Quantitative Economics 7 (2016), 257–287 1759-7331/20160257

Bayesian estimation of a dynamic stochastic general equilibrium model with asset prices

Martin Kliem
Economic Research Center, Deutsche Bundesbank

Harald Uhlig
Department of Economics, University of Chicago

This paper presents a novel Bayesian method for estimating dynamic stochastic general equilibrium (DSGE) models subject to a constrained posterior distribution for the implied Sharpe ratio. We apply our methodology to a DSGE model with habit formation in consumption and leisure, and show that the constrained estimation produces both reasonable asset-pricing and business-cycle implications. Next, we estimate the Smets–Wouters model subject to the same Sharpe ratio constraint. The results move the model closer to reproducing observed risk premia, but at increasing cost to its macroeconomic performance.

Keywords. Bayesian estimation, stochastic steady state, prior choice, Sharpe ratio.

JEL classification. C11, E32, E44, G12.

1. Introduction

This paper presents a novel Bayesian method for estimating dynamic stochastic general equilibrium (DSGE) models subject to constraints on the posterior distribution for some implied statistic of interest, such as the Sharpe ratio, that is, the market price for risk. This is akin to imposing a version of the Hansen–Jaganathan bound; see Hansen and Jagannathan (1997). We use the log-linearized law of motion to calculate a version of the Sharpe ratio as minus the ratio of the correlation between the log stochastic discount factor and the log return on capital, divided by the standard deviation of the log return on capital; see (18). Starting from an initial, unconstrained prior, we construct a constrained prior so that the resulting implied posterior for some variable of interest (the

Martin Kliem: martin.kliem@bundesbank.de
Harald Uhlig: huhlig@uchicago.edu
We would like to thank Fabio Canova, Bong Geun Choi, Lars Hansen, Ivan Jacaard, John Geweke, Holger Müller, B. Ravikumar, Tom Sargent, Frank Schorfheide, Andrea Tamoni, Mu-Chun Wang, and Tao Zha as well as three anonymous referees for their helpful comments. Moreover, we have benefited from presentations of earlier drafts at a variety of conferences and talks. This research was previously supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk" and is currently supported by NSF Grant SES-1227280. The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the Deutsche Bundesbank or its staff.

DOI: 10.3982/QE396

Sharpe ratio, for instance) coincides with some a priori chosen distribution, and such that the constrained prior is proportional to the original prior, conditional on that variable. We apply our methodology to a DSGE model with habit formation in consumption and leisure, real wage rigidities, and capital adjustment costs, extending Uhlig (2007). As a constraint, we use a sharply peaked density centered at 0.2, a previous estimate of the Sharpe ratio. We show that the estimation subject to this constraint produces a quantitative model with both reasonable asset-pricing and business-cycle implications. Next, we estimate the Smets–Wouters model (see Smets and Wouters (2007)) subject to a Sharpe ratio constraint centered around the more modest value of 0.075. The results move the model closer to reproducing observed risk premia, but at increasing cost to its macroeconomic performance.

It can be challenging to specify a dynamic stochastic general equilibrium (DSGE) model with reasonable macroeconomic implications as well as asset-pricing implications. Considerable advances have been made here in the literature in recent years. One branch of the literature exploits Epstein–Zin preference specifications developed by Epstein and Zin (1989, 1991), and includes Tallarini (2000), Rudebusch and Swanson (2012), Guvenen (2009), and Piazzesi and Schneider (2007). Another branch of the literature has pursued habit-formation specifications; see, for example, Abel (1990), Ljungqvist and Uhlig (2000), Campbell and Cochrane (1999), Boldrin, Christiano, and Fisher (2001), Uhlig (2007), Ljungqvist and Uhlig (2015), and Campbell, Pflueger, and Viceira (2015). Finally, smartly chosen tail risk distributions and disaster risks have been successfully utilized by Barro (2006), Gabaix (2012), and Gourio (2012).

But even if a model is cleverly crafted for that purpose, the estimation of such a model may not necessarily lead to parameters that allow one to deliver on both. In essence, the practical problem appears to boil down to having just a single observation on the size of the risk premium, while there are many observations helping to identify parameters crucial for the macroeconomic dynamics of the model. Thus, rather than calculating the implied posterior for the risk premium among many other statistic, our methodology instead imposes a constraint. The constraint is expressed as a density and therefore is a bit softer than an exact numerical value: this allows the use of standard sampling techniques.

