Rare Shocks, Great Recessions

Vasco Cúrdia, Marco Del Negro, Daniel Greenwald[*]

Federal Reserve Bank of New York and *New York University*

February 14, 2012; First Draft: March 2011

WORK IN PROGRESS

Abstract

We estimate a DSGE model where rare large shocks can occur, by replacing the commonly used Gaussian assumption with a Student-t distribution. We show that the latter is favored by the data in the context of a Smets and Wouters-type model estimated on macro variables, even if we allow for low frequency variation in the shocks' volatility. The evidence is even stronger when we introduce financial frictions as in Bernanke, Gertler and Gilchrist (1999), and correspondingly include a measure of interest rate spreads among the observables. We provide some evidence that introducing Student-t shocks reduces the importance of low-frequency time-variation in volatility. In particular, we show that the Great Recession of 2008-09 does not result in significant increases in estimated time-varying volatility (i.e., it is not a reversal of the Great Moderation) but is largely the outcome of large shocks.

---

1 Introduction

Financial crises do not happen every decade. 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 large shocks. This cannot be done using the normal distribution, which is the standard assumption in the DSGE literature. We estimate a linearized DSGE model assuming that shocks are generated from a Student-t distribution, which is designed to capture fat tails. The number of degrees of freedom in the Student-t distribution, which determines the likelihood of having rare large shocks, is estimated from the data. We show that estimating DSGE models with Student-t distributed shocks is a fairly straightforward extension of current methods (as described, for instance, in An and Schorfheide (2007)). In fact, the Gibbs sampler a simple extension of Geweke (1993)'s Gibbs sampler for a linear model to the DSGE framework. In addition to allowing for fat tails, we introduce time-varying volatilities following the approach in Justiniano and Primiceri (2008). The paper is also 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-t distributed shocks.

We show that the normality assumption for the distribution of shocks in DSGE models is counterfactual, even when these models are estimated using the standard set of macroeconomic variables over the post-Great Moderation period. To some extent this is not too surprising given that we had evidence, even before the recent recession, that macro variables do not quite conform to the gaussian assumptions (see Christiano (2007)). The normality assumption is even more counterfactual when we include financial variables — such as spreads — into the set of observables, which is a natural step if one is to introduce financial frictions into the model.

There are of course ways to introduce departures from normality in the shocks distribution other than assuming that shocks are Student-t distributed. Two important papers in the DSGE literature, Justiniano and Primiceri (2008) and Fernández-Villaverde and Rubio-Ramírez (2007), have introduced stochastic volatility in the

estimation of DSGE models. In both papers stochastic volatility captures low frequency movements in volatility however, as they focused on the sources of the Great Moderation. In particular, in Justiniano and Primiceri (2008) this is by assumption, as they postulate a random walk as the law of motion of the volatilities. Low frequency movements in volatility, while potentially important to describe the data, cannot capture rare large shocks. Most importantly, we argue that the presence of large shocks may potentially distort the assessment of these low frequency movements. At the same time. ignoring low frequency movements in volatiltiy may bias the results toward finding evidence in favor of fat tails. For this reasons it is important to allow for both fat tails and low frequency movements in volatility in the estimation.

The top panel of Figure 1 shows the time series of the smoothed "entrepreneurial risk" shocks (in absolute value) from a model that includes financial frictions modeled as in Bernanke et al. (1999) and Christiano et al. (2009). The model is estimated under the assumption of normality on the post-war data. These shocks capture mean-preserving variations in the volatility of idiosyncratic entrepreneurial productivity, and lead to corresponding changes in spreads. We will therefore often refer to these shocks as "spread" shocks. The shocks are normalized, that is, expressed in standard deviations units. The solid line is the median, and the dashed lines are the posterior 90% bands. Figure 1 shows that the size of the entrepreneurial risk shock in 2008Q4, the quarter immediately following Lehman Brothers' collapse, is between 6 and 7 standard deviations. The bottom panel of Figure 1 shows the evolution of the smoothed monetary policy shocks (again, normalized, and in absolute value).

These two plots illustrate why we need both fat tails and time-varying volatility. Imagine estimating a model with only 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 we will show, such model will produce a time series of volatilities peaking around 1980, and then again during the Great Recession. Viceversa, imagine estimating a model with only fat tails but no slow moving time variation in volatility on shocks that fit the pattern shown in the bottom panel. Here the model would interpret

all the large shocks in the late 70s and 80s as rare shocks, ignoring the evident clustering of these shocks – that is, the fact that the volatility of monetary policy shocks was low in the sixties, high during the Volcker period, and then low again.

The main focus of this paper is twofold. We ask whether there is evidence of fat tails, and whether tail events are important for the macroeconomy. Next, we ask whether allowing for Student-t distributed shocks affects the inference about time-variation in volatility. In particular, we investigate whether the Great Recession is best captured as a reversal of the Great Moderation, or as the outcome of rare large shocks.

Our results so far can be summarized as follows. First, perhaps not surprisingly in light of the plots for the shocks in Figure 1, we find very strong evidence in favor of Student-t distributed shocks. The estimated degree of freedom for the Student-t distribution vary across shocks, but for some shocks, notably spread shocks, are quite low, indicating substantial Kurtosis. For this reason, allowing for Student-t distributed shocks substantially improves the model's fit, more than just allowing for (random walk) time variation in volatilities. Second, rare large shocks, and in particular rare large shocks to the financial sector, explain an important part of the Great Recession. This finding shows that we do not just need fat tails to fit financial variables such as spreads, but that these tail events have an important impact on the macroeconomy. Last, we show that slow-moving stochastic volatility is less important in the presence of rare large shocks. In fact, whenever we allow for Student-t distributed shocks adding stochastic volatility does not improve the marginal likelihood. This result is preliminary, and may well depend on our imposing a random walk as the law of motion of the volatilities. In any case, even allowing for fat tails we find evidence that for some shocks, e.g. policy shocks, stochastic volatility is important.

It is important to point out a number of caveats regarding our analysis. For one, in the current draft we allow for excess Kurtosis but nor for skewness. The shocks plots in in Figure 1 make it plain that most large shocks occur during recessions. We plan to address this issue in future drafts. A recent paper by Müller (2011)

describes some of the dangers associated with departures from Gaussianity when the alternative shock distribution is also misspecified. Next, our model is linear. It may well then be that what we capture as large rare shocks are Gaussian shocks whose effect is amplified through a non-linear propagation mechanism. This is obviously an important line of research. At the same time, it is no excuse for sticking to gaussianity in linear models, as much of the literature has done so far. Moreover, this paper attempts to contribute to that literature by identifying what sort of non-linearities, and in which sector, may be most helpful. Finally, we should remark that our log-linear DSGE model with Student-t distributed shocks cannot be taken as a literal description of reality for the reason that if the model is correct expectations of the variables in levels are not defined. We are not particularly disturbed by this implication because we do not take the model literally. Rather, we interpret it a useful approximation meant to investigate specific questions, such as the extent to which, and the sources of, time-variation in macroeconomic volatility.

The next section discusses Bayesian inference. The section first describes the procedure used to estimate a DSGE model with Student-t distributed shocks, and then combines Student-t distributed shocks with time-variation in volatilities. Section 3 describes the model, as well as our set of observables. As far as results go, the paper at this point is more of an estended abstract than an actual paper. We provide evidence that the Student-t distribution assumption is favored by the data relative to Gaussianity, and this evidence is surprisngly stronger when spreads are included among the observables. We also show that the normality assumption leads to downplay the importance of large shocks relative to the Student-t distribution. We plan to do much more, including addressing some of the questions raised in the introduction, in the next few months.

2   Bayesian Inference

The first part of the section describes the estimation of a DSGE model with both Student-t distributed shocks and time-varying volatilities. The Gibbs sampler com-

bines the algorithm proposed by Geweke (1993)'s for a linear model with Student-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 Justiniano and Primiceri (2008). Section A.2 discusses the computation of the marginal likelihood.

The model consists of the standard measurement and transition equations:

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

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

for $t = 1, .., T$, where y_t, s_t, and ε_t are $n \times 1$, $k \times 1$, and $\bar{q} \times 1$ vector of observables, states, and shocks, respectively. Call $p(\theta)$ 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 a 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).

Define

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

where the parameters $\sigma_{1:\bar{q}}$ (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:

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

We consider two types of priors for ρ_q:

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

In the SV-UR case $\tilde{\sigma}_{q,t}$ follows a random walk as in Justiniano and Primiceri (2008), 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 very 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-AR 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 high values of ρ_q. As a consequence, $\sigma_{q,t}$ and $\tilde{h}_{q,t}$ play very different roles in (3): $\sigma_{q,t}$ allows for slow-moving trends in volatility, while $\tilde{h}_{q,t}$ allows for large shocks. Finally, to close the model we make the following distributional assumptions on the initial conditions $\tilde{\sigma}_{q,0}, \ q = 1, .., \bar{q}$:

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

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.

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)

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

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

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

where

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

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

The sampler consists of six blocks.

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

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

$$\begin{aligned} p(\theta|&\tilde{h}_{1:T}, \tilde{\sigma}_{1:T}, \lambda_{1:\bar{q}}, \rho_{1:\bar{q}}, \omega_{1:\bar{q}}^2, y_{1:T}) \\ &\propto \left[\int p(y_{1:T}|s_{1:T}, \theta) p(s_{1:T}|\varepsilon_{1:T}, \theta) p(\varepsilon_{1:T}|\tilde{h}_{1:T}, \tilde{\sigma}_{1:T}, \theta) \cdot d(s_{1:T}, \varepsilon_{1:T}) \right] p(\theta) \\ &= p(y_{1:T}|\tilde{h}_{1:T}, \tilde{\sigma}_{1:T}, \theta) p(\theta) \end{aligned}$$
$$(17)$$

where

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

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

$$\varepsilon_t|\tilde{h}_{1:T}, \tilde{\sigma}_{1:T} \sim \mathcal{N}(0, \Delta_t), \tag{18}$$

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

(1.2) 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).

(2) 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

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

which implies

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

(3) 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

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

This is a non-standard distribution, hence the draw is obtained from a Metropolis-Hastings step.

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

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

using the algorithm developed by Kim et al. (1998), which we briefly describe in appendix A.3.

