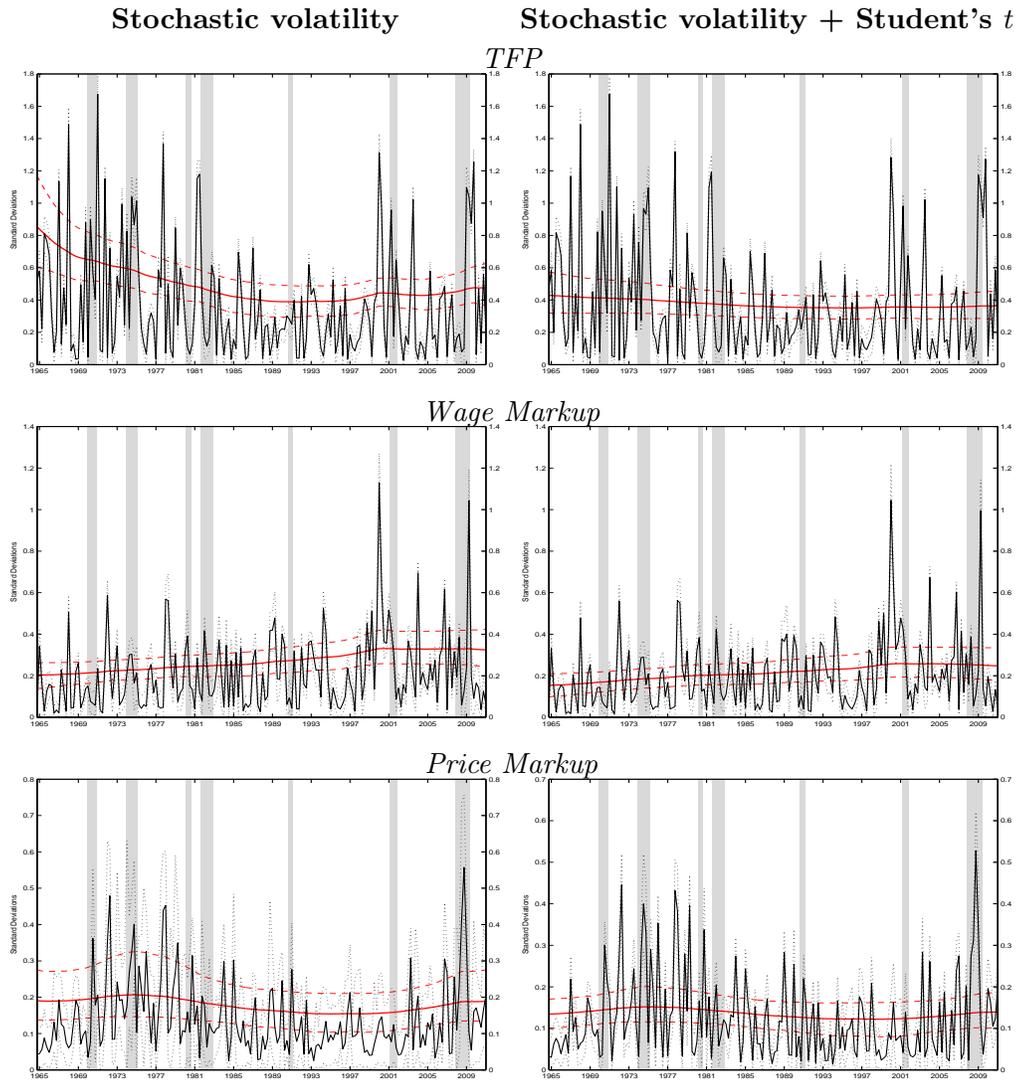
Stochastic volatility   **Stochastic volatility + Student's t**

TFP

Wage Markup

Price Markup

Notes: Estimation with Student's t distribution with $\underline{\lambda} = 6$. The solid line is the median, and the dashed lines are the posterior 90% bands. Black line is the absolute value of the shock, and the red line is the stochastic volatility component.

B.4 Robustness to Different Assumptions and Estimation Approches for the Stochastic Volatility Component

B.4.1 JPR Version of Algorithm

Table B.18: Log-Marginal Likelihoods, JPR Algorithm

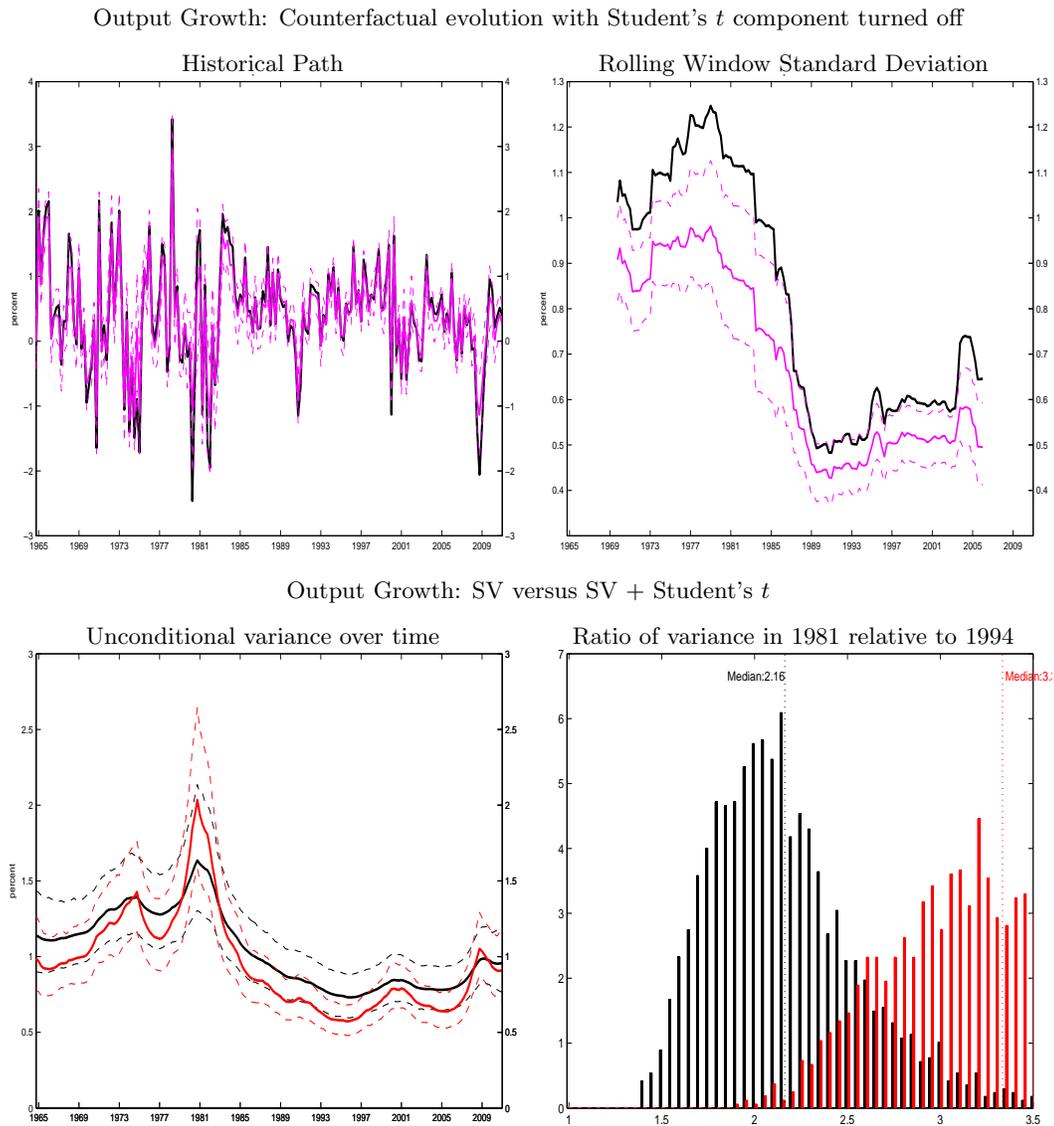|	Constant Volatility	Stochastic Volatility
Gaussian shocks		
	-1117.9	-975.2
Student's t distributed shocks		
$\underline{\lambda} = 15$	-999.8	-945.1
$\underline{\lambda} = 9$	-990.6	-936.3
$\underline{\lambda} = 6$	-975.9	-928.7