Our procedure, which expands Kliem (2009), adds to the existing literature on endogenous prior choice for Bayesian estimation of DSGE models pioneered by Del Negro and Schorfheide (2008). Similarly to us, these authors propose a methodology to construct prior distributions for DSGE model parameters from beliefs about statistics of interest, which are themselves functions of the parameters. More specifically, they focus on the model-implied pth-order vector autoregression. Christiano, Trabandt, and Walentin (2011) pointed out how to modify the approach if some additional statistics of interest are provided: in the context here, that would be the Sharpe ratio. Like Del Negro and Schorfheide (2008) and Christiano, Trabandt, and Walentin (2011), our approach amounts to multiplying the original prior with a term that involves the statistics of interest as a function of the parameters. The main difference between our approach and Christiano, Trabandt, and Walentin (2011) lies in modifying the prior, so that the marginal posterior for the parameter of interest has a particular shape. This is advantageous if one wants to avoid using the data twice: the prior view about, say, the Sharpe ratio truly comes from having observed the entire data already, and it may then be odd to use these data again to calculate a posterior for that statistic of interest. The posterior will be informative about the other aspects of the model, and that will be the focus of interest. Furthermore Del Negro and Schorfheide (2008) and Christiano, Trabandt, and Walentin (2011) focus on statistics from a hypothetical pre-sample and thus their sample moments, while we allow for a free implied posterior for the statistic of interest. Our emphasis here is on constraining asset-pricing implications, and it may be instructive to compare these approaches, when applying the other two with that focus in mind.

To demonstrate our methodology and to set ourselves a bit of a challenge, we apply our method to two models, which do not easily account for asset-pricing observations. The first and main application in Section 3 is the habit-formation model in Uhlig (2007), extended to include additional shocks. Unsurprisingly and in line with a large literature, the constrained estimation delivers a high degree of risk aversion with respect to short-term fluctuations in consumption and thereby a low degree of intertemporal substitution. This is simply an unavoidable consequence of employing a log-linearized habit-formation model and seeking to match risk premia on asset markets.

The novel contribution of the methodology and insights here is to respect that consumption fluctuations now arise endogenously in the model, rather than being exogenously given as in much of the asset-pricing literature: our estimation methodology then seeks to find the optimal compromise between matching macroeconomic facts and risk premia, taking into account these general equilibrium repercussions.

The constrained estimator does so for the model in Section 3 by increasing the persistence of the response of the investment-to-output ratio, using the capital stock as a long-run buffer to smooth out short-term consumption fluctuations. This increase in the long-run volatility of the investment–output ratio may provide an intriguing connection to the long-run risk literature for Epstein–Zin preferences; see Bansal and Yaron (2004) and Hansen, Heaton, and Li (2008). The estimates further suggest that labor rigidities rather than external habits in consumption play an important role for producing our results, in line with Uhlig (2007). Remarkably, the volatility of the risk-free return is similar in size to what is observed in the data and the "risk-free rate puzzle" (Weil (1989)) is avoided. The Bayesian posterior odds ratio between the unconstrained and the constrained model is around 6.6, demonstrating that the constrained model is not "obviously false" from the unconstrained perspective.

This is in rather marked contrast to the constrained estimation of the Smets–Wouters model, pursued in Section 4. Now, the posterior odds ratio of the unconstrained model over the constrained model is $3.7e^{24}$, even though we set ourselves the more modest target of only matching half of the observed Sharpe ratio. Clear gaps open in the macroeconomic implications of the unconstrained and constrained model. For illustration, we then additionally impose a second constraint in Section 4.2, asking the estimated model to also deliver an upwardly sloping term structure of interest rates. It turns out that the method now hits a rather hard wall in terms of delivering on both the Sharpe ratio and the term premium, and cannot fully deliver the constraints one is asking for.

This, of course, should not be a surprise. If one wishes to construct a DSGE model for the purpose of, say, analyzing monetary policy or fiscal policy, and wishes that model to account for asset-pricing features such as the Sharpe ratio, it is worth introducing features that it will help to do so. Our procedure is certainly not a "cure-all." Rather, our methodology is useful for finding a good parameterization, when the model has a chance to succeed in principle, and for finding the least bad parameterization, when it does not.

The paper is organized as follows. Section 2 introduces our methodology. Section 3 applies our method to the habit-formation model of Uhlig (2007). Section 4 examines our procedure when applied to the Smets and Wouters (2007) model, including a multivariate constraint application for illustration. Section 5 concludes the paper. An additional appendix, available in a supplementary file on the journal website, http://qeconomics.org/supp/396/supplement.pdf, provides auxiliary results. Technical material is available in a supplementary file on the journal website, http://qeconomics.org/supp/396/code_and_data.zip.

2. Constructing the constrained prior

Our methodology can broadly be understood as a Bayesian estimation subject to a constraint: we shall describe it in general terms first, and then provide additional details for our core application to constraining asset-pricing implications in a DSGE model.