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

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

where $\tilde{\sigma}_{q,1}^{*~2} = \tilde{\sigma}_{q,1}^2$ in the SV-UR case, and $\tilde{\sigma}_{q,1}^{*~2} = \tilde{\sigma}_{q,1}^2 / (1 - \rho_q^2)$ in the SV-S case. This implies that

$$\omega_q^2 | \tilde{\sigma}_{1:T}, \rho_{1:\bar{q}}, \cdots \sim \mathcal{IG} \left(\frac{\nu+T}{2}, \frac{1}{2} \left(\nu\underline{\omega}^2 + \sum_{t=2}^T (\tilde{\sigma}_{q,t} - \rho_q \tilde{\sigma}_{q,t-1})^2 + \tilde{\sigma}_{q,1}^{*~2} \right) \right), \text{i.i.d. across } q.$$

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

$$p(\tilde{\sigma}_{1:T} | \rho_{1:\bar{q}}, \omega_{1:\bar{q}}^2) p(\rho_{1:\bar{q}} | \omega_{1:\bar{q}}^2) \propto \prod_{q=1}^{\bar{q}} \mathcal{N}(\bar{\rho}_q, \bar{V}_q) \mathcal{I}(\rho_q) p(\tilde{\sigma}_1 | \rho_q, \omega_q^2),$$

where $\bar{V}_q = (\bar{v}_\rho^{-1} + \omega_q^{-2} \sum_{t=2}^T \tilde{\sigma}_{q,t-1}^2)^{-1}$ and $\bar{\rho}_q = \bar{V}_q \left(\bar{v}_\rho^{-1} \bar{\rho} + \omega_q^{-2} \sum_{t=2}^T \tilde{\sigma}_{q,t} \tilde{\sigma}_{q,t-1} \right)$. This probability 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), where $\mathcal{N}(\bar{\rho}_q, \bar{V}_q) \mathcal{I}(\rho_q)$ is the proposal density and the acceptance ratio simplifies to $\frac{p(\tilde{\sigma}_1 | \rho_q^*, \omega_q^{2~j})}{p(\tilde{\sigma}_1 | \rho_q^{j-1}, \omega_q^{2~j})}$, with ρ^{j-1} and ρ^* being the draw at the $(j-1)^{\text{th}}$ iteration and the proposed draw, respectively.

3   The DSGE Model

The economy is described by a medium-scale New Keynesian model with price and wage rigidities, capital accumulation, investment adjustment costs, variable capital utilization, and habit formation. The model is based on work of Smets and Wouters (2003), Smets and Wouters (2007), and Christiano et al. (2005). The specific version is taken from Del Negro et al. (2007), except for the monetary policy rule. We further consider an extension in which we add credit frictions to this framework, following the "financial accelerator" model described in Bernanke et al. (1999). The actual implementation of the credit frictions follows closely that of Christiano et al. (2009). For brevity we only present the log-linearized equilibrium conditions and refer the reader to the above referenced papers for the derivation of these conditions from

assumptions on preferences and technologies. All variables that appear subsequently are expressed as log-deviations from this steady state.

Firms. The economy is populated by a continuum of firms that combine capital and labor to produce differentiated intermediate goods. These firms have access to the same Cobb-Douglas production function with capital elasticity α and total factor productivity Z_t. Total factor productivity is assumed to be non-stationary, and its growth rate $z_t = \ln(Z_t/Z_{t-1})$ follows the autoregressive process

$$z_t = (1 - \rho_z)\gamma + \rho_z z_{t-1} + \sigma_z \epsilon_{z,t}. \tag{19}$$

Output, consumption, investment, capital, and the real wage can be detrended by Z_t. In terms of the detrended variables the model has a well-defined steady state.

The intermediate goods producers hire labor and rent capital in competitive markets and face identical real wages, w_t, and rental rates for capital, r_t^k. Cost minimization implies that all firms produce with the same capital-labor ratio

$$k_t - L_t = w_t - r_t^k \tag{20}$$

and have marginal costs

$$mc_t = (1 - \alpha)w_t + \alpha r_t^k. \tag{21}$$

The intermediate goods producers sell their output to perfectly competitive final good producers, which aggregate the inputs according to a CES function. Profit maximization of the final good producers implies the following demand curve

$$y_t(j) - y_t = -\left(1 + \frac{1}{\lambda_f e^{\widetilde{\lambda}_{f,t}}}\right)(p_t(j) - p_t). \tag{22}$$

Here $y_t(j) - y_t$ and $p_t(j) - p_t$ are quantity and price for good j relative to quantity and price of the final good. The price p_t of the final good is determined from a zero-profit condition for the final good producers. We assume that the price elasticity of the intermediate goods is time-varying. Since this price elasticity affects the mark-up that intermediate goods producers can charge over marginal costs, we refer to $\widetilde{\lambda}_{f,t}$ as mark-up shock. Following Calvo (1983), we assume that in every period a fraction of the intermediate goods producers ζ_p is unable to re-optimize their prices.

A fraction ι_p of these firms adjust their prices mechanically according to lagged inflation, while the remaining fraction $1 - \iota_p$ adjusts to steady state inflation π^*. All other firms choose prices to maximize the expected discounted sum of future profits, which leads to the Phillips curve:

$$\pi_t = \frac{\beta}{1 + \iota_p \beta} \mathbb{E}_t[\pi_{t+1}] + \frac{\iota_p}{1 + \iota_p \beta} \pi_{t-1} + \frac{(1 - \zeta_p \beta)(1 - \zeta_p)}{\zeta_p(1 + \iota_p \beta)} mc_t + \frac{1}{\zeta_p} \lambda_{f,t}, \qquad (23)$$

where π_t is inflation and β is the discount rate. Our assumption on the behavior of firms that are unable to re-optimize their prices implies the absence of price dispersion in the steady state. As a consequence, we obtain a log-linearized aggregate production function of the form

$$y_t = (1 - \alpha)L_t + \alpha k_t. \qquad (24)$$

Equations (21), (20), and (24) imply that the labor share lsh_t equals marginal costs in terms of log-deviations: $lsh_t = mc_t$.

There is a representative, competitive, capital producer who produces capital by combining existing capital (purchased from the households at the nominal price Q_t^k) with investment (bought from final goods producers) into new capital and then sell it to the households. Optimal investment satisfies the first-order condition:

$$i_t = \frac{1}{1 + \beta}\left[i_{t-1} - z_t\right] + \frac{\beta}{1 + \beta}\mathbb{E}_t[i_{t+1}] + \frac{1}{(1 + \beta)S'' e^{2\gamma}}q_t^k + \mu_t, \qquad (25)$$

where i_t is investment, and μ_t is a stochastic disturbance to the price of installed capital relative to consumption. Investment in our model is subject to adjustment costs, and S'' denotes the second derivative of the investment adjustment cost function at steady state.

Households. There is a continuum of households with identical preferences, which are separable in consumption, leisure, and real money balances. Households' preferences display (internal) habit formation in consumption, that is, period t utility is a function of $\ln(C_t - hC_{t-1})$, where C_t is the level of consumption. Households supply monopolistically differentiated labor services. These services are aggregated according to a CES function that leads to a demand elasticity $1 + 1/\lambda_w$ (see Equation (22)). The composite labor services are then supplied to the intermediate goods

producers at real wage w_t. To introduce nominal wage rigidity, we assume that in each period a fraction ζ_w of households is unable to re-optimize their wages. A fraction ι_w of these households adjust their $t-1$ nominal wage by $\pi_{t-1}e^\gamma$, where γ represents the average growth rate of the economy, while the remaining fraction $1 - \iota_p$ adjusts to steady state wage growth $\pi^* e^\gamma$. All other households re-optimize their wages. First-order conditions imply that

$$
\begin{aligned}
\tilde{w}_t &= \zeta_w \beta \mathbb{E}_t \Big[\tilde{w}_{t+1} + \Delta w_{t+1} + \pi_{t+1} + z_{t+1} - \iota_w \pi_t \Big] \\
&\quad + \frac{1 - \zeta_w \beta}{1 + \nu_l (1 + \lambda_w)/\lambda_w} \left(\nu_l L_t - w_t - \xi_t + \frac{1}{1 - \zeta_w \beta} \phi_t \right),
\end{aligned}
\tag{26}
$$

where $\tilde{w}_t$ is the optimal real wage relative to the real wage for aggregate labor services, w_t, and ν_l would be the inverse Frisch labor supply elasticity in a model without wage rigidity ($\zeta_w = 0$) and differentiated labor. Moreover, ξ_t denotes the marginal utility of consumption defined below and ϕ_t is a preference shock that affects the intratemporal substitution between consumption and leisure. The real wage paid by intermediate goods producers evolves according to

$$
w_t = w_{t-1} - \pi_t - z_t + \iota_w \pi_{t-1} + \frac{1 - \zeta_w}{\zeta_w} \tilde{w}_t.
\tag{27}
$$

Households are able to insure the idiosyncratic wage adjustment shocks with state contingent claims. As a consequence they all share the same marginal utility of consumption ξ_t, which is given by the expression:

$$
(e^\gamma - h\beta)(e^\gamma - h)\xi_t = -(e^{2\gamma} + \beta h^2)c_t + \beta h e^\gamma \mathbb{E}_t[c_{t+1} + z_{t+1}] + h e^\gamma (c_{t-1} - z_t), \tag{28}
$$

where c_t is consumption. In addition to state-contingent claims households accumulate three types of assets: one-period nominal bonds that yield the return R_t, capital $\bar{k}_t$, and real money balances.[1]

The first order condition with respect to bond holdings delivers the standard Euler equation:

$$
\xi_t = \mathbb{E}_t[\xi_{t+1}] + R_t - \mathbb{E}_t[\pi_{t+1}] - \mathbb{E}_t[z_{t+1}].
\tag{29}
$$

---

[1]Since preferences for real money balances are assumed to be additively separable and monetary policy is conducted through a nominal interest rate feedback rule, money is block exogenous and we will not use the households' money demand equation in our empirical analysis.