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

Table B.19: Posterior of the Student's t Degrees of Freedom, JPR Algorithm

	Without Stochastic Volatility			*With Stochastic Volatility*		
	$\underline{\lambda} = 15$	$\underline{\lambda} = 9$	$\underline{\lambda} = 6$	$\underline{\lambda} = 15$	$\underline{\lambda} = 9$	$\underline{\lambda} = 6$
Gvmt (g)	10.8	7.7	6.1	14.0	9.9	7.6
	(3.5,18.4)	(3.1,12.4)	(2.8,9.4)	(4.8, 23.0)	(4.2, 15.5)	(3.7, 11.6)
Discount (b)	8.6	6.8	5.7	7.9	6.4	5.4
	(3.4,14.0)	(3.2,10.5)	(2.8,8.4)	(2.6, 13.8)	(2.7, 10.2)	(2.6, 8.0)
MEI (μ)	11.0	8.0	6.5	3.4	7.6	6.2
	(3.7,18.4)	(3.3,12.7)	(3.1,9.8)	(2.9, 17.6)	(3.1, 12.0)	(2.9, 9.3)
TFP (z)	5.3	4.5	3.9	9.9	7.1	5.6
	(2.0,8.7)	(2.0,6.9)	(2.0,5.8)	(2.9, 17.3)	(2.7, 11.5)	(2.5, 8.7)
Price Markup (λ_f)	10.5	7.5	6.1	15.3	10.6	8.2
	(3.4,17.9)	(3.1,12.0)	(2.9,9.3)	(5.6, 24.7)	(4.6, 16.5)	(4.0, 12.3)
Wage Markup (λ_w)	10.9	8.1	6.5	11.9	8.6	6.9
	(3.8,18.1)	(3.5,12.6)	(3.2,9.7)	(4.1, 19.9)	(3.6, 13.4)	(3.4,10.3)
Policy (r^m)	3.2	3.0	2.9	15.0	10.5	8.1
	(1.7,4.6)	(1.7,4.3)	(1.7,4.1)	(5.4, 24.5)	(4.5, 16.4)	(3.9, 12.2)

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

Figure B.7: Results using JPR Algorithm

Output Growth: Counterfactual evolution with Student's t component turned off

Output Growth: SV versus SV + Student's t

Notes: Top panels: Black lines are the historical evolution of the variable, and magenta lines are the median counterfactual evolution of the same variable if we shut down the Student-t distributed component of all shocks. The rolling window standard deviation uses 20 quarters before and 20 quarters after a given quarter. Southwest panel: Black line is the unconditional standard deviation in the estimation with both stochastic volatility and Student-t components, while the red line is the unconditional variance in the estimation with stochastic volatility component only. Southeast panel: Black bars correspond to the posterior histogram of the ratio of volatility in 1981 over the variance in 1994 for the estimation with both stochastic volatility and Student-t components, while the red bars are for the estimation with with stochastic volatility component only.

B.4.2   SV-S Specification

Table B.20: Log-Marginal Likelihoods, SV-S Specification

	Constant Volatility	Stochastic Volatility
Gaussian shocks		
	-1117.9	-1100.2
Student's t distributed shocks		
$\underline{\lambda} = 6$	-975.9	-971.0

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

Table B.21: Posterior of the Student's t Degrees of Freedom, SV-S Specification

	Without Stochastic Volatility $\underline{\lambda} = 6$	*With Stochastic Volatility* $\underline{\lambda} = 6$
Gvmt (g)	6.1 (2.8,9.4)	7.2 (3.4, 11.0)
Discount (b)	5.7 (2.8,8.4)	4.1 (2.3, 5.8)
MEI (μ)	6.5 (3.1,9.8)	5.7 (2.8, 8.6)
TFP (z)	3.9 (2.0,5.8)	4.2 (2.1, 6.4)
Price Markup (λ_f)	6.1 (2.9,9.3)	6.6 (3.2, 10.0)
Wage Markup (λ_w)	6.5 (3.2,9.7)	6.1 (3.0, 9.0)
Policy (r^m)	2.9 (1.7,4.1)	3.2 (1.7, 4.6)

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

Table B.22: Posterior of SV Persistence Parameter, SV-S pecification

	Without Student t	With Student t
Gvmt (g)	0.495	0.879
	(0.091, 0.803)	(0.470, 1.000)
Discount (b)	0.711	0.475
	(0.295, 1.000)	(0.129, 0.781)
MEI (μ)	0.472	0.473
	(0.130, 0.782)	(0.133, 0.778)
TFP (z)	1.000	0.473
	(0.999, 1.000)	(0.126, 0.777)
Price Markup (λ_f)	0.477	0.477
	(0.132, 0.789)	(0.125, 0.785)
Wage Markup (λ_w)	0.514	0.475
	(0.200, 1.000)	(0.129, 0.781)
Policy (r^m)	0.990	0.518
	(0.977, 1.000)	(0.205, 1.000)

Notes: Numbers shown for the posterior mean and the 90% intervals of the SV persistence parameter.

Figure B.8: Results using SV-S specification.

Output Growth: Counterfactual evolution with Student's t component turned off

Output Growth: SV versus SV + Student's t

Notes: Top panels: Black lines are the historical evolution of the variable, and magenta lines are the median counterfactual evolution of the same variable if we shut down the Student-t distributed component of all shocks. The rolling window standard deviation uses 20 quarters before and 20 quarters after a given quarter. Southwest panel: Black line is the unconditional standard deviation in the estimation with both stochastic volatility and Student-t components, while the red line is the unconditional variance in the estimation with stochastic volatility component only. Southeast panel: Black bars correspond to the posterior histogram of the ratio of volatility in 1981 over the variance in 1994 for the estimation with both stochastic volatility and Student-t components, while the red bars are for the estimation with with stochastic volatility component only.