2.1 *Bayesian estimation subject to a posterior density constraint*

Suppose one wishes to impose the constraint (or a priori belief) that a vector of values or statistics $\omega \in \mathbb{R}^m$, which can be expressed as a function of the parameters $\omega = \Omega(\theta)$, lies within bounds $a \leq \omega \leq b$, where $a, b \in \mathbb{R}^m$. Given data X, a straightforward modification of some unconstrained prior would then be to constrain its domain to the set

$$p(\theta|X) \propto \begin{cases} p(\theta)p(X|\theta), & \text{if } \omega(\theta) \in Q, \\ 0, & \text{if } \omega(\theta) \notin Q, \end{cases} \tag{1}$$

$$\text{where } Q \equiv \{\theta : a \leq \omega(\theta) \leq b\}. \tag{2}$$

We wish to refine this approach for several reasons, building on insights learned from the specific application to constraining the Sharpe ratio in a DSGE model. First, since such a model will often favor parameters with a low Sharpe ratio, the approach above will lead to a pile-up of the constrained parameters at the lower bound a: the results will then be rather sensitive to specifying that lower bound. Second, our a priori beliefs regarding the Sharpe ratio are more appropriately formulated as a probability distribution, or, more precisely, are given by an uncertain estimate of the Sharpe ratio, given observed data, rather than imposing that it lie within some interval. Finally, the procedure becomes both more flexible and more reasonable. We argue that it is more appealing to learn the implications of the model if the Sharpe ratio is constrained to be high with considerable probability, rather than imposing that it be within some interval. We believe this insight applies more generally to other statistics of interest as well.

Consequently, we use the following approach. The model generates data X: the particular observation at hand is the realization $\bar{X}$. We start from some *unconstrained prior* $p(\theta)$ in the parameter vector θ. We then construct a *constrained prior* $\tilde{p}_{\bar{X}}(\theta)$, such that the following two properties hold:

P1. The implied posterior for the statistic of interest $\omega = \Omega(\theta) \in \mathbb{R}^m$ coincides with some a priori given distribution $F(\omega)$, with density $f(\omega)$ at the given observation $\bar{X}$.

P2. The constrained prior is proportional to the unconstrained prior, given any ω.

This approach should be viewed as a modification or refinement of the standard approach of Bayesian estimation subject to a constraint, described above. Rather than imposing an inequality constraint on a parameter or function of the parameters, we impose a probability density. We believe that P2 is a particularly appealing property and a useful benchmark. It says that prior views regarding the parameters remain unchanged, conditional on any particular value for ω. We only wish to impose that some values of ω are more likely than others, compared to the unconstrained prior. While that necessarily shifts the prior regarding the underlying parameters due to the reweighting of different ω's, we do not want to additionally impose a change in our prior views regarding various combinations of these parameters if they all result in the same value of ω.

More formally, we start with the following criteria:

1. A proper prior $p(\theta)$, which we call the unconstrained prior.
2. A likelihood function $\ell(X|\theta)$.
3. A mapping $\Omega(\theta) = \omega \in \mathbb{R}^m$.
4. A differentiable probability distribution function $F(\omega)$ on $\mathbb{R}$, that is, an increasing function with $\lim_{\omega\to-\infty} F(\omega) = 0$, $\lim_{\omega\to\infty} F(\omega) = 1$, and[1] the corresponding probability density function

$$f(\omega) = \frac{\partial^m F(\omega)}{\partial\omega_1 \partial\omega_2 \cdots \partial\omega_m}. \tag{3}$$

Given any data X, calculate the implied unconstrained posterior for ω,

$$G(\omega|X) = \int_{\theta|\Omega(\theta)\leq\omega} p(\theta|X)\, d\theta, \tag{4}$$

as well as its density $g(\omega|X)$ defined analogously to (3). Define the *transformation function*

$$h_{\bar{X}}(\omega) = \frac{f(\omega)}{g(\omega|\bar{X})}. \tag{5}$$

Note that h depends on the particular observation $\bar{X}$, indicated by the subscript. We shall use this notation throughout whenever that dependence needs to be indicated. It

[1]We assume throughout that the mathematical conditions for performing the differentiation and integration steps are satisfied.

is important that $h_{\bar{X}}$ does not vary with the data when defining the constrained prior below in (6), that is, $h(\omega)$ is a function of ω only from here onward.