Households sell existing capital to capital producers (after depreciation) and purchase new capital from them, which is then rented to the intermediate-goods firms. Capital accumulates according to the law of motion

$$\bar{k}_t = (2 - e^\gamma - \delta)\left[\bar{k}_{t-1} - z_t\right] + (e^\gamma + \delta - 1)\left[i_t + S'' e^{2\gamma}(1 + \beta)\mu_t\right], \qquad (30)$$

where δ is the depreciation rate of capital. The price of capital evolves according to:

$$q_t^k = \beta e^{-\gamma}(1 - \delta)I\!\!E_t\left[q_{t+1}^k\right] + I\!\!E_t\left[(1 - (1-\delta)\beta e^{-\gamma})r_{t+1}^k - (R_t - \pi_{t+1})\right]. \qquad (31)$$

Capital utilization u_t is variable and r_t^k in the previous equation represents the rental rate of effective capital

$$k_t = u_t + \bar{k}_{t-1} - z_t. \qquad (32)$$

The optimal degree of utilization is determined by

$$u_t = \frac{r_*^k}{a''} r_t^k. \qquad (33)$$

Here a'' is the derivative of the per-unit-of-capital cost function $a(u_t)$ evaluated at the steady state utilization rate. The aggregate resource constraint is given by:

$$y_t = (1 + g_*)\left[\frac{c_*}{y_*}c_t + \frac{i_*}{y_*}\left(i_t + \frac{r_*^k}{e^\gamma - 1 + \delta}u_t\right)\right] + g_t, \qquad (34)$$

where c_*/y_* and i_*/y_* are the steady state consumption-output and investment-output ratios, respectively, and $g_*/(1 + g_*)$ corresponds to the government share of aggregate output. The process g_t can be interpreted as exogenous government spending shock. It is assumed that fiscal policy is passive in the sense that the government uses lump-sum taxes to satisfy its period budget constraint. Finally, all stochastic processes described above are assumed to be AR(1) processes with normally distributed errors.

The government. The central bank follows a standard feedback rule:

$$R_t = \rho_R R_{t-1} + (1 - \rho_R)\left(\psi_1\left(\pi_t^{4q} - \pi_t^*\right) + \psi_2 y_t^{4q}\right) + \sigma_r \epsilon_{R,t}, \qquad (35)$$

where R_t and π_t^{4q} represent the interest rate and 4-quarter inflation (in deviations from steady state inflation), respectively, and y_t^{4q} is the 4-quarter growth rate of

output (in deviations from steady state growth rate), π_t^* is a time-varying inflation target that evolves according to the law of motion

$$\pi_t^* = \rho_{\pi^*}\pi_{t-1}^* + \sigma_{\pi^*}\epsilon_{\pi_*,t}, \tag{36}$$

and $\epsilon_{R,t}$ is an i.i.d. shock.

Including financial frictions. We add to this model credit frictions as in Christiano et al. (2009). In order to that we change the ownership of capital from the households to entrepreneurs. In particular, we now consider a continuum of entrepreneurs. Each entrepreneur buys installed capital from the capital producers at the end of period $t-1$ using her own net worth and a loan from the banking sector. In the next period she rents capital out to firms, earning a rental rate per unit of effective capital. In period t she is subject to an i.i.d. (across entrepreneurs and over time) shock that increases or shrinks her capital. In addition, after observing the shock she can choose a level of utilization by paying a cost in terms of general output. At the end of period t the entrepreneurs sells undepreciated capital to the capital producers. Because now we have entrepreneurs holding capital, the evolution of the value of capital in (31) needs to be replaced with

$$\hat{\tilde{R}}_t^k - \pi_t = \frac{r_*^k}{r_*^k + 1 - \delta}\hat{r}_t^k + \frac{1-\delta}{r_*^k + 1 - \delta}\hat{q}_t^k - \hat{q}_{t-1}^k, \tag{37}$$

where $\hat{\tilde{R}}_t^k$ is the entrepreneur's nominal return on capital — which in the baseline model without credit frictions is equal to the risk free nominal interest rate.

Because banks cannot observe the individual realizations of the entrepreneur's shock. However, in case of default by the entrepreneur the bank will pay for monitoring costs and extract a fraction of remaining value of the entrepreneur's assets. Banks pool their loans in such a way that the expected return on the full portfolio of loans equals their opportunity cost — the interest rate paid on deposits from the households, which in turn equals the risk free rate in the economy. The optimal debt contract in this environment implies a spread between the expected return on capital and the risk free rate given by

$$\mathbb{E}_t\left[\hat{\tilde{R}}_{t+1}^k - \hat{R}_t\right] = \zeta_{sp,b}\left(\hat{q}_t^k + \hat{\bar{k}}_t - \hat{n}_t\right) + \tilde{\sigma}_{\omega,t}, \tag{38}$$

where $\hat{n}_t$ is entrepreneur's net-worth, $\left(\hat{q}_t^k + \hat{\bar{k}}_t - \hat{n}_t\right)$ is a measure of entrepreneur's leverage, and $\tilde{\sigma}_{\omega,t}$ is a shock to the volatility of the idiosyncratic shock to each entrepreneur (the higher this parameter the higher the expected fraction of entrepreneurs that default, everything else constant). We label it as spread shock and assume that it follows a stationary AR(1) process.

A fraction $1 - \gamma_*$ of entrepreneurs exits the economy and fraction γ_* survives to continue operating for another period. A fraction of the total net worth owned by exiting entrepreneurs is consumed upon exit and the remaining fraction of their networth is transferred as a lump sum to the households. Each period new entrepreneurs enter and receive a net worth transfer. Because this transfer is small, this exit and entry process ensures that entrepreneurs do not accumulate enough net worth to escape the financial frictions. Aggregate entrepreneurs' net worth evolves accordingly as:

$$
\begin{aligned}
\hat{n}_t &= \zeta_{n,\tilde{R}^k}\left(\hat{\tilde{R}}_t^k - \pi_t\right) - \zeta_{n,R}\left(\hat{R}_{t-1} - \pi_t\right) + \zeta_{n,qK}\left(\hat{q}_{t-1}^k + \hat{\bar{k}}_{t-1}\right) + \zeta_{n,n}\hat{n}_{t-1} \\
&\quad -\gamma_*\frac{v_*}{n_*}\hat{z}_t - \zeta_{n,\tilde{\sigma}_\omega}\tilde{\sigma}_{\omega,t-1},
\end{aligned} \tag{39}
$$

where the coefficients $(\zeta_{n,\tilde{R}^k}, \zeta_{n,R}, \zeta_{n,n}, \zeta_{n,\tilde{\sigma}_\omega})$ are all functions of standard parameters in the baseline model and $(\gamma_*, \bar{F}, \zeta_{sp,b})$, with $\bar{F}$ denoting the steady state fraction of entrepreneurs that default. We calibrate the first two of these non-standard parameters and estimate the latter one, which controls how important is the endogenous relationship between the real economy and the degree of financial frictions, expressed in terms of the spread between the return on capital and the risk free rate in the economy.

State-Space Representation of the DSGE Model. We use the method in Sims (2002) to solve the log-linear approximation of the DSGE model. 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:

$$
s_t = \mathcal{T}(\theta)s_{t-1} + \mathcal{R}(\theta)\epsilon_t, \tag{40}
$$

which summarizes the evolution of the states s_t, and of the measurement equations:

$$y_t = \mathcal{Z}(\theta)s_t + \mathcal{D}(\theta), \tag{41}$$

which maps the states onto the vector of observables y_t, where $\mathcal{D}(\theta)$ represents the vector of steady state values for these observables. Specifically, for our standard set of macro time series the set of measurement equations is:

$$
\begin{aligned}
&\text{Real output growth (\%, annualized)}\\
&400(\ln RGDP_t - \ln RGDP_{t-1}) &=&\quad 400(y_t - y_{t-1} + z_t) + \gamma\\[4pt]
&\text{Hours (\%)}\\
&100\ln L_t &=&\quad 100(L_t + \ln L^{adj})\\[4pt]
&\text{Labor Share (\%)}\\
&100\ln LSH_t &=&\quad 100(L_t + w_t - \widehat{y}_t + \ln lsh_*)\\[4pt]
&\text{Inflation (\%,annualized)}\\
&400(\ln P_t - \ln P_{t-1}) &=&\quad 400\pi_t + \pi^*\\[4pt]
&\text{Interest Rates (\%,annualized)}\\
&400\ln R_t &=&\quad 400R_t + R_*,
\end{aligned}
\tag{42}
$$

where $RGDP_t$, L_t, LSH_t, P_t, and R_t represent real per capita GDP, total per capita hours, the labor share, the price level, and the interest rate, respectively.[2] The quantities γ, π^*, and R_* are the annualized (in percent) steady state real output growth, inflation, and nominal interest rate, respectively, lsh_* is the steady states of the labor share and the parameter L^{adj} captures the units of measured hours (it can be viewed as a re-parameterization of the steady state associated with the time-varying preference parameter ϕ_t that appears in the households' utility function). As a measure of inflation we use the core personal consumption expenditure (PCE)

---

[2]Relative to Smets and Wouters (2007) we use the labor share as opposed to growth in wages, both because it contains level information relative to wages, which among other things allows us to estimate the capital share, and because it provides a measure of real marginal costs, a key variable for inflation in New Keynesian models. We also choose to exclude consumption and investment. Del Negro et al. (2007) find that the non-stationarity of the great ratios involving consumption and investment is one of the main sources of misspecification of DSGE models. Such misspecification may in principle "pollute" the outcome of the model comparison done in this paper. Since consumption and investment as observables do not necessarily play a key role in the exercise at hand we therefore chose to exclude them. Del Negro and Schorfheide (2008), a paper about model comparison, take the same route.

price index. Appendix A.1 provides further details on the data. In our benchmark specification we use 97 quarters of data spanning the post-Volcker disinflation period: 1984Q2 to 2008Q2.

Whenever we consider the extended model with credit frictions, the set of equations (42) is augmented to include:

$$SP_t = SP_* + 400 I\!\!E_t \left[\hat{\tilde{R}}^k_{t+1} - \hat{R}_t \right],$$ (43)

where SP_t is the annualized spread between the Baa bond rate yield and the 10-year Treasury yield, and SP_* the annualized steady state level of the spread.

4 Results

4.1 Prior and Posterior for the DSGE Parameters

Table 1 shows the priors for the DSGE model parameters. The top panel shows the prior for the the nominal rigidities parameters, d_p and d_w. We depart from the literature and form a prior in terms of the expected duration of no-change spells, as opposed to the standard parameterization in terms of ζ_p and ζ_w, the probability of not being able to adjust prices (or wages). We do that because when the usual prior, which is expressed in terms of a Beta distribution with a domain in the $(0, 1)$ interval, is either very informative or allows for very long durations that are at odds with the data. This is because of the nonlinearity of the mapping between ζ and d, which is $d = \frac{1}{1-\zeta}$. As ζ approaches one d goes to infinity. Therefore a relatively flat prior on the $(0, 1)$ interval, such as the one used by Smets and Wouters (2007), puts non negligible mass on very high durations. We choose a Gamma prior on d_p and d_w with mean 3 and standard deviation 1, which at least for prices is broadly consistent with the empirical micro-level literature.