B.5   Subsample Analysis

Table B.23: Log-Marginal Likelihoods, Sub-samples

	Sample Ending in 2004Q4		*Sample Starting in 1984Q1*		*Sample Starting in 1991Q4*	
	Constant Volatility	Stochastic Volatility	Constant Volatility	Stochastic Volatility	Constant Volatility	Stochastic Volatility
Gaussian shocks						
	-964.0	-936.5	-521.3	-526.5	-378.1	-382.1
Student's t distributed shocks						
$\underline{\lambda} = 15$	-881.6	-870.1	-476.8	-479.56	-348.5	-341.9
$\underline{\lambda} = 9$	-870.6	-849.0	-471.1	-469.9	-339.5	-333.1
$\underline{\lambda} = 6$	-858.8	-844.1	-460.4	-462.2	-329.9	-328.1

Notes: The parameter $\underline{\lambda}$ represents the prior mean for the degrees of freedom in the Student's t distribution.
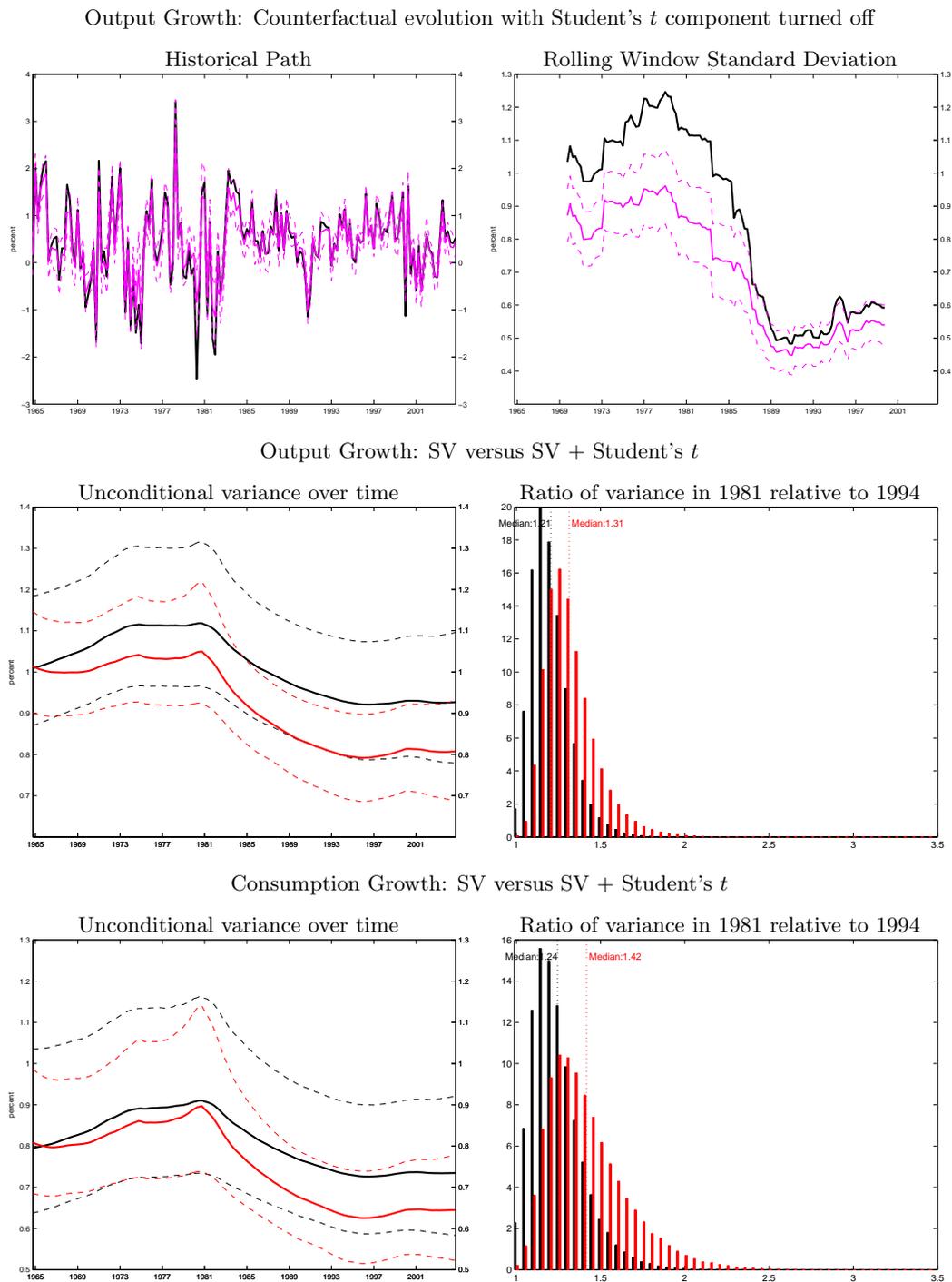
B.5.1 Sample Ending in 2004Q4

Table B.24: Posterior of the Student's t Degrees of Freedom, Sample Ending in 2004Q4