We claim that the following prior has the two desired properties listed above. Define the constrained prior $\tilde{p}_{\bar{X}}$ as

$$\tilde{p}_{\bar{X}}(\theta) = C_{\bar{X}}^{-1} p(\theta) h_{\bar{X}}\big(\Omega(\theta)\big), \tag{6}$$

where

$$C_{\bar{X}} = \int p(\theta) h_{\bar{X}}\big(\Omega(\theta)\big)\, d\theta \tag{7}$$

is the integration constant. Note that (6) implies the second of the desired properties, that is, that the constrained prior is proportional to the unconstrained prior for any given ω. We need to verify the first desired property. The implied posterior for ω, given some observation X, is given by the distribution function

$$\tilde{G}_{\bar{X}}(\omega|X) = \int_{\theta|\Omega(\theta)\le\omega} \tilde{p}_{\bar{X}}(\theta|X)\, d\theta. \tag{8}$$

We need to show that $\tilde{G}_{\bar{X}}(\cdot|\bar{X}) = F(\cdot)$.

To that end, let $\tilde{g}_{\bar{X}}(\tilde{\omega}|X)$ be the density function for the multivariate distribution function $\tilde{G}_{\bar{X}}$, defined analogously to (3). Moreover, let $\Delta = (\Delta_1, \dots, \Delta_m)$ be a *volume element* or vector of small positive entries $\Delta_i > 0$. Substituting the explicit expression for the posterior, it follows that

$$\begin{aligned}
&\tilde{g}_{\bar{X}}(\tilde{\omega}|X)\Delta_1\Delta_2\cdots\Delta_m \\
&\quad\approx \int_{\theta|\tilde{\omega}\le\Omega(\theta)\le\tilde{\omega}+\tilde{\Delta}} \left(\frac{p(\theta)h_{\bar{X}}\big(\Omega(\theta)\big)\ell(X|\theta)}{\int p(\theta)h_{\bar{X}}\big(\Omega(\theta)\big)\ell(X|\theta)\, d\theta} \right) d\theta \\
&\quad\approx \frac{h_{\bar{X}}(\tilde{\omega})}{\int p(\theta)h_{\bar{X}}\big(\Omega(\theta)\big)\ell(X|\theta)\, d\theta} \int_{\theta|\tilde{\omega}\le\Omega(\theta)\le\tilde{\omega}+\Delta} p(\theta)\ell(X|\theta)\, d\theta \\
&\quad\propto \frac{f(\tilde{\omega})}{g_{\bar{X}}(\tilde{\omega}|\bar{X})} \int_{\theta|\tilde{\omega}\le\Omega(\theta)\le\tilde{\omega}+\Delta} p(\theta)\ell(X|\theta)\, d\theta.
\end{aligned} \tag{9}$$

Likewise,

$$g(\tilde{\omega}|X)\Delta_1\Delta_2\cdots\Delta_m \propto \int_{\theta|\tilde{\omega}\le\Omega(\theta)\le\tilde{\omega}+\tilde{\Delta}} p(\theta)\ell(X|\theta)\, d\theta. \tag{10}$$

Therefore, at $X = \bar{X}$,

$$\tilde{g}_{\bar{X}}(\tilde{\omega}|\bar{X})\Delta_1\Delta_2\cdots\Delta_m \propto f(\tilde{\omega})\Delta_1\Delta_2\cdots\Delta_m. \tag{11}$$

Integrating and recognizing that both sides of this equation are probability densities delivers

$$\tilde{G}_{\bar{X}}(\tilde{\omega}|\bar{X}) = F(\tilde{\omega})$$

as claimed.

These calculations also show what happens in the limit as $F(\omega)$ approaches a point mass at a particular value for $\omega = \omega^*$, as it would in the case of estimation subject to an exact constraint. In that case, $f(\omega)$ and $h_{\bar{X}}(\omega)$ approach a multiple of a Dirac delta function at $\omega = \omega^*$, and the dependence on the sample $\bar{X}$ disappears, except in insisting that the constrained prior $\tilde{p}_{\bar{X}}$ integrates to unity. This should alleviate concerns that our procedure uses the data twice: instead, it should be viewed as imposing a soft constraint. Moreover, near the point-mass limit, one nearly does not need to know $g(\omega|\bar{X})$ unless one wishes to calculate posterior odds of models and therefore needs $C_{\underline{X}}$: all that one needs is to check that it is not zero, and then to normalize $\tilde{p}$ to integrate to unity. This is of importance in cases where the unconstrained prior assigns very little mass to the area near ω^*, making it hard to calculate $g(\omega|X)$ accurately. Of course, if the constraint is imposed rather softly, the dependence on $\bar{X}$ and the precise calculation of $g(\omega|X)$ matters relatively more.