The next panel shows the prior for the the policy rule parameters, namely the responses of interest rates to inflation (ψ_1) and economic activity (ψ_2) – 4-quarter

output growth in the baseline specification – in the policy rule, persistence (ρ_r), and steady state inflation target (π^*). The prior on π^* is centered using pre-sample information on inflation, as in Del Negro and Schorfheide (2008). The prior on ψ_1 and ψ_2 are centered at 2 and .2 respectively, and imply a fairly strong response to inflation and a moderate response to output. Priors on variance of i.i.d. policy shocks σ_r is centered at .15. In general the priors on the standard deviations of the shocks are chosen so that overall variance of endogenous variables is roughly close to that observed in the pre-sample 1959Q3-1984Q1, informally following the approach in Del Negro and Schorfheide (2008).

Priors on remaining parameters are shown in the bottom panel of Table 1. The priors on "Endogenous Propagation and Steady State" are all chosen as in Del Negro and Schorfheide (2008). Specifically, the prior for the habit persistence parameter h is centered at 0.7, which is the value used by Boldrin et al. (2001). The prior for a'' implies that in response to a 1% increase in the return to capital, utilization rates rise by 0.1 to 0.3%. These numbers are considerably smaller than the one used by Christiano et al. (2005). For the model with credit frictions the key new parameters are the elasticity of the spread with respect to the leverage, $\zeta_{sp,b}$, the survival rate for the entrepreneurs, γ_*, and the steady state default rate, $\bar{F}$. The last two are calibrated while the former is estimated. Following Gilchrist et al. (2009), we set γ_* to 0.99. We set $\bar{F}$ to imply an annualized default rate of 3%, as in Bernanke et al. (1999). The prior for the spread elasticity, $\zeta_{sp,b}$, is a beta distribution with mean 0.05 (as in Bernanke et al. (1999)) and standard deviation of 0.02. The steady state spread has a Gaussian prior with mean 2 and standard deviation of 0.5, in annual percentage terms.

We again depart from the literature and form a prior on the persistence of shocks in terms of the half life of the shock h, as opposed to the standard autocorrelation coefficient ρ. We do that for three reasons. First, half-lives are in our view more easily interpretable than autocorrelation coefficients. Second, the usual prior, which is expressed in terms of a Beta distribution with a domain in the $(0, 1)$ interval,

is either very informative or allows for very long half-livesthat are at odds with the data. Specifically, a relatively flat prior on the $(0, 1)$ interval, such as the one used by Smets and Wouters (2007), puts non negligible mass on extremely high half-lives. Third, from a computational perspective if the shock is very persistent the estimated value of ρ is close to the boundary, and that creates computational issues in the MCMC procedure used to obtain draws from the posterior. No such boundary problems arise when the parameter is expressed in terms of the half-life of the shock. We choose a Gamma prior with mean 6 and standard deviation of 5.8. The prior for the standard deviations of the shocks is quite flat, and the mean is chosen so to produce standard deviations of the observables that are not at odds with those observed in the data.

Table 2 presents the posterior mean for the standard parameters of the DSGE model for the different specifications with Gaussian shocks, Student-t distributed shocks, stochastic volatility, and both.

4.2 Evidence against Gaussianity

In this section we offer some evidence in support of the hypothesis of fat tails for the shocks. It is generally agreed that financial crises are not very common but when they do take place they can be very disruptive with significant increases in interest rate spreads and subsequent disruptions in economic activity. Therefore we use the model with credit frictions to extract the smoothed shocks over the sample period under the assumption of Gausianity, and search for evidence of whether that assumption is validated. Figure 1 shows the absolute value of the smoothed shocks over the sample period (with solid line representing the median, dotted lines the 90% bands, and the grey vertical shades marking NBER recession periods).

In the top panel we show this for shocks to entrepreneurial risk. It is remarkable that there are nine instances in the sample in which the absolute deviation is higher than two standard deviations, with one instance reaching as high as nearly seven

standard deviations. Furthermore the mean of the excess kurtosis across draws is 13, which would be equivalent to a Student-t distribution with 4.5 degrees of freedom — highly suggestive of fat tails and far from what the Gaussian assumption. Confirming the idea that big financial shocks are associated with big disruptions in the economic activity, the biggest deviations coincide with NBER recessions.

The question is then whether this is a characteristic intrinsic to financial shocks, or whether it also applies to the innovations to the other exogenous disturbances in the model. We thus show in the bottom panel of figure 1 the absolute value of the smoothed shocks to the interest rate rule, which we label as policy shocks. Once more there are several instances with deviations far in excess of two standard deviations, even if fewer than in the previous case. The mean excess kurtosis is 4.2, which would be equivalent to a Student-t distribution with 5.4 degrees of freedom.

Interestingly in the case of policy shocks the biggest deviations occur in periods in which there are an abnormal number of large deviations, with larger shocks in the late 70s and early 80s. This suggests a possible trend in the sequence of shocks, which may be due to slow-moving time-varying volatility as opposed to rare but large shocks — in the case of entrepreneurial risk shocks the large deviations appear to be more sporadic, hence less prone to a trend. In the next section we discuss in depth the role of Student-t distributed shocks vis-a-vis slow-moving stochastic volatility.

4.3   Fat Tails and Stochastic Volatility

In the previous section we showed some evidence of rare large shocks and time varying volatility based on historical shocks extracted from standard gaussian estimation. In this section we conduct a horse-race between models estimated with Student-t distributed shocks and/or stochastic volatility. Table 4 shows the log marginal likelihood for the different shock assumptions for both the baseline (Panel A) and financial frictions (Panel B). For each model we consider specifications with and without stochastic volatility (time-varying $\sigma_{q,t}$) and with different prior means

for the degrees of freedom of the Student-t distributed component of the shocks (with $\lambda_q = \infty$ being the gaussian case). [3]

The main result to take out of this table is that in the financial frictions model the data strongly favors the assumption of Student-t distributed shocks vis-a-vis the Gaussian assumption. Furthermore, whether we complement that with stochastic volatility or not does not change the fit substantially. In the Baseline model there is strong evidence in favor of Student-t distributed shocks with and without stochastic volatility. In this case the fit is improved if we have both rather than only one of these features.

In the financial frictions model the fit improves as we reduce the prior mean on the degrees of freedom for the Student-t distributed component of the shocks and this leads to a gain of 170 log-points in the case without stochastic volatility and about 55 log-points in the case with stochastic volatility. In this case the fit is better with Student-t distributed shocks alone, rather than combined with stochastic volatility, which may be related to the significant uncertainty in the estimation of the stochastic volatility component for several of the shocks.

In order to understand these results, Table 5 shows the posterior mean of the degrees of freedom for the Student-t distributed component of each shock in the specifications with and without stochastic volatility.[4] In the model with financial frictions (Panel B) we get relatively low posterior means for the degrees of freedom, especially if we do not consider stochastic volatility. In this case we get several shocks with very low degree of freedom and none deviates significantly from the prior mean on the upside. Even if we consider stochastic volatility in the shocks, we still have a similar pattern, with the highest posterior mean at 7.5 for the leisure shock.

---

[3]The reason why we consider different prior means is that the scope for the degrees of freedom is very large and the prior mean choice does affect the posterior shape to some extent.

[4]In this table we show only the results with the prior mean for each model that leads to the best fit for each model: prior mean of 6 degrees of freedom for the model with financial frictions and prior mean of 9 degrees of freedom for the baseline model.

Importantly the shocks to the spread have very fat tails even when we allow for stochastic volatility (posterior mean of 2.9). This confirms the idea that shocks originating in the financial sector are not very frequent but can nevertheless be very large. On the other hand, the monetary policy shocks are estimated with fat tails in the absence of stochastic volatility but less so when we allow for the two to coexist — the posterior mean of the degrees of freedom of the Student-t distributed component change from 2.5 to 7.2.

It is intuitive to expect some reduction in the degree of fat tails in the shocks once we also consider stochastic volatility. However, it is less obvious how to interpret the result that once we add Student-t distributed shocks the stochastic volatility no longer seems to further improve the fit of the model. In order to understand this we show in Figure **??** two panels with the absolute value of the smoothed shock history for the spread shock, for the estimation with Student-t distributed shocks (but no stochastic volatility). The first panel shows the shock itself (in absolute value), $|\varepsilon_t|$, while the second panel shows the absolute value of $|\eta_t| = |\tilde{h}_t^{1/2}\varepsilon_t|$, which we refer to as the "tamed" shocks. This latter one is the residual shock once the Student-t distributed component is removed. In the first panel there is strong evidence in favor some form of time varying volatility, whether rare shocks of stochastic volatility. Once we remove the rare shocks' component the residual no longer exhibits that much evidence of time variation in the shock size. This may suggest that once Student-t distributed shocks are considered, there is not too much a of a role for stochastic volatility.

The fact that rare shocks appear to expunge stochastic volatility from its role in explaining spread shocks does not mean that it applies to all types of shocks. Figure **??** also shows a similar comparison but for the shock to the monetary policy rule. In this case the Student-t distributed shock component surely reduced the degree of time variation in the innovations to monetary policy, but the residual still exhibit fairly strong evidence of slow moving stochastic volatility with higher volatility in the 1970s and early 1980s and lower afterwards.

Given the marginal likelihood results shown in Table 4 it looks like there are

more shock types similar to the spread shock than to the monetary policy shock, so that all in all the data does not see enough evidence of slow moving stochastic volatility on top of rare shocks to justify adding that component to all shocks. (One may wonder that a better specification would be to consider stochastic volatility only for some shocks but not for all of them.)

So far we showed some evidence, based on the Student-t distributed shock estimation that once we consider rare shocks, the (slow moving) stochastic volatility component for some shocks is no longer obviously useful in fitting the data. What is left to check is how much does that component actually change if we consider first the stochastic volatility estimation and then the stochastic volatility with rare large shocks estimation. Figures 3 present the results for the spread and monetary policy shocks, respectively. In each case we compare the cases with and without Student-t distributed shocks.