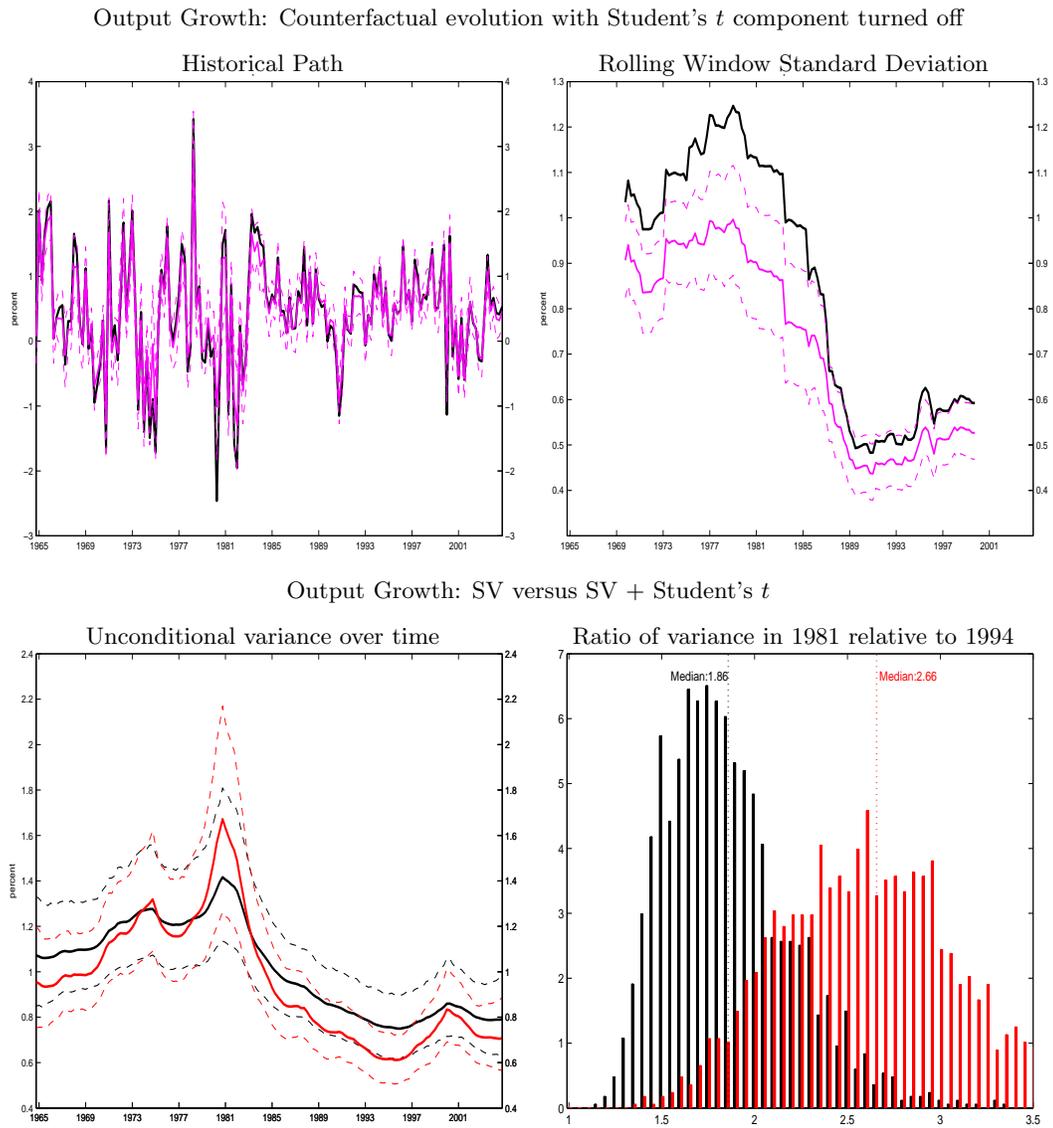|	Without Stochastic Volatility			With Stochastic Volatility		
	$\underline{\lambda}=15$	$\underline{\lambda}=9$	$\underline{\lambda}=6$	$\underline{\lambda}=15$	$\underline{\lambda}=9$	$\underline{\lambda}=6$
Gvmt (g)	10.8	7.7	6.1	11.2	8.0	6.3
	(3.5,18.4)	(3.1,12.4)	(2.8,9.4)	(3.7,18.8)	(3.3,12.7)	(2.9,9.6)
Discount (b)	8.6	6.8	5.7	9.4	7.2	5.7
	(3.4,14.0)	(3.2,10.5)	(2.8,8.4)	(3.3,15.5)	(3.1,11.2)	(2.8,8.6)
MEI (μ)	11.0	8.0	6.5	11.3	8.3	6.6
	(3.7,18.4)	(3.3,12.7)	(3.1,9.8)	(3.8,19.0)	(3.4,13.1)	(3.1,10.1)
TFP (z)	5.3	4.5	3.9	6.5	5.0	4.3
	(2.0,8.7)	(2.0,6.9)	(2.0,5.8)	(2.3,11.0)	(2.2,7.9)	(2.1,6.5)
Price Markup (λ_f)	10.5	7.5	6.1	11.7	8.5	7.0
	(3.4,17.9)	(3.1,12.0)	(2.9,9.3)	(4.0,19.5)	(3.5,13.4)	(3.2,10.7)
Wage Markup (λ_w)	10.9	8.1	6.5	12.1	8.8	6.9
	(3.8,18.1)	(3.5,12.6)	(3.2,9.7)	(4.2,20.1)	(3.7,13.7)	(3.4,10.3)
Policy (r^m)	3.2	3.0	2.9	9.4	7.0	5.6
	(1.7,4.6)	(1.7,4.3)	(1.7,4.1)	(2.6,16.6)	(2.5,11.4)	(2.4,8.8)

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

Figure B.9: Results for sub-sample ending in 2004Q4.

Output Growth: Counterfactual evolution with Student's t component turned off

Output Growth: SV versus SV + Student's t

Consumption Growth: SV versus SV + Student's t

Notes: Top panels: Black lines are the historical evolution of the variable, and magenta lines are the median counterfactual evolution of the same variable if we shut down the Student-t distributed component of all shocks. The rolling window standard deviation uses 20 quarters before and 20 quarters after a given quarter. Southwest panel: Black line is the unconditional standard deviation in the estimation with both stochastic volatility and Student-t components, while the red line is the unconditional variance in the estimation with stochastic volatility component only. Southeast panel: Black bars correspond to the posterior histogram of the ratio of volatility in 1981 over the variance in 1994 for the estimation with both stochastic volatility and Student-t components, while the red bars are for the estimation with with stochastic volatility component only.

Figure B.10: Results using JPR Algorithm — sub-sample ending in 2004Q4

Output Growth: Counterfactual evolution with Student's t component turned off

Historical Path

Rolling Window Standard Deviation

Output Growth: SV versus SV + Student's t

Unconditional variance over time

Ratio of variance in 1981 relative to 1994

Notes: Top panels: Black lines are the historical evolution of the variable, and magenta lines are the median counterfactual evolution of the same variable if we shut down the Student-t distributed component of all shocks. The rolling window standard deviation uses 20 quarters before and 20 quarters after a given quarter. Southwest panel: Black line is the unconditional standard deviation in the estimation with both stochastic volatility and Student-t components, while the red line is the unconditional variance in the estimation with stochastic volatility component only. Southeast panel: Black bars correspond to the posterior histogram of the ratio of volatility in 1981 over the variance in 1994 for the estimation with both stochastic volatility and Student-t components, while the red bars are for the estimation with with stochastic volatility component only.