The following algorithm provides a numerical implementation of our methodology:

1. Estimate the model by sampling from the unconstrained posterior $p(\theta|X)$.
2. Approximate the implied unconstrained probability density function $g(\omega|\bar{X})$.
3. Calculate the transformation function $h_{\bar{X}}(\omega)$ per (5).
4. Estimate the model with the constrained prior by sampling from $\tilde{p}_{\bar{X}}(\theta|X)$.
5. For Bayesian posterior odds ratio calculations and model comparisons, for example, calculate the normalization constant $C_{\bar{X}}$ of (7).

The choice of the imposed implied posterior for the variable of interest is entirely up to the researcher employing this methodology and simply a generalization of textbook estimation subject to constraint. From a pure methodological perspective, it should not matter much whether such constraints are imposed a priori so as to express strong beliefs about the properties of some parameter, or a posteriori so as to learn more about certain features of a model. From the perspective of econometrics as conversation and rhetoric (see McCloskey (1983)), we view this as a practical and appealing way to explore and communicate model properties within a range of specifications of interest, rather than being forced to examine model estimates, that deliver noncredible implications for key variables. For these practical reasons, one may then wish to obtain a data-driven implied posterior for the variable of interest. Indeed, the main difference between our approach and Christiano, Trabandt, and Walentin (2011) lies in modifying the prior, so that the marginal posterior for the parameter of interest has a particular shape. This is advantageous if one wants to avoid using the data twice: the prior view about, say, the Sharpe ratio truly comes from having observed the entire data already, and it may then be odd to use these data again to calculate a posterior for that statistic of interest. The posterior will be informative about the other aspects of the model, and that will be the focus of interest.

2.2 *Constraining the sharpe ratio*

While we have described the method with the potential of imposing a multivariate constraint, we shall focus on a univariate constraint in the bulk of the application below, except for Section 4.2. In particular, we shall focus on constraining the Sharpe ratio on the returns to capital investments.

As is well known, many estimated DSGE models typically have poor asset-pricing implications, unless designed otherwise or—and that is the perspective in this paper—unless the estimation pays particular attention to these aspects. From the perspective of model applications, it is interesting to know the properties of the model if the estimation procedure is constrained to deliver certain asset-pricing implications. In our specific example, we seek to constrain our prior to deliver particular implications regarding the Sharpe ratio, that is, the market price for risk.

We impose the constraint by exploiting the implications for that market price for risk, given the log-linearized dynamics. Following Campbell (1994) and Uhlig (1999), log-linearize the given (and possibly detrended) model and solve for the recursive law of motion,

$$\hat{y}_t = A\hat{h}_{t-1} + B\varepsilon_t, \tag{12}$$

where $\hat{y}_t = \log(y_t) - \log(y^{ss})$ is a vector containing all log-linearized model variables and $\hat{h}_t = \log(h_t) - \log(h^{ss})$ is the vector containing all log-linearized state variables of the model, with y^{ss} and h^{ss} as their corresponding steady-state values. The entries in the matrices A and B can typically be interpreted as elasticities and are functions of the model parameters θ,

$$A = A(\theta), \qquad B = B(\theta).$$

Following Lettau and Uhlig (2002) and using the representation (12), we can decompose the log pricing kernel into its conditional expectation and its innovations,

$$\hat{M}_{t+1} = E_t[\hat{M}_{t+1}] + b_M\varepsilon_{t+1}, \tag{13}$$

where b_M indicates the row vector of matrix B with respect to the pricing kernel. Let $\Sigma = E_t[\varepsilon_{t+1}\varepsilon'_{t+1}]$ be the variance–covariance matrix of ε_t: we assume it to be diagonal, but the formulas apply more generally. The matrix Σ is a function of the underlying parameters θ,

$$\Sigma = \Sigma(\theta).$$

The conditional variance, σ_M^2 of the pricing kernel is

$$\sigma_M^2 = b_M \Sigma b'_M. \tag{14}$$

Similarly, we can solve for the conditional variances $\sigma^2_{R^k}$ of $\hat{R}^k_t$, the log-linearized return on some asset (where we use the superscript k to refer to the typical case of calculating the return of investing in capital). The conditional covariance of the pricing kernel and the return on the asset, σ_{MR^k}, is

$$\sigma_{MR^k} = b_M \Sigma b'_{R^k}, \tag{15}$$

with b_{R^k} the row vector of matrix B for the return on the asset $\hat{R}^k$.