In the case of the spread shock, shown in Figure 3, allowing for a Student-t distributed component in the shock strips the stochastic volatility component of some of the high frequency movements. Furthermore notice that there is substantial uncertainty around the median, suggesting that once we account for rare large shocks there is scant evidence of a stochastic volatility component.

In the case of the monetary policy shock, the two panels are more similar. The only difference is the scale, with the presence of Student-t distributed shocks reducing the scale of the peak in volatility in the 1970s, suggesting that this period had both large shocks and persistent shocks as well. The fact that the stochastic volatility retains its pattern is consistent with our previous conjecture that for monetary policy shocks there is some evidence that we need stochastic volatility.

For other shocks the results are more mixed, but in general the stochastic volatility component is estimated with significant uncertainty, which may suggest why the overall fit — as measured by the marginal likelihood — is higher in the case with Student-t distributed shocks and no stochastic volatility.

4.4 Do Fat Tails Matter for the Macroeconomy?

We have so far run a horse race between Student-t distributed shocks and stochastic volatility with the former coming on top in terms of fit for this particular model. Now the question is how important is this for the description of the data. In order to get to this we first notice that the evidence in favor of rare large shocks is especially strong for the financial frictions model and the the spread shock in particular shows very strong evidence of rare large shocks. Therefore the obvious next step is to judge how important is this particular shock for the economy.

Figure 4 shows the contribution of the spread shocks to the history of GDP growth and hours worked. The black line shows the evolution of these variables in deviations from the steady state, and the vertical bars show in each period the median contribution of spread shocks to that variable at that point in time (whether the shock originated in that period or before). In other words, it is the counterfactual evolution of these variables if we were to shut down all other shocks. What we observe from this figure is that the spread shock explains a large part of the recent contraction throughout the Great Recession, which is widely attributed to financial conditions disruptions. In the rest of the sample there is less evidence of its role. However, there is some non-negligible contribution of this shock to the contractions of these variables in both the 1974 and 1981.

Another way to evaluate how important are Student-t distributed shocks is to run another counterfactual experiment. In this case we extract the shocks hitting the economy and then eliminate the Student-t distributed component from them so that we get the evolution of these variables in the absence of rare large shocks. Figure 5 shows this counterfactual for the same two variables, GDP growth and hours worked. The black line is the actual data and the pink line is the counterfactual evolution of the variables in the absence of rare large shocks.

In the case of GDP growth (top panel), the two lines are fairly similar. However, the peaks and troughs (especially the latter) are much more tame in the absence of the Student-t distributed component, especially obvious in the three contractions in

which the spread is especially important: 1974, 1981 and 2008-2009. In particular, in the absence of the Student-t distributed component, the "great recession" would not be that much worse than the 1991 recession.

When it comes to hours worked, the Student-t distributed component explains a large share of the fluctuations, especially in explaining the downturns. This is very clear for the period from 1975-1985 and then from 2001-2009. The first was the period of large shocks to the monetary policy rule and the latter the financial crisis, both of which the Student-t distributed shocks capture fairly well.

5 Conclusions

To be written

References

An, Sungbae and Frank Schorfheide, "Bayesian Analysis of DSGE Models–Rejoinder," *Econometric Reviews*, 2007, *26*, 211 – 219.

Bernanke, Ben S., Mark Gertler, and Simon Gilchrist, "The Financial Accelerator in a Quantitative Business Cycle Framework," in John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*, Vol. 1C, Amsterdam: North-Holland, 1999, chapter 21, pp. 1341–93.

Boldrin, Michele, Lawrence Christiano, and Jonas Fisher, "Habit Persistence, Asset Returns, and the Business Cycle," *American Economic Review*, 2001, *91*, 149–166.

Calvo, Guillermo, "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, 1983, *12*, 383–398.

Chib, Siddhartha and Edward Greenberg, "Bayes inferences in regression models with ARMA(p,q) errors," *Journal of Econometrics*, 1994, *64*, 183–206.

_ **and Srikanth Ramamurthy**, "DSGE Models with Student-t errors," *mimeo, Washington University in St. Louis*, 2011.

Christiano, Lawrence J., "Comment on 'On the Fit of New Keynesian Models' by Del Negro ,Schorfheide, Smets ,Wouters," *Journal of Business and Economic Statistics*, 2007, *25* (2), 143–151.

Christiano, Lawrence, Martin Eichenbaum, and Charles Evans, "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 2005, *113*, 1–45.

_ **, Roberto Motto, and Massimo Rostagno**, "Financial Factors in Economic Fluctuations," *Unpublished*, May 2009.

Del Negro, Marco and Frank Schorfheide, "Forming priors for DSGE models (and how it affects the assessment of nominal rigidities)," *Journal of Monetary Economics*, 2008, *55* (7), 1191 – 1208.

_ , _ , **Frank Smets, and Raphael Wouters**, "On the Fit of New Keynesian Models," *Journal of Business and Economic Statistics*, 2007, *25* (2), 123 – 162.

Durbin, James and Siem Jan Koopman, "A Simple and Efficient Simulation Smoother for State Space Time Series Analysis," *Biometrika*, 2002, *89* (3), 603–616.

Fernández-Villaverde, Jesús and Juan F. Rubio-Ramírez, "Estimating Macroeconomic Models: A Likelihood Approach," *Review of Economic Studies*, 2007, *74* (4), 1059–1087.

Geweke, John, "Bayesian Treatment of the Independent Student-t Linear Model," *Journal of Applied Econometrics,*, 1993, *8*, S19–S40.

_ , "Priors for Macroeconomic Time Series and Their Application," *Econometric Theory*, 1994, *10*, 609–632.

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

_ , *Contemporary Bayesian Econometrics and Statistics*, Wiley, 2005.

Gilchrist, Simon, Alberto Ortiz, and Egon Zakrajsek, "Credit Risk and the Macroeconomy: Evidence from an Estimated DSGE Model," *Unpublished*, July 2009.

Justiniano, Alejandro and Giorgio Primiceri, "The Time-Varying Volatility of Macroeconomic Fluctuations," *American Economic Review*, 2008, *98* (3), 604 – 641.

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

Müller, Ulrich K., "Risk of Bayesian Inference in Misspecified Models, and the Sandwich Covariance Matrix," *Mimeo, Princeton University*, 2011.

Sims, Christopher A., "Solving Rational Expectations Models," *Computational Economics*, 2002, *20*, 1–20.

Smets, Frank and Raf Wouters, "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, 2003, *1* (5), 1123 – 1175.

_ **and** _ , "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 2007, *97* (3), 586 – 606.

A Appendix

A.1 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 (42). Real output is obtained by dividing the nominal series (*GDP*) by population 16 years and older (*LN16N*), and deflating using the chained-price GDP deflator (*JGDP*). We compute quarter-to-quarter output growth as log difference of real GDP per capita and multiply the growth rates by 400 to convert them into annualized percentages. Our measure of hours worked is computed by taking total hours worked reported in the National Income and Product Accounts (NIPA), which is at annual frequency, and interpolating it using growth rates computed from hours of all persons in the non-farm business sector (*LXNFH*). We divide hours worked by *LN16N* to convert them into per capita terms. We then take the log of the series multiplied by 100 so that all figures can be interpreted as percentage changes in hours worked. The labor share is computed by dividing total compensation of employees (*YCOMP*) obtained from the NIPA by nominal GDP. We then take the log of the labor share multiplied by 100. Inflation rates are defined as log differences of the personal consumption expenditure (PCE) core price index (*JCXFE*) and converted into annualized percentages. The nominal rate corresponds to the effective Federal Funds Rate (*FFED*), also in percent. The spread is extracted by taking the difference between the Baa bond yield (*FBAA*) and the 10-year treasury yield (*FCM10*), annualized and in percentages.

A.2 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:

$$
\begin{aligned}
p(y_{1:T}) &= \int p(y_{1:T}|s_{1:T},\theta)p(s_{1:T}|\varepsilon_{1:T},\theta)p(\varepsilon_{1:T}|\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\theta) \\
&\qquad p(\tilde{h}_{1:T}|\lambda_{1:\bar{q}})p(\tilde{\sigma}_{1:T}|\omega_{1:\bar{q}}^2)p(\lambda_{1:\bar{q}})p(\omega_{1:\bar{q}}^2)p(\theta) \\
&\qquad\qquad d(s_{1:T},\varepsilon_{1:T},\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\lambda_{1:\bar{q}},\rho_{1:\bar{q}},\omega_{1:\bar{q}}^2,\theta), \\
&= \int p(y_{1:T}|\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\theta)p(\tilde{h}_{1:T}|\lambda_{1:\bar{q}})p(\tilde{\sigma}_{1:T}|\omega_{1:\bar{q}}^2) \\
&\qquad p(\lambda_{1:\bar{q}})p(\omega_{1:\bar{q}}^2)p(\theta)d(\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\lambda_{1:\bar{q}},\rho_{1:\bar{q}},\omega_{1:\bar{q}}^2,\theta)
\end{aligned}
\tag{44}
$$

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 $f(\theta,\tilde{h}_{1:T},\tilde{\sigma}_{1:T},\lambda_{1:\bar{q}},\rho_{1:\bar{q}},\omega_{1:\bar{q}}^2)$ is any distribution with support contained in the support of the posterior density such that

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

it follows from the definition of the posterior density that:

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

We follow Justiniano and Primiceri (2008) in choosing

$$
f(\theta,\tilde{h}_{1:T}) = f(\theta) \cdot p(\tilde{h}_{1:T}|\lambda_{1:\bar{q}})p(\tilde{\sigma}_{1:T}|\omega_{1:\bar{q}}^2)p(\lambda_{1:\bar{q}})p(\omega_{1:\bar{q}}^2),
\tag{45}
$$

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

$$
\hat{p}(y_{1:T}) = \left[\frac{1}{n_{sim}} \sum_{j=1}^{n_{sim}} \frac{f(\theta^j)}{p(y_{1:T}|\tilde{h}_{1:T}^j,\tilde{\sigma}_{1:T}^j,\theta^j)p(\theta^j)} \right]^{-1}
\tag{46}
$$

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 (45), 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 Justiniano and Primiceri (2008) we found that this method delivers very similar results across different chains.