Figure B.11: $\tilde{h}_{q,t}^{-1/2}$ — sub-sample ending in 2004Q4
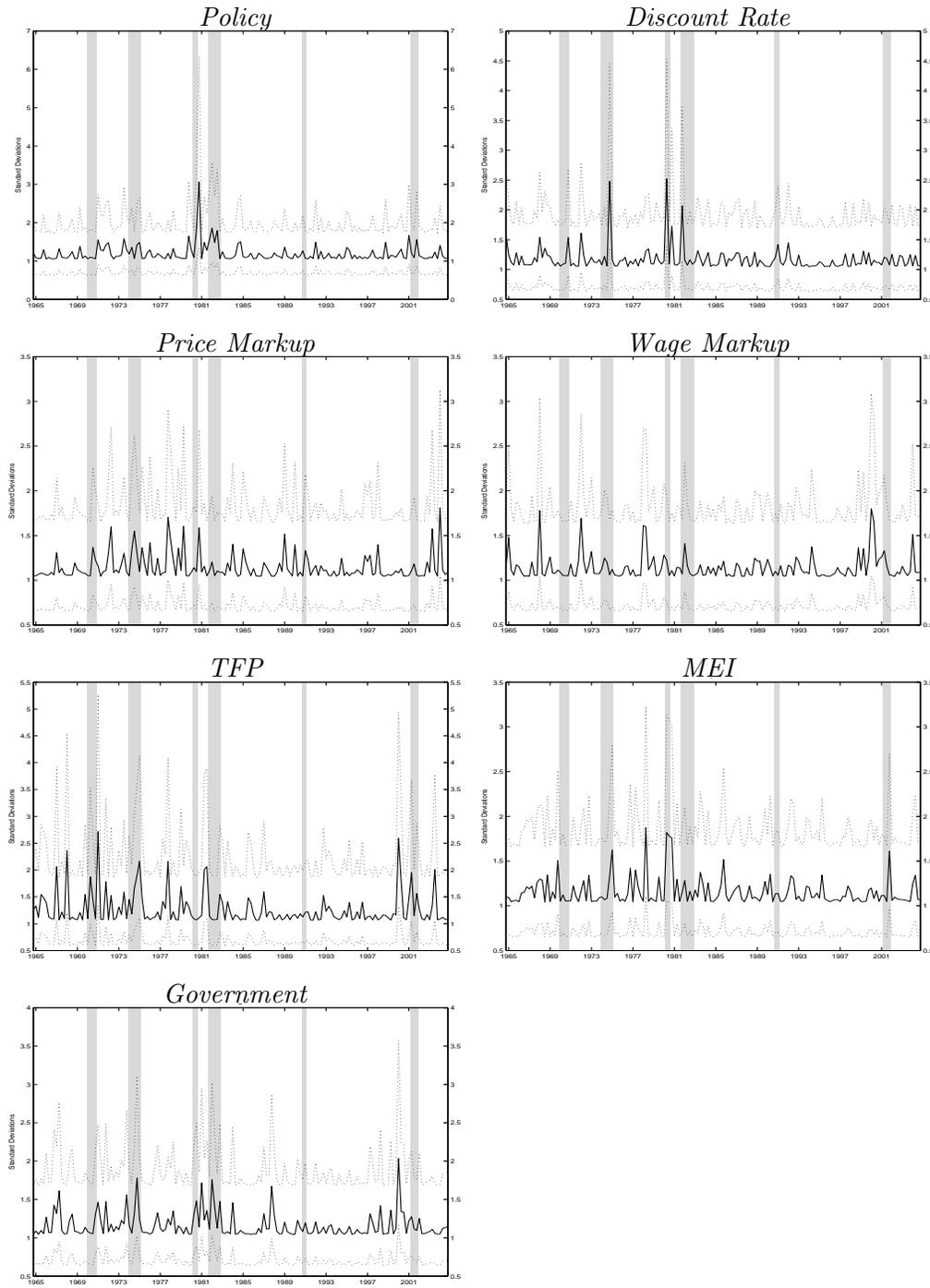
Policy

Discount Rate

Price Markup

Wage Markup

TFP

MEI

Government

Notes: Estimation with Student's t distribution with $\underline{\lambda} = 6$. The solid line is the median, and the dashed lines are the posterior 90% bands. Black line is the absolute value of the shock, and the red line is the stochastic volatility component.

Figure B.12: Shocks (absolute values) and smoothed stochastic volatility component, $\sigma_{q,t}$ — sub-sample ending in 2004Q4
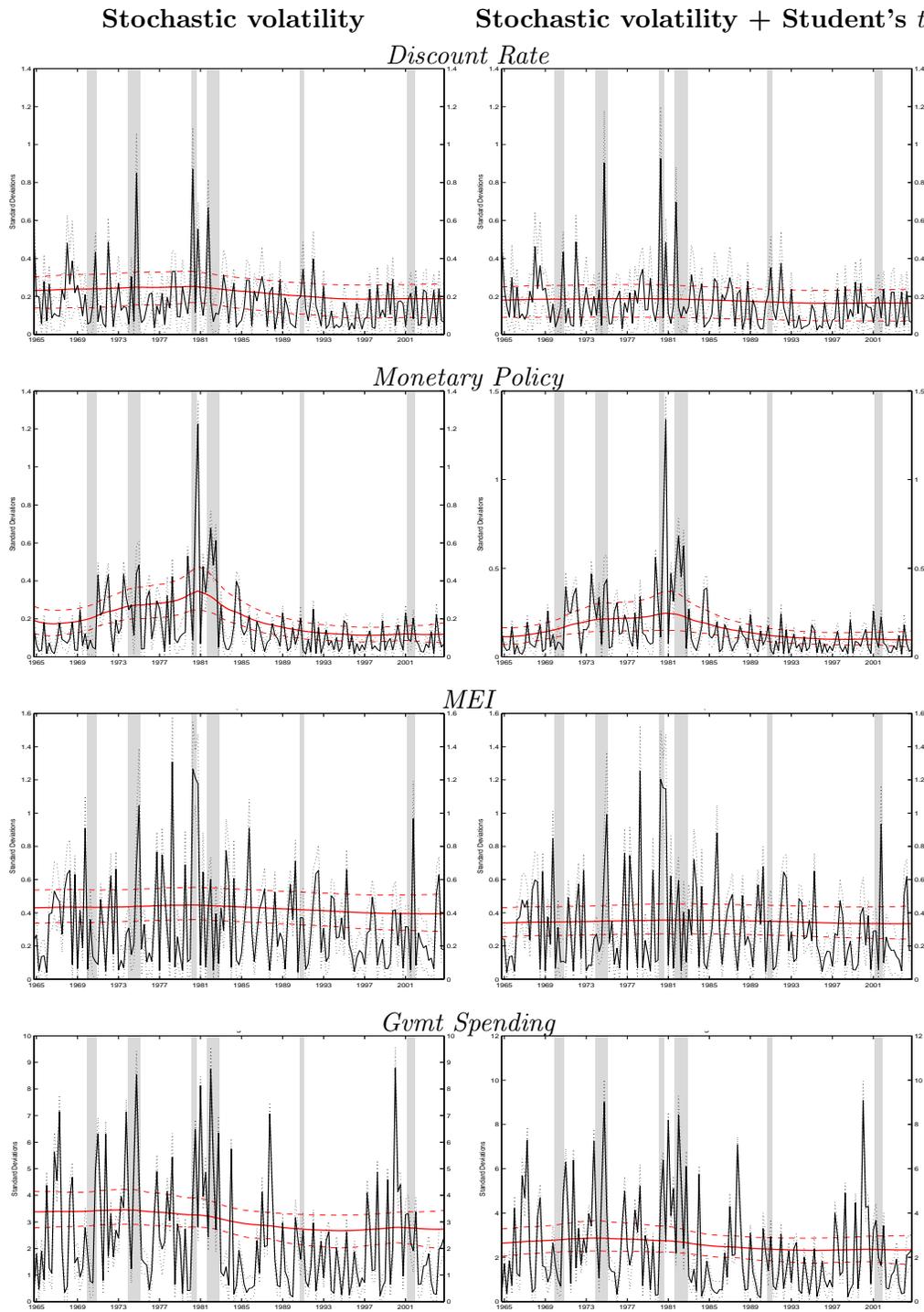
Stochastic volatility   **Stochastic volatility + Student's t**

Discount Rate

Monetary Policy

MEI

Gvmt Spending

Notes: Estimation with Student's t distribution with $\underline{\lambda} = 6$. The solid line is the median, and the dashed lines are the posterior 90% bands. Black line is the absolute value of the shock, and the red line is the stochastic volatility component.