Given the log-linear law of motion and exploiting the joint normality of the innovations of the logs of the variables, one can now calculate the implied risk premium on any asset, as in Lettau and Uhlig (2002). This is similar to the second-order "update" of the steady state in Schmitt-Grohe and Uribe (2004), in light of the first-order dynamics, and exploits that $E[X] = E[\exp(\log(X))] = \exp(E[\log(X)] + \mathrm{Var}[\log(X)])$ for a normally distributed random variable $\log(X)$. The formulas below would be exact if the log-linear law of motion and joint normality were exact: the approximation error is in the latter. The implied risk premium on the return of the asset over the risk-free return satisfies[2]

$$\log E_t\big[R^k_{t+1}\big] - \log R^f_t = -\sigma_{MR^k}. \tag{16}$$

Likewise, for the Sharpe ratio ω, defined here as in Lettau and Uhlig (2002) as

$$\omega = \frac{\log E_t\big[R^k_{t+1}\big] - \log R^f_t}{\sigma_{R^k}}, \tag{17}$$

we obtain

$$\omega = -\frac{\sigma_{MR^k}}{\sigma_{R^k}}. \tag{18}$$

This provides the target for our constraint in the constrained estimation procedure. Equation (18) provides us with the mapping $\Omega(\theta) = \omega \in \mathbb{R}$, which is required in the method described above. Put differently, we shall impose constraints on the Sharpe ratio for investing in capital, as this can be measured in the data using stock returns.

As Hansen and Jagannathan (1997) have shown and as was subsequently investigated further in Campbell and Cochrane (2000) and Lettau and Uhlig (2002), the highest possible Sharpe ratio is equal to σ_M when the correlation between the pricing kernel and the return of capital is equal to -1. Alternatively, one could therefore use the mapping

$$\Omega(\theta) = \omega = \sigma_M$$

if one wishes to impose a constraint on the maximal market price for risk, that is, to impose a version of the Hansen–Jaganathan bound; see Hansen and Jagannathan (1997).

To implement the procedure practically, one needs the densities $f(\omega)$ and $g(\omega|\bar{X})$. We proceed by approximating a sample histogram from the unconstrained posterior with a gamma density, to calculate $g(\cdot)$. With that, one obtains an approximation to $h_{\bar{X}}(\omega)$, which is easy to calculate accurately, but one may feel uncomfortable that the

[2]For example, $E_t[R^k_{t+1}] = E_t[\exp(E_t[\log R^k_{t+1}] + b_{R^k}\varepsilon_{t+1})] = \exp(E_t[\log R^k_{t+1}] + \sigma^2_{R^k}/2)$. A few more similar calculations deliver the equations in the text.

approximation is accurate around the desired peak $\omega^* = \text{argmax}_\omega f(\omega)$ of $f(\cdot)$. As an alternative, one could pursue a kernel density estimation procedure for $g(\omega|\bar{X})$ in particular around ω^*, but if ω^* is very unlikely under the unconditional prior, this may result in long computations. In the extreme case, there is no parameter combination resulting in ω near ω^*, in which case $g(\omega|\bar{X})$ is zero in a neighborhood of ω^* and the true $h_{\bar{X}}(\omega)$ is not defined. With a kernel density approximation procedure for $g(\omega|\bar{X})$, this case may be computationally hard to distinguish from the fundamentally rather different case of a very low density. Our (potentially rather mistaken) gamma approximation to $g(\omega|\bar{X})$ avoids these numerical challenges and potentially intense calculations at this stage, and instead produces a numerical challenge in the next step, when sampling from the constrained prior $\tilde{p}_{\bar{X}}(\theta)$. We sample from this prior by first obtaining its peak and calculating the Hessian there, for the purpose of constructing a normal density approximation and starting the Markov chain Monte Carlo (MCMC) algorithm. Consider first the extreme case that the true $g(\omega|\bar{X})$ is zero around ω^*. The maximization procedure then attempts to essentially climb $h_{\bar{X}}(\omega)$, but then stops at a sharp boundary, encountering a nondifferentiability at the peak of $\tilde{p}_{\bar{X}}(\theta)$ and before reaching ω^*. For the same reason, it may even encounter difficulties of finding the peak, when using versions of the Newton–Raphson algorithm. While these numerical challenges are solvable in principle, it now is already clear that a parameterization with the desired ω^* cannot be obtained. So, rather than doggedly proceeding with the original $f(\omega)$, we recommend adjusting this function and the target ω^* to a more modest value, where these adverse conditions do not hold, and restarting the calculations. Even in the less extreme case, where ω^* is reachable in principle, but the true $g(\omega|\bar{X})$ falls off rather sharply before reaching that value, the numerical challenges in computing the peak and the Hessian there may be rather similar in practice, and our recommendation of adjusting the target ω^* to a more modest value stays the same. One simply needs to recognize that our econometric approach is not a cure-all to fundamental deficiencies of the model itself. Some researchers may wish to proceed in this slightly more benign case anyhow, however.