A.3 Drawing the stochastic volatilities

We draw the stochastic volatilities using the procedure in Kim et al. (1998), which we briefly describe. Taking squares and then logs of (3) one obtains:

$$\varepsilon_{q,t}^* = 2\tilde{\sigma}_{q,t} + \eta_{q,t}^* \tag{47}$$

where

$$\varepsilon_{q,t}^* = \log(\sigma_q^{-2}\tilde{h}_{q,t}\varepsilon_{q,t}^2 + c), \tag{48}$$

$c = .001$ being an offset constant, and $\eta_{q,t}^* = \log(\eta_{q,t}^2)$. If $\eta_{q,t}^*$ were normally distributed, $\sigma_{q,1:T}$ could be drawn using standard methods for state-space systems. In fact, $\eta_{q,t}^*$ is distributed as a $\log(\chi_1^2)$. Kim et al. (1998) address this problem by approximating the $\log(\chi_1^2)$ with a mixture of normals, that is, expressing the distribution of $\eta_{q,t}^*$ as:

$$p(\eta_{q,t}^*) = \sum_{k=1}^{K} \pi_k^* \mathcal{N}(m_k^* - 1.2704, \nu_k^{*\,2}) \tag{49}$$

The parameters that optimize this approximation, namely $\{\pi_k^*, m_k^*, \nu_k^*\}_{k=1}^{K}$ and K, are given in Kim et al. (1998). Note that these parameters are independent of the specific application. The mixture of normals can be equivalently expressed as:

$$\eta_{q,t}^*|\varsigma_{q,t} = k \sim \mathcal{N}(m_k^* - 1.2704, \nu_k^{*\,2}), \ Pr(s_{i,t} = k) = \pi_k^*. \tag{50}$$

Hence step (4) of the Gibbs sampler actually consists in two steps:

(4.1) Draw from $p(\varsigma_{1:T}|\tilde{\sigma}_{1:T}, \varepsilon_{1:T}, \tilde{h}_{1:T}, s_{1:T}\lambda_{1:\bar{q}}, \rho_{1:\bar{q}}, \omega_{1:\bar{q}}^2, y_{1:T})$ using (49) for each q. Specifically:

$$Pr\{\varsigma_{q,t} = k|\tilde{\sigma}_{1:T}, \varepsilon_{1:T}, \tilde{h}_{1:T} \dots\} \propto \pi_k^* \nu_k^{*-1} \exp\left[-\frac{1}{2\nu_k^{*\,2}}(\eta_{q,t}^* - m_k^* + 1.2704)^2\right]. \tag{51}$$

where from (47) $\eta_{q,t}^* = \varepsilon_{q,t}^* - 2\tilde{\sigma}_{q,t}$.

(4.2) Draw from $p(\tilde{\sigma}_{1:T}|\varsigma_{1:T}, \varepsilon_{1:T}, \tilde{h}_{1:T}, s_{1:T}\lambda_{1:\bar{q}}, \rho_{1:\bar{q}}, \omega^2_{1:\bar{q}}, y_{1:T})$ using Durbin and Koopman (2002), where (47) is the measurement equation and (8) is the transition equation.

Note that in principle we should make it explicit that we condition on $\varsigma_{1:T}$ in the other steps of the Gibbs sampler as well. In practice, all other conditional distributions do not depend on $\varsigma_{1:T}$, hence we omit the term for simplicity.

Table 1: Priors on DSGE Model Parameters

Parameter	Domain	Density	Para (1)	Para (2)	5%	95%
Nominal Rigidities						
d_p	$[0,1)$	Gamma	3.000	1.000	1.390	4.549
d_w	$[0,1)$	Gamma	3.000	1.000	1.418	4.548
Policy						
ψ_1	$\mathbb{R}^+$	Gamma	2.00	0.25	1.590	2.407
ψ_2	$\mathbb{R}^+$	Gamma	0.20	0.10	0.046	0.346
ρ_r	$[0,1)$	Beta	0.50	0.200	0.170	0.827
π^*	$\mathbb{R}$	Normal	2.004	0.991	0.370	3.646
σ_r	$\mathbb{R}^+$	InvGamma	0.150	4.000	0.080	0.298
h_{π_*}	$\mathbb{R}^+$	Gamma	40.00	5.000	31.630	47.954
σ_{π_*}	$\mathbb{R}^+$	InvGamma	0.03	7.000	0.019	0.048
Endogenous Propagation and Steady State						
α	$[0,1)$	Beta	0.330	0.020	0.297	0.363
s''	$\mathbb{R}^+$	Gamma	4	1.500	1.592	6.304
h	$[0,1)$	Beta	0.700	0.150	0.471	0.946
a'	$\mathbb{R}^+$	Gamma	0.200	0.100	0.045	0.346
ν_l	$\mathbb{R}^+$	Gamma	1.000	0.250	0.591	1.395
ζ_{sp}	$[0,1)$	Beta	0.051	0.019	0.023	0.082
r^*	$\mathbb{R}^+$	Gamma	1.5	1	0.106	2.883
γ	$\mathbb{R}^+$	Gamma	2.000	0.750	0.822	3.162
g^*	$\mathbb{R}^+$	Gamma	0.300	0.100	0.143	0.459
spr_*	$\mathbb{R}^+$	Gamma	2.000	0.249	1.579	2.394
Exogenous Processes						
h_z	$\mathbb{R}^+$	Gamma	6.000	5.800	0.100	13.676
h_ϕ	$\mathbb{R}^+$	Gamma	6.000	5.800	0.100	13.676
h_{λ_f}	$\mathbb{R}^+$	Gamma	6.000	5.800	0.100	13.676
h_μ	$\mathbb{R}^+$	Gamma	6.000	5.800	0.100	13.676
h_g	$\mathbb{R}^+$	Gamma	6.000	5.800	0.100	13.676
h_{σ_w}	$\mathbb{R}^+$	Gamma	6.000	5.800	0.100	13.676
σ_z	$\mathbb{R}^+$	InvGamma	0.300	4.000	0.158	0.592
σ_ϕ	$\mathbb{R}^+$	InvGamma	8.000	4.000	4.280	15.775
σ_{λ_f}	$\mathbb{R}^+$	InvGamma	0.200	4.000	0.105	0.395
σ_μ	$\mathbb{R}^+$	InvGamma	0.750	4.000	0.399	1.481
σ_g	$\mathbb{R}^+$	InvGamma	0.500	4.000	0.271	0.999
σ_{σ_w}	$\mathbb{R}^+$	Gamma	.050	4.000	0.026	0.099

Notes: Para (1) and Para (2) correspond to means and standard deviations for the Beta, Gamma, and Normal distributions and to s and ν for the Inverse Gamma distribution, where $p_{\mathcal{IG}}(\sigma|\nu,s) \propto \sigma^{-\nu-1}e^{-\nu s^2/2\sigma^2}$. The last two columns report the 5th and 95th quintile of the prior distribution.

Table 2: Posterior Means of the DSGE Model Parameters

Parameter	Prior Mean	Std.Dev.	Posterior Mean Gaussian	TD	SV	SV+TD
α	0.330	0.020	0.335	0.337	0.336	0.337
d_p	3.000	1.000	7.393	7.387	7.360	7.520
S''	4.000	1.500	3.157	2.971	3.554	2.738
h	0.700	0.150	0.697	0.778	0.732	0.728
a''	0.200	0.100	0.300	0.328	0.292	0.319
ν_l	1.000	0.250	1.917	2.023	2.074	2.192
d_w	3.000	1.000	11.065	9.568	10.583	9.617
r^*	1.750	0.250	1.303	1.394	1.442	1.349
ψ_1	2.000	0.250	2.012	2.641	2.432	2.357
ψ_2	0.200	0.100	0.170	0.159	0.166	0.170
ρ_r	0.500	0.200	0.599	0.642	0.602	0.621
π^*	2.000	1.000	4.304	4.190	4.175	4.247
spr_*	2.000	0.250	1.465	1.204	1.629	1.266
ζ_{sp}	0.050	0.020	0.063	0.066	0.042	0.067
γ	2.000	0.750	1.298	1.410	1.741	1.431
g^*	0.300	0.020	0.298	0.299	0.297	0.297
L^{adj}	253.500	5.000	232.135	228.346	236.055	228.186
h_z	6.000	5.800	1.091	1.274	1.148	1.159
h_ϕ	6.000	5.800	0.600	0.583	0.591	0.560
h_{π_*}	40.000	5.000	41.063	39.402	41.089	40.676
h_{λ_f}	6.000	5.800	0.739	0.754	0.818	0.644
h_μ	6.000	5.800	27.284	19.329	21.013	21.010
h_g	6.000	5.800	16.126	17.008	15.126	16.971
h_{sigw}	6.000	5.800	28.526	28.480	21.206	26.445
σ_z	0.300	4.000	1.053	0.903	1.060	0.846
σ_ϕ	8.000	4.000	3.924	3.679	3.584	3.428
σ_{π_*}	0.030	7.000	0.046	0.053	0.058	0.048
σ_{λ_f}	0.200	4.000	0.099	0.089	0.091	0.087
σ_μ	0.750	4.000	0.351	0.346	0.385	0.336
σ_g	0.500	4.000	0.250	0.211	0.236	0.211
σ_r	0.200	4.000	0.259	0.115	0.077	0.078
σ_{sigw}	0.050	4.000	0.081	0.043	0.089	0.035

Notes: We consider a prior mean of 6 degrees of freedom for the Student-t distributed component. The stochastic volatility component assumes a prior mean for the size of the shocks to volatility of $(0.01)^2$.

Table 3: Marginal Likelihoods

	Baseline	Financial Frictions
	Post-84 sample	*Post-84 sample, with Spreads*
Gaussian Shocks		
	-681.78	-716.26
Student-t distributed Shocks		
$\lambda = 15$	-679.00	-690.29
$\lambda = 9$	-676.43	-683.02
$\lambda = 6$	-678.49	-677.45

Notes: The parameter λ represents the degrees of freedom in the Student-t distribution. The marginal likelihoods of the *Baseline* and *Financial Frictions* models are not comparable, as the set of observables is different.