Figure B.12 — Continued

Stochastic volatility   **Stochastic volatility + Student's t**

TFP
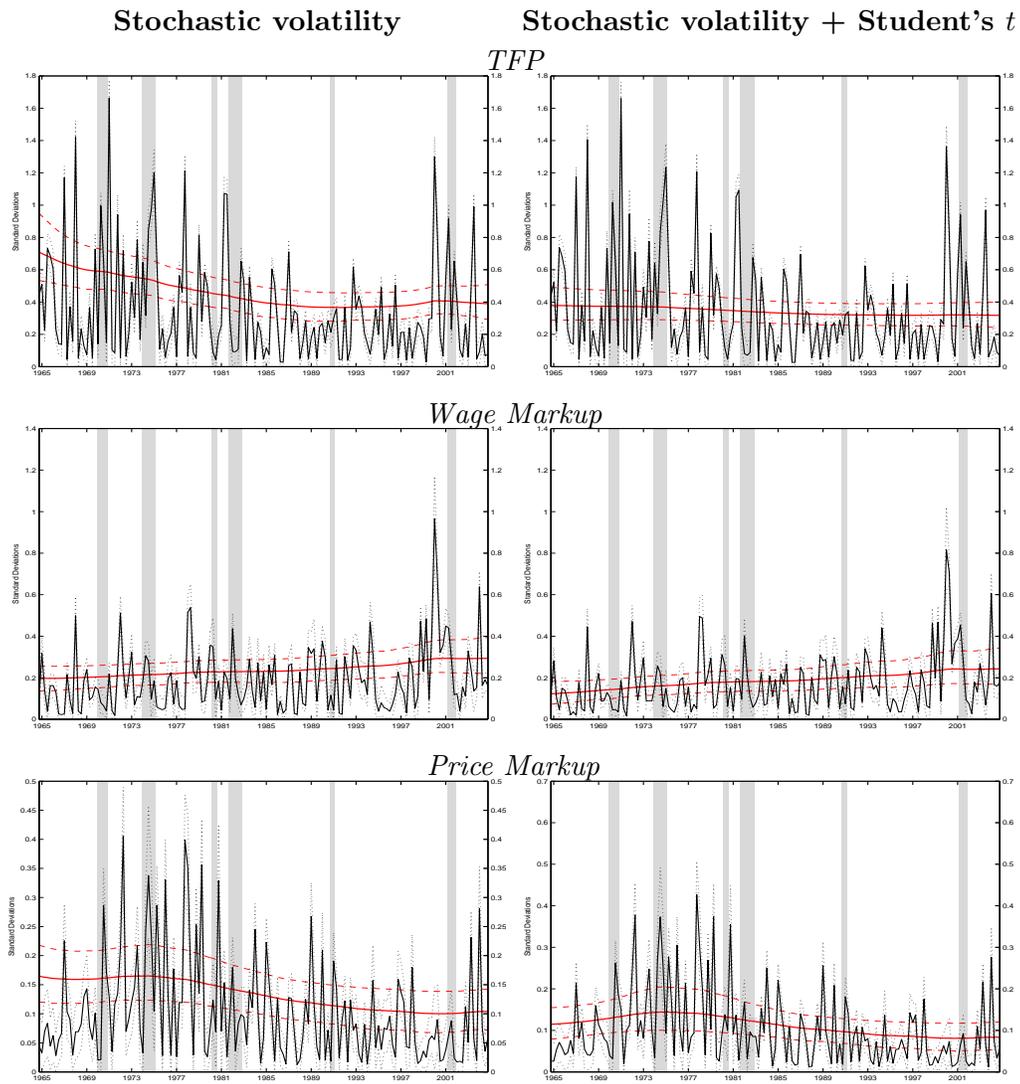

Wage Markup

Price Markup

Notes: Estimation with Student's t distribution with $\underline{\lambda} = 6$. The solid line is the median, and the dashed lines are the posterior 90% bands. Black line is the absolute value of the shock, and the red line is the stochastic volatility component.

Table B.25: Comparison of Parameter Estimates: Subsample ending in 2004Q4 vs. full sample