The particular implementation in the context of our first application is discussed in Section 3.3 and shown in Figure 1. Note that by choosing a Gamma distribution, we implicitly impose that the Sharpe ratio is positive: this seems reasonable on economic grounds. However, for our application the choice of the Gamma distribution itself could also be substituted by, for example, a log-normal distribution or another distribution that imposes a positive Sharpe ratio. We approximate $f(\omega)$ by following our estimates for the Sharpe ratio and assume that $f(\omega)$ is Gamma distributed with mean 0.2049. For our second application in Section 4, we found that $g(\omega|\bar{X})$ is numerically ill-behaved around our initial Sharpe ratio target. Given the practical considerations explained above and to avoid moving the macroeconomic implications too far, we reduce the mean to a third of that value and set ourselves a more modest goal in terms of asset-pricing implications. Moreover, it turns out that we thus need to impose some $f(\omega)$, which is considerably tighter than would be justified on estimation uncertainty of the Sharpe ratio alone. However, this ensure that the tails of h die out and that we therefore guarantee a proper prior, and, therefore, we choose a standard deviation of 0.001 for $f(\omega)$. Since our main aim is a flexible implementation of estimation subject to constraint rather than the imposition of a particular data-driven prior, we view this as a feature, not a bug. Moreover, given that the exercise here is to obtain a model estimate in line with the observed Sharpe ratio as well as to illustrate our estimation methodology, we view this as perfectly adequate: after all, this is part of the prior specification, and not necessarily entirely data-driven.

Finally, we calculate the normalization constant $C_{\bar{X}} = \int p(\theta) h_{\bar{X}}(\Omega(\theta))\, d\theta$ of (7). To that end, we apply a MCMC algorithm and use the modified harmonic mean estimator by Geweke (1999) to compute the integral. If one can show that the Sharpe ratio only depends on a few parameters of the model, and the unconditional prior is a product of densities in these parameters and a density of the other parameters, then one can focus the calculation of the constant entirely on these few parameters, since that is where all the adjustment takes place. This Rao–Blackwellization can then help to achieve accuracy more easily.[3] Note that this constant is only used for model comparison, though, since the MCMC-based method for constrained-posterior inference does not require knowledge of the normalizing constant. If such a model comparison is intended as an overall "rough" diagnostic, one may not need to compute the normalizing constant with great precision, or other methods, such as a Laplace approximation, may suit the purpose as well.

Mathematically, one can obtain these restrictions on ω from the log-linearized dynamics, even if one calculates the log linearization around a nonstochastic steady state without a risk premium for the various assets. One may feel that doing so builds in a logical inconsistency, however. If the risk premium on investing in, say, capital is imposed to be fairly high, the steady-state level of capital should be fairly low, so as to generate this additional risk premium on average. Therefore, we follow Juillard (2011) and Coeurdacier, Rey, and Winant (2011) by evaluating the model around its risky steady state, which incorporates information about the stochastic nature of the economic environment. For the numerical solution, we impose that

$$E\big[R_t^f\big] = E\left[\exp\left(-\log \bar{M} - E_t[\hat{M}_{t+1}] - \frac{\sigma_M^2}{2}\right)\right] \tag{19}$$

as well as

$$E\big[R_{t+1}^k\big] = E\left[\exp\left(-\log \bar{M} - E_t[\hat{M}_{t+1}] - \frac{\sigma_M^2}{2} - \sigma_{MR^k}\right)\right]. \tag{20}$$

Given $\bar{M}$ as well as variances and covariances, note that the right-hand side can once again be calculated using (12) and the assumption of normally distributed shocks.

Because the conditional second moments depend on the policy function, the stochastic steady state is computationally more demanding than the deterministic steady state. In particular, the calculation requires the steady state and the corresponding dynamics to be jointly determined (see, e.g., de Groot (2013)). Therefore, we get a fixed point problem by solving our model accurately with respect to the stochastic steady state. For this reason, we use an iterative procedure. We start with the deterministic steady state to obtain our policy function. The resulting steady-state adjustment

[3] We are grateful to the co-editor for pointing this out.

yields a new set of policy functions, etcetera. As discussed in Canton (2002), a few iterations suffice to achieve convergence and to resolve the fixed point problem to a reasonable degree of accuracy.

3. Application 1: Habit formation and asset pricing

We apply our methodology to a dynamic stochastic general equilibrium model with external habit formation or "catching up with the Joneses" (see Abel (1990)) in both consumption and leisure, building on Uhlig (2007). We extend the model by adding shocks to permit estimation as well as to solve the model around the stochastic steady state. We briefly describe the model here for the sake of completeness.