Table 4: Marginal Likelihoods

Panel A: Baseline Model

	No Stoch. Vol.	S.Vol. ($\underline{\omega}^2 = (0.01)^2$)
Gaussian	-1503.944	-1423.765
Student-t (λ_q prior mean 9)	-1428.752	-1404.383
Student-t (λ_q prior mean 6)	-1424.361	-1411.316

Panel B: Financial Frictions Model

	No Stoch. Vol.	S.Vol. ($\underline{\omega}^2 = (0.01)^2$)	S.Vol. ($\underline{\omega}^2 = (0.06)^2$)
Gaussian	-1586.542	-1531.557	-1505.186
Student-t (λ_q prior mean 9)	-1422.333	-1446.649	-1456.306
Student-t (λ_q prior mean 6)	-1416.991	-1427.833	-1449.519

Notes: The parameter λ represents the degrees of freedom in the Student-t distribution. The marginal likelihoods of the *Baseline* and *Financial Frictions* models are not comparable, as the set of observables is different.

Table 5: Student's t Degrees of Freedom Posterior Mean

Panel A: Baseline Model

	z	ϕ	λ_f	μ	g	r
Student-t	7.00	6.14	9.04	7.09	9.36	4.00
SV + Student-t	8.61	5.69	10.99	9.72	9.96	11.24

Panel B: Financial Frictions Model

	z	ϕ	λ_f	μ	g	r	σ_ω
Student-t	6.14	7.12	6.59	5.14	6.95	2.49	2.64
SV + Student-t	6.08	7.50	6.46	5.49	6.81	7.17	2.88

Notes: Specifications shown with degrees of freedom's prior mean of 9 for the *Baseline* model, and 6 for the *Financial Frictions* model.

Figure 1: Smoothed Shocks under Gaussianity (Absolute Value, Standardized)

Entrepreneurial risk

Policy
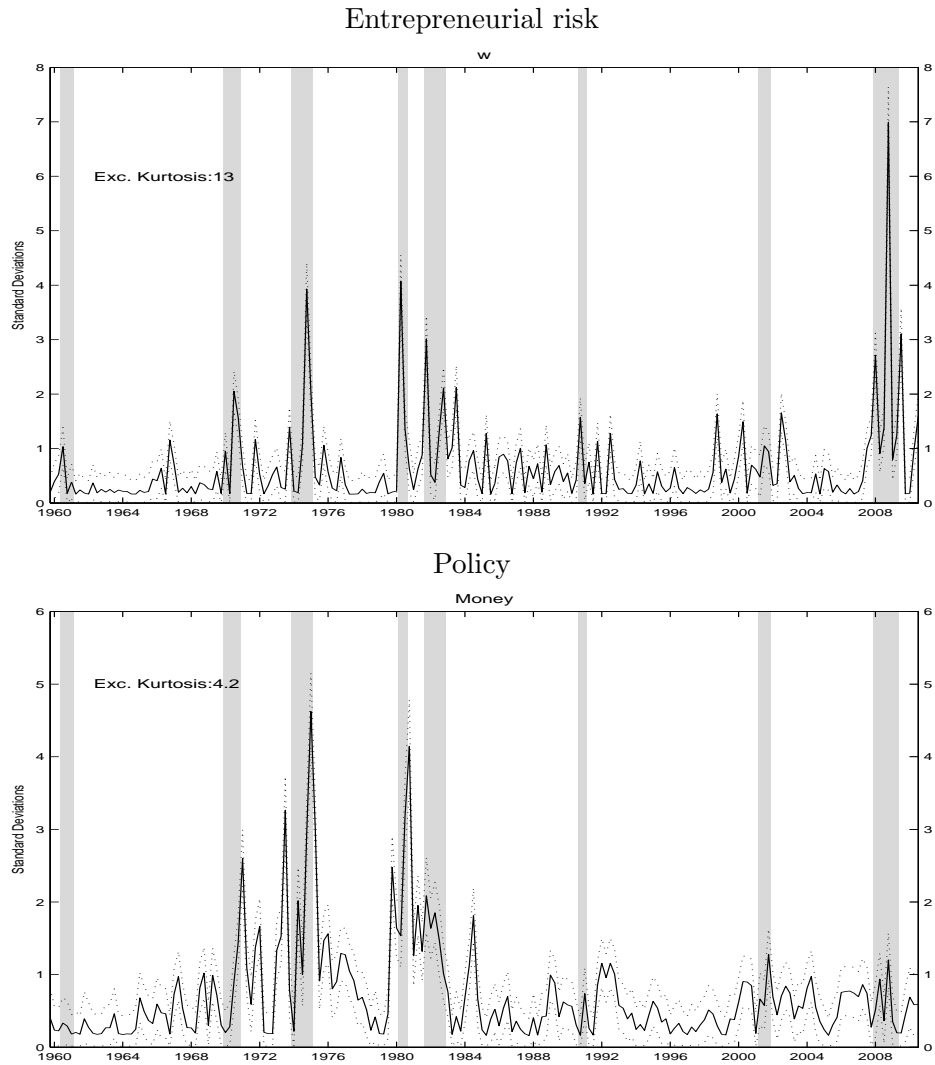
Notes: The solid line is the median, and the dashed lines are the posterior 90% bands.

Figure 2: Shocks and "Tamed" Shocks (Absolute Value, Standardized)
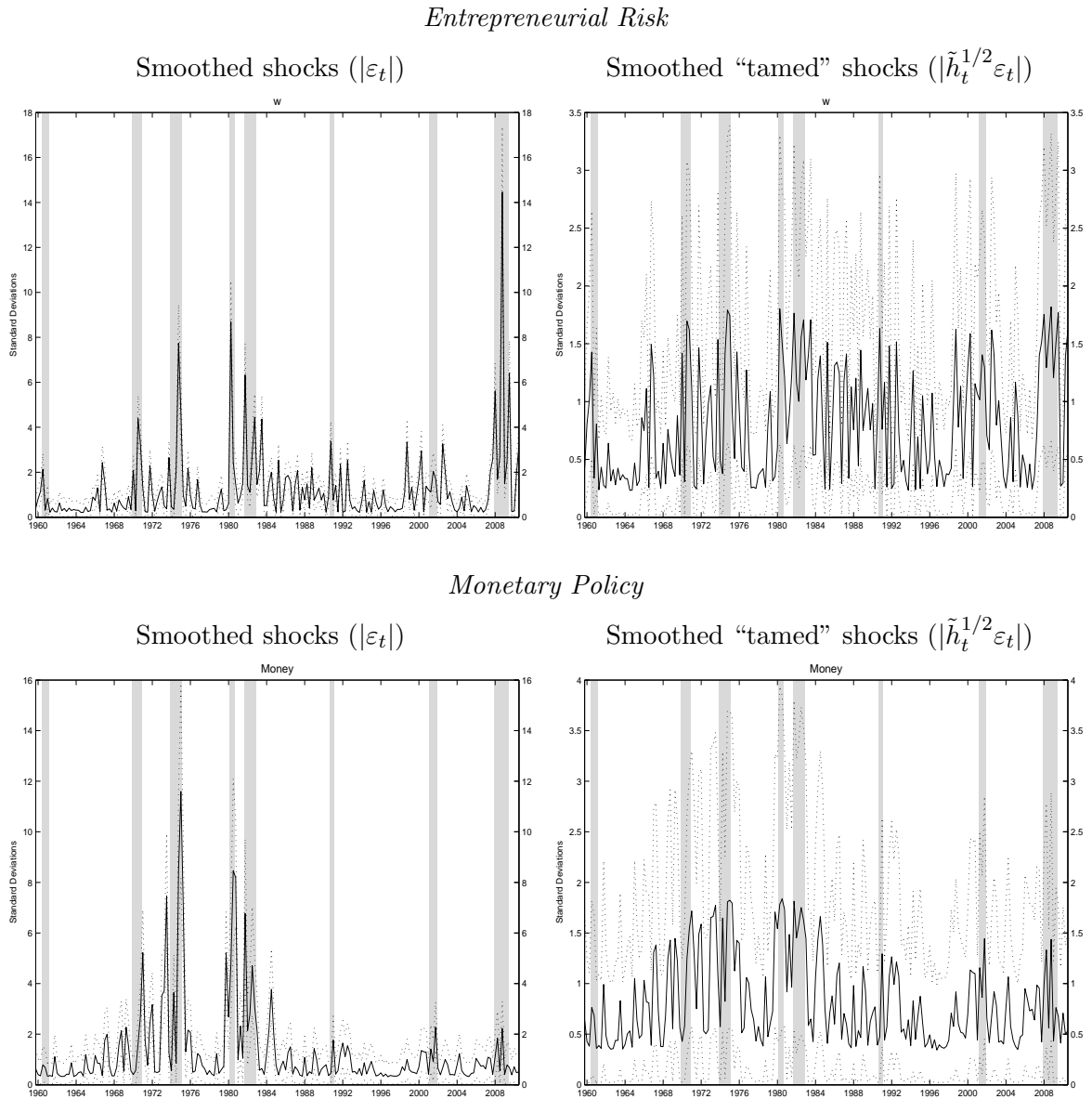
Entrepreneurial Risk

Smoothed shocks ($|\varepsilon_t|$) Smoothed "tamed" shocks ($|\tilde{h}_t^{1/2}\varepsilon_t|$)

Monetary Policy

Smoothed shocks ($|\varepsilon_t|$) Smoothed "tamed" shocks ($|\tilde{h}_t^{1/2}\varepsilon_t|$)

Notes: Estimation with Student-t distribution with $\underline{\lambda} = 6$ of the Financial Frictions Model. The solid line is the median, and the dashed lines are the posterior 90% bands. Shocks are expressed in units of the standard deviation σ_q.