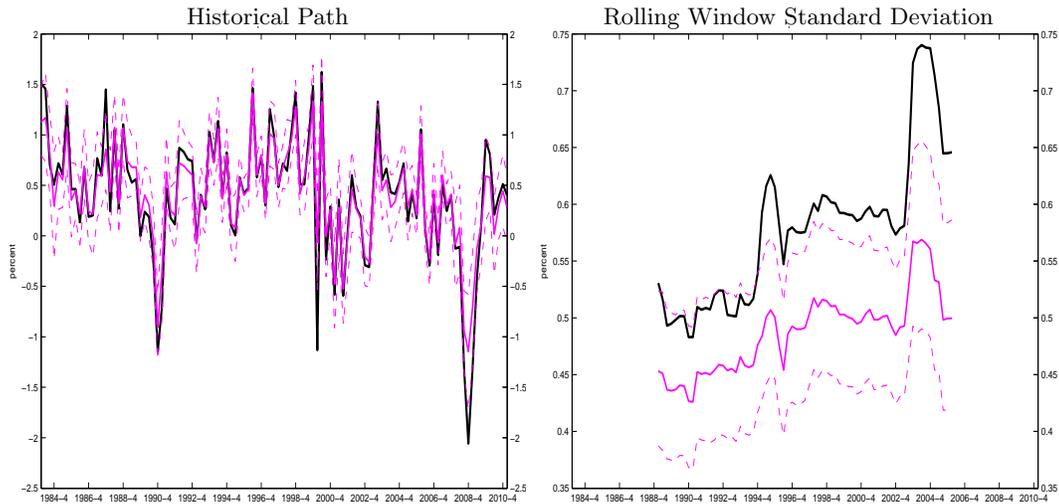|	Full Sample				To 2004Q4			
	Base	SV	TD	SVTD	Base	SV	TD	SVTD
α	0.15	0.134	0.15	0.135	0.174	0.17	0.181	0.179
ζ_p	0.734	0.78	0.808	0.846	0.696	0.776	0.763	0.82
ι_p	0.315	0.344	0.383	0.286	0.291	0.306	0.313	0.289
Φ	1.58	1.518	1.575	1.551	1.704	1.686	1.719	1.677
S''	4.686	5.013	5.07	5.651	6.114	6.693	7.169	6.516
h	0.611	0.609	0.582	0.571	0.703	0.703	0.734	0.68
ψ	0.714	0.734	0.67	0.666	0.626	0.637	0.547	0.521
ν_l	2.088	2.212	2.3	2.476	2.349	2.67	2.601	2.636
ζ_w	0.803	0.826	0.83	0.843	0.756	0.802	0.812	0.829
ι_w	0.541	0.547	0.495	0.511	0.584	0.554	0.528	0.501
β	0.206	0.184	0.202	0.175	0.167	0.151	0.162	0.161
ψ_1	1.953	1.866	1.82	1.884	2.066	2.071	1.957	1.934
ψ_2	0.083	0.073	0.115	0.116	0.09	0.104	0.12	0.138
ψ_3	0.245	0.217	0.213	0.184	0.238	0.207	0.193	0.189
π^*	0.683	0.719	0.706	0.808	0.709	0.784	0.756	0.847
σ_c	1.236	1.109	1.248	1.274	1.406	1.426	1.471	1.42
ρ	0.835	0.854	0.875	0.875	0.825	0.852	0.867	0.874
γ	0.306	0.321	0.356	0.389	0.415	0.421	0.431	0.434
$\bar{L}$	-44.17	-46.67	-43.38	-44.73	-42.897	-43.542	-43.348	-44.265
ρ_g	0.977	0.977	0.982	0.988	0.98	0.981	0.98	0.981
ρ_b	0.758	0.845	0.844	0.852	0.285	0.335	0.279	0.453
ρ_μ	0.748	0.753	0.791	0.806	0.735	0.739	0.746	0.772
ρ_z	0.994	0.991	0.987	0.981	0.963	0.961	0.959	0.963
ρ_{λ_f}	0.791	0.797	0.811	0.83	0.891	0.831	0.893	0.863
ρ_{λ_w}	0.981	0.952	0.962	0.923	0.969	0.951	0.928	0.902
ρ_{rm}	0.154	0.219	0.219	0.227	0.145	0.15	0.183	0.179
σ_g	2.892	3.169	2.387	2.665	3.091	3.388	2.566	2.654
σ_b	0.125	0.122	0.072	0.1	0.232	0.23	0.192	0.163
σ_μ	0.43	0.454	0.325	0.3	0.435	0.43	0.35	0.334
σ_z	0.493	0.869	0.362	0.473	0.463	0.717	0.331	0.38
σ_{λ_f}	0.164	0.191	0.163	0.127	0.136	0.166	0.111	0.117
σ_{λ_w}	0.281	0.203	0.213	0.151	0.255	0.197	0.188	0.12
σ_{rm}	0.228	0.243	0.133	0.095	0.235	0.196	0.133	0.112
η_{gz}	0.787	0.775	0.786	0.765	0.747	0.736	0.74	0.736
η_{λ_f}	0.67	0.749	0.815	0.734	0.73	0.681	0.771	0.775
η_{λ_w}	0.948	0.914	0.924	0.865	0.887	0.876	0.833	0.797

B.5.2 Sample Starting in 1984Q1

Table B.26: Posterior of the Student's t Degrees of Freedom, Sample Starting in 1984Q1
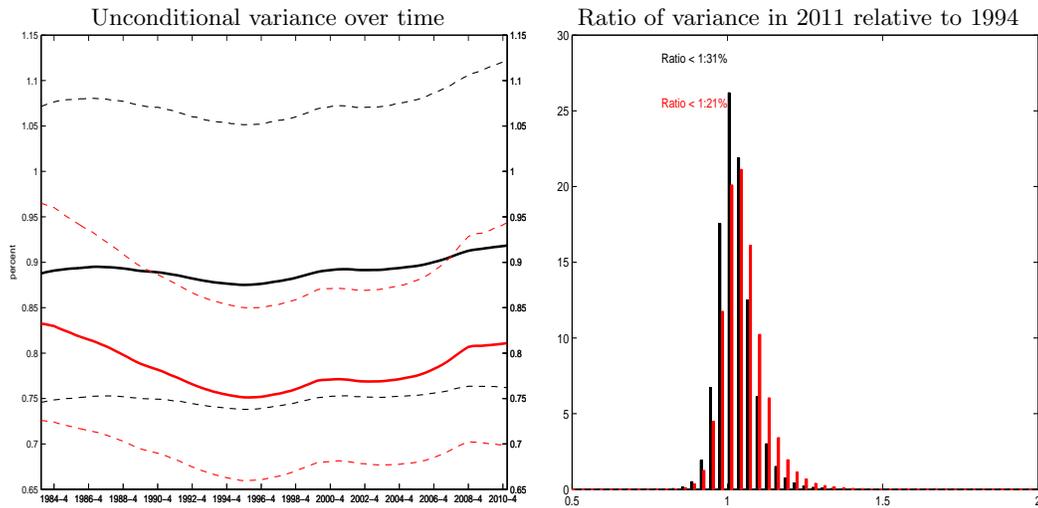
	Without Stochastic Volatility			With Stochastic Volatility		
	$\underline{\lambda} = 15$	$\underline{\lambda} = 9$	$\underline{\lambda} = 6$	$\underline{\lambda} = 15$	$\underline{\lambda} = 9$	$\underline{\lambda} = 6$
Gvmt (g)	7.6	5.8	4.8	7.6	5.7	4.7
	(2.4, 13.2)	(2.3, 9.3)	(2.2, 7.4)	(2.4, 13.2)	(2.3, 9.1)	(2.1, 7.3)
Discount (b)	9.6	6.9	5.6	9.6	7.0	5.6
	(2.9, 16.4)	(2.7, 11.2)	(2.5, 8.6)	(2.9, 16.4)	(2.7, 11.3)	(2.5, 8.7)
MEI (μ)	10.0	7.6	6.1	10.0	7.5	6.0
	(3.3, 17.0)	(3.1, 12.0)	(2.8, 9.3)	(3.4, 17.0)	(3.0, 11.8)	(2.8, 9.2)
TFP (z)	6.8	5.2	4.3	7.5	5.6	4.6
	(2.1, 11.8)	(2.1, 8.4)	(2.0, 6.7)	(2.3, 13.1)	(2.2, 9.1)	(2.1, 7.2)
Price Markup (λ_f)	8.9	6.4	5.1	10.3	7.0	5.5
	(2.5, 15.7)	(2.2, 10.4)	(2.2, 8.0)	(2.8, 18.0)	(2.6, 11.6)	(2.3, 8.6)
Wage Markup (λ_w)	9.5	7.2	5.8	10.4	7.6	6.1
	(3.1, 16.2)	(2.9, 11.4)	(2.7, 8.8)	(3.3, 17.7)	(3.1, 12.2)	(2.8, 9.4)
Policy (r^m)	10.6	7.6	6.1	10.9	7.7	6.0
	(3.3, 18.0)	(3.0, 12.3)	(2.7, 9.3)	(3.3, 18.5)	(3.0, 12.4)	(2.7, 9.3)

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

Figure B.13: Results for sub-sample starting in 1984Q1
Output Growth: Counterfactual evolution with Student's t component turned off