Figure 3: Shocks (Absolute Values) and Smoothed stochastic volatility component,
$\sigma_q \sigma_{q,t}$

Entrepreneurial Risk
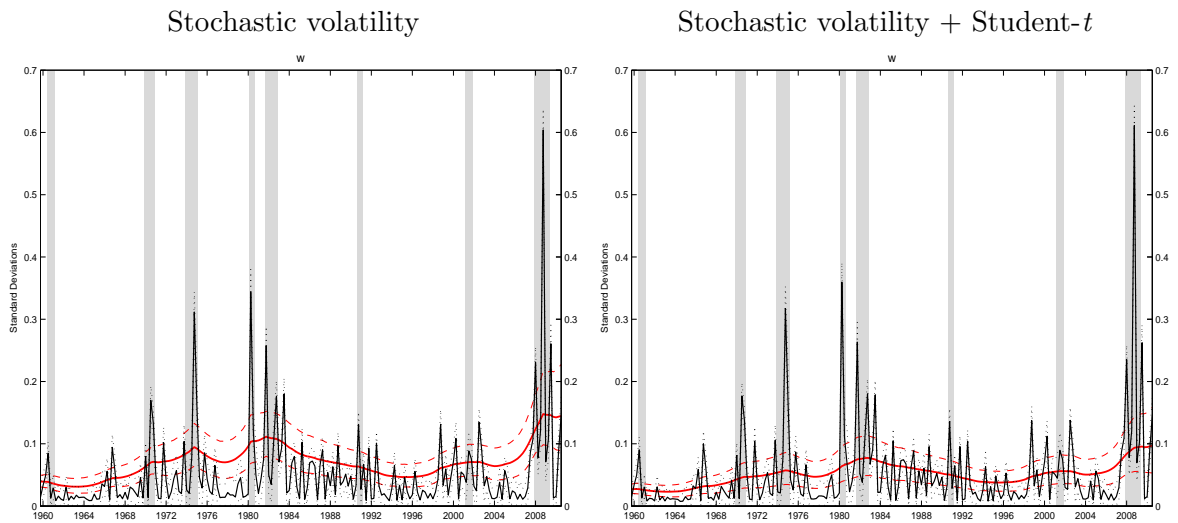
Monetary Policy
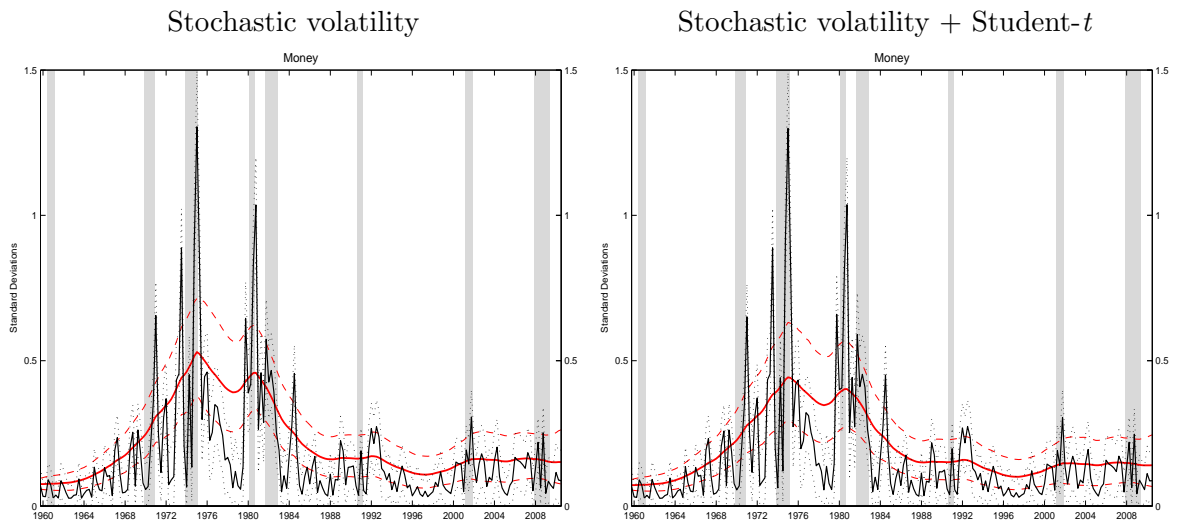
Notes: Estimation with Student-t distribution with $\underline{\lambda} = 6$ of the Financial Frictions Model. The solid line
is the median, and the dashed lines are the posterior 90% bands.

Figure 4: Contribution of the spread shock to output and hours worked fluctuations, *Financial Frictions Model*, estimation with Student-*t* distributed shocks.
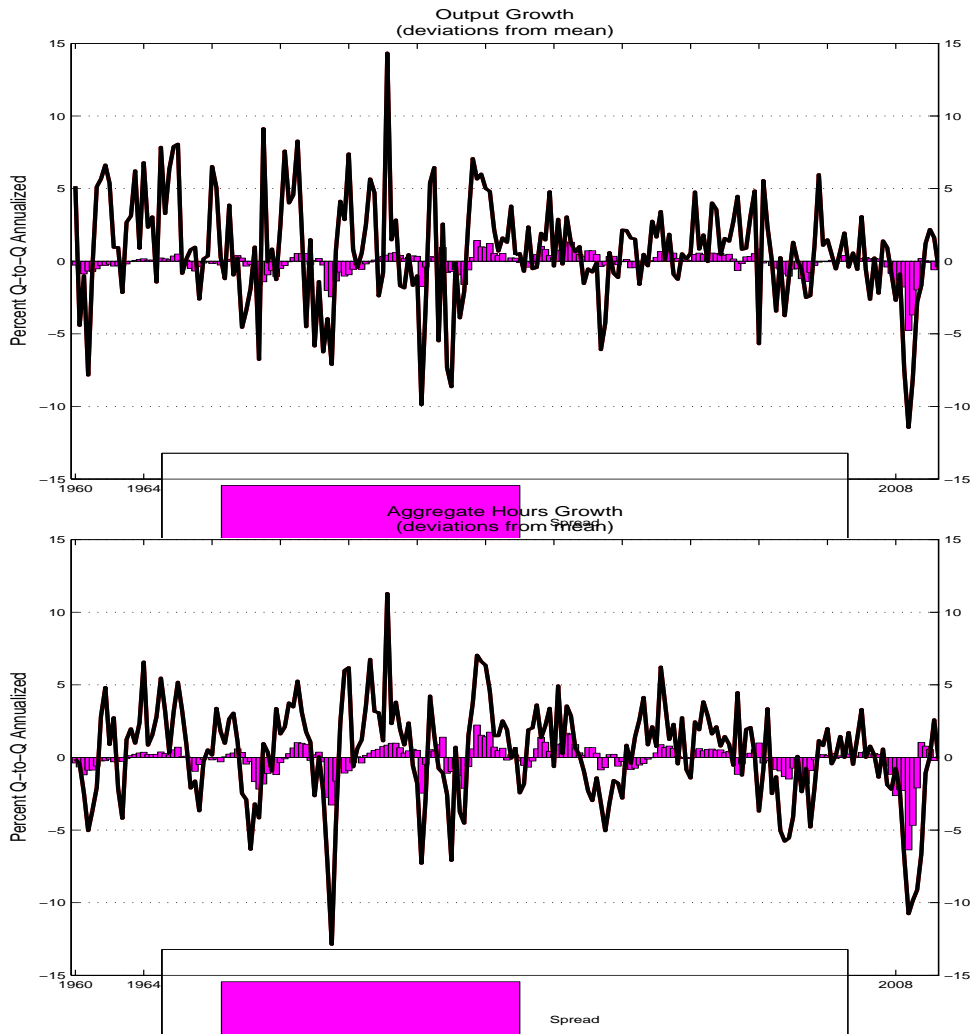
Notes: $\lambda = 6$. The black line is the historical deviations from the mean for the variable, and the pink line is the median counterfactual evolution of the same variable if only spread shocks had taken place (shutting down all other shocks).

Figure 5: Counterfactual evolution of output and hours worked when the Student-t distributed component is turned off, *Financial Frictions Model*, estimation with Student-t distributed shocks.
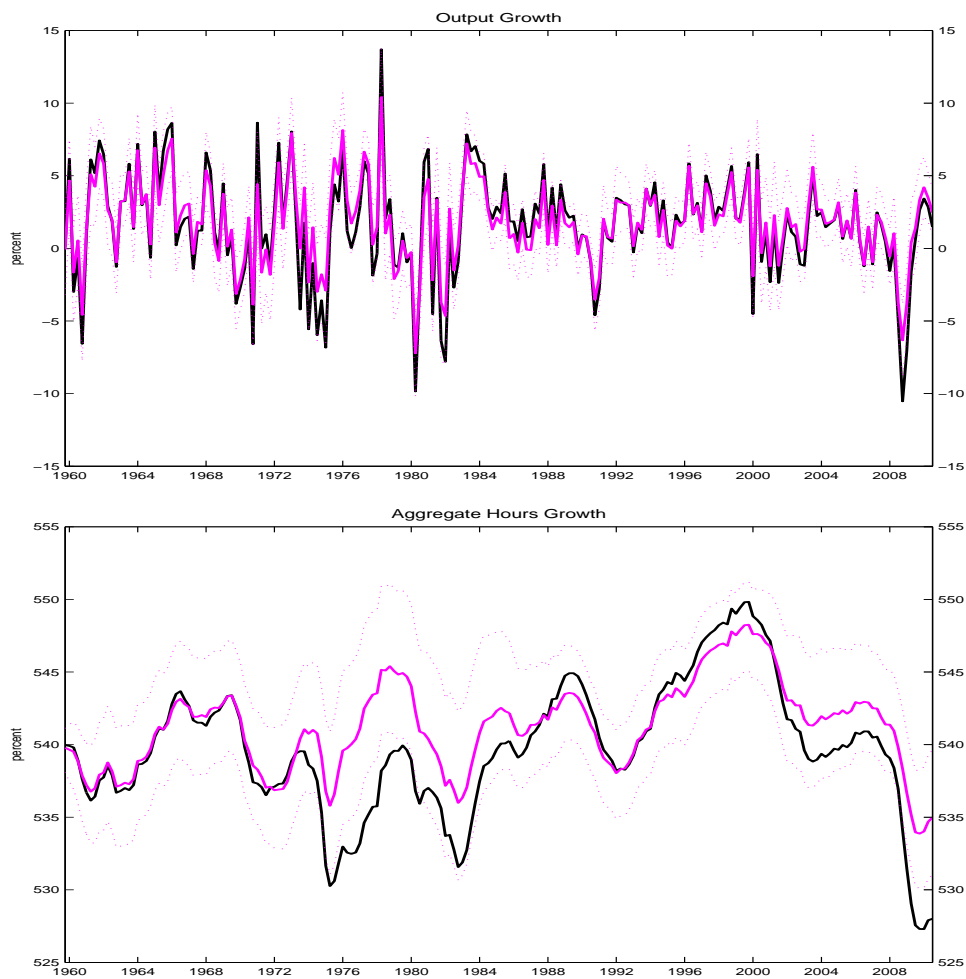
Notes: $\lambda = 6$. The black line is the historical evolution of the variable, and the pink line is the median counterfactual evolution of the same variable if we shut down the Student-t distributed component of all shocks.