Historical Path

Rolling Window Standard Deviation

Output Growth: SV versus SV + Student's t

Unconditional variance over time

Ratio of variance in 2011 relative to 1994

Notes: Top panels: Black lines are the historical evolution of the variable, and magenta lines are the median counterfactual evolution of the same variable if we shut down the Student-t distributed component of all shocks. The rolling window standard deviation uses 20 quarters before and 20 quarters after a given quarter. Southwest panel: Black line is the unconditional standard deviation in the estimation with both stochastic volatility and Student-t components, while the red line is the unconditional variance in the estimation with stochastic volatility component only. Southeast panel: Black bars correspond to the posterior histogram of the ratio of volatility in 2011 over the variance in 1994 for the estimation with both stochastic volatility and Student-t components, while the red bars are for the estimation with with stochastic volatility component only.
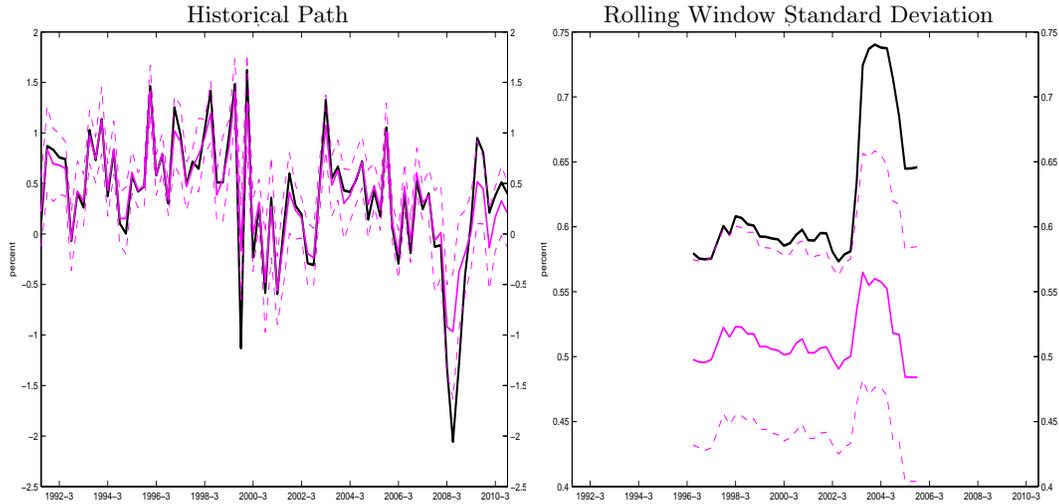
B.5.3   Sample Starting in 1991Q4

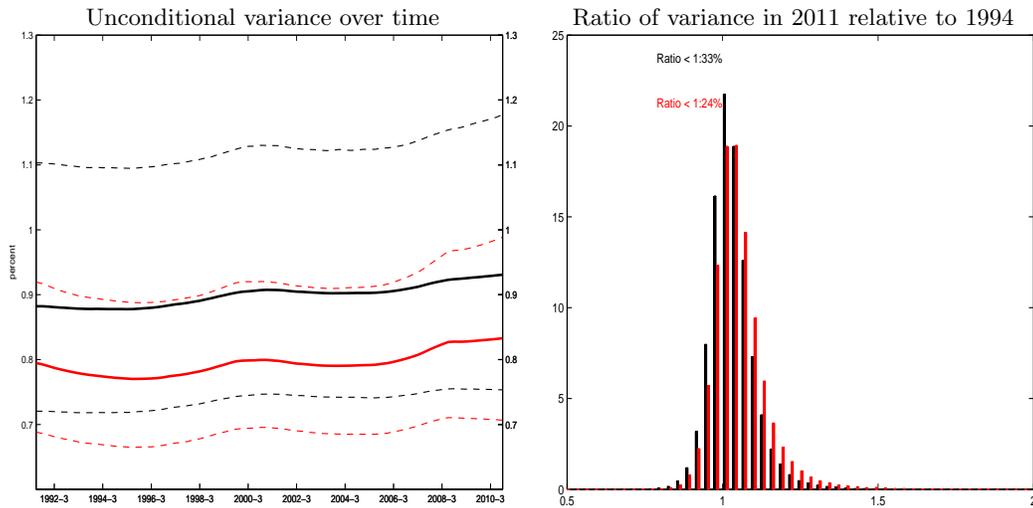Table B.27: Posterior of the Student's t Degrees of Freedom, Sample Starting in 1991Q4

	Without Stochastic Volatility			With Stochastic Volatility		
	$\underline{\lambda}=15$	$\underline{\lambda}=9$	$\underline{\lambda}=6$	$\underline{\lambda}=15$	$\underline{\lambda}=9$	$\underline{\lambda}=6$
Gvmt (g)	9.9	6.9	5.4	10.0	7.1	5.5
	(2.7,17.1)	(2.5,11.4)	(2.2,8.5)	(2.7,17.5)	(2.5,11.6)	(2.3,8.7)
Discount (b)	9.8	7.4	5.6	10.1	7.3	5.8
	(3.0,16.9)	(2.8,12.0)	(2.4,8.7)	(3.0,17.4)	(2.7,11.8)	(2.5,9.1)
MEI (μ)	7.0	5.5	4.6	7.2	5.6	4.6
	(2.3,11.9)	(2.1,8.8)	(2.0,7.2)	(2.3,12.3)	(2.2,9.1)	(2.0,7.2)
TFP (z)	7.6	5.4	4.2	7.9	5.6	4.4
	(1.9,13.9)	(1.8,9.1)	(1.7,6.8)	(2.0,14.3)	(1.9,9.4)	(1.8,7.1)
Price Markup (λ_f)	6.2	5.0	3.7	10.0	6.9	5.4
	(1.6,11.2)	(1.7,8.4)	(1.5,6.0)	(2.1,18.2)	(1.9,11.7)	(1.9,8.9)
Wage Markup (λ_w)	10.6	7.4	5.7	10.7	7.4	5.8
	(3.0,18.3)	(2.6,12.1)	(2.4,9.0)	(3.0,18.5)	(2.6,12.2)	(2.4,9.1)
Policy (r^m)	11.7	7.8	5.9	11.7	7.9	5.9
	(3.1,20.2)	(2.6,12.9)	(2.4,9.4)	(3.1,20.1)	(2.7,13.0)	(2.4,9.4)

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

Figure B.14: Results for sub-sample starting in 1991Q4
Output Growth: Counterfactual evolution with Student's t component turned off

Output Growth: SV versus SV + Student's t

Notes: Top panels: Black lines are the historical evolution of the variable, and magenta lines are the median counterfactual evolution of the same variable if we shut down the Student-t distributed component of all shocks. The rolling window standard deviation uses 20 quarters before and 20 quarters after a given quarter. Southwest panel: Black line is the unconditional standard deviation in the estimation with both stochastic volatility and Student-t components, while the red line is the unconditional variance in the estimation with stochastic volatility component only. Southeast panel: Black bars correspond to the posterior histogram of the ratio of volatility in 2011 over the variance in 1994 for the estimation with both stochastic volatility and Student-t components, while the red bars are for the estimation with with stochastic volatility component only.