3.1 *The model*

Output y_t is produced by a competitive sector of firms. Each firm produces output y_t with capital k_{t-1} in place from the previous period as well as labor n_t per the Cobb–Douglas production function

$$y_t = k_{t-1}^{\theta}\big(e^{z_{P,t}} n_t\big)^{1-\theta}, \tag{21}$$

where $z_{P,t}$ is a productivity or technology parameter. We assume it to follow a random walk with drift, $z_{P,t} = \gamma + z_{P,t-1} + \varepsilon_{P,t}$, $\varepsilon_{P,t} \sim N(0, \sigma_P)$, with γ reflecting the trend.

There is a representative household with the utility function

$$U = E\left[\sum_{t=0}^{\infty} \beta^t \frac{\big((c_t - H_t)\big(A + \big(e^{z_{L,t}} l_t - F_t\big)^{\nu}\big)\big)^{1-\eta} - 1}{1-\eta}\right]. \tag{22}$$

The discount factor β, and A, ν, and η are parameters, which we assume to satisfy $\nu > 0$ and $\eta > \nu/(\nu+1)$ so as to assure monotonicity and concavity; see Uhlig (2007). The variables c_t and l_t denote consumption and leisure of the particular household. The utility depends on the economy-wide average level of consumption habit and leisure habit, H_t and F_t, evolving according to

$$H_t = e^{\gamma}\big((1-\rho_c)\chi C_{t-1} + \rho_c H_{t-1}\big), \tag{23}$$

$$F_t = (1-\rho_l)\psi L_{t-1} + \rho_l F_{t-1}, \tag{24}$$

where C_t and L_t are aggregate average levels of consumption and leisure: in equilibrium, $C_t = c_t$ and $L_t = l_t$. The parameters ρ_c, ρ_l, χ, and ψ determine the persistence and importance of the habit features. The variable $z_{L,t}$ represents a labor supply shock. We assume it to follow an (autoregression) AR(1) process, $z_{L,t} = \pi_L z_{L,t-1} + \varepsilon_{L,t}$, $\varepsilon_{L,t} \sim N(0, \sigma_L)$. Total time endowment is normalized to unity, so that total labor supply is $n_t = 1 - L_t$.

The budget constraint of the agent is

$$c_t + x_t + T_t = d_t k_{t-1} + w_t n_t. \tag{25}$$

Capital accumulation is affected by a depreciation rate δ and investment adjustment costs $g(\cdot)$,

$$k_t = \left(1 - \delta + g\left(e^{z_{I,t}}\frac{x_t}{k_{t-1}}\right)\right)k_{t-1}. \tag{26}$$

Following Jermann (1998), we assume the adjustment cost function $g(\cdot)$ to satisfy $g(\tilde{\delta}) = \delta + \exp(\gamma) - 1$, $g'(\tilde{\delta}) = 1$, and $g''(\tilde{\delta}) = -1/\zeta\ \forall \zeta > 0$, where $\tilde{\delta}$ is defined as $\tilde{\delta} = \exp(\gamma) + \delta - 1$, to adjust for trend growth. Adjustment costs are affected by the parameter $z_{I,t}$, following the AR(1) process: $z_{I,t} = \pi_I z_{I,t-1} + \varepsilon_{I,t}$, $\varepsilon_{I,t} \sim N(0, \sigma_I)$. Given initial capital k_{-1}, the household maximizes its utility by choosing leisure l_t, consumption c_t, and investments x_t subject to the constraints (25) and (26), taking as given the exogenous habits H_t and F_t and their aggregate evolution, real wages w_t, dividends d_t, and lump-sum taxes T_t.

The agent's first-order condition for labor supply yields the frictionless wage or the marginal rate of substitution, $w_t^f = U_L/U_c$. As motivated in Uhlig (2007), we assume real wage rigidities as postulated by, for example, Hall (2005), Shimer (2005), and Blanchard and Galí (2007). More precisely, we assume that

$$w_t = \left(e^{\gamma} w_{t-1}\right)^{\mu}\left(e^{\varpi + z_{W,t}} w_t^f\right)^{1-\mu}. \tag{27}$$

The parameter $\varpi > 0$ represents an average wage markup to ensure that $w > w^f$ locally around the steady state, and that therefore the labor market is (typically) demand constrained. The wage markup $z_{W,t}$ follows an AR(1) process: $z_{W,t} = \pi_W z_{W,t-1} + \varepsilon_{W,t}$, $\varepsilon_{W,t} \sim N(0, \sigma_W)$. The parameter μ reflects the degree of real wage stickiness. In the special case of $\mu = \varpi = z_{W,t} = 0$, there are no frictions and wages are fully flexible.

Finally, there is a government, financing an exogenously given stream of expenditures G_t with lump sum taxes, $G_t = T_t$. We assume that $G_t = \bar{g}\exp(z_{P,t-1} + z_{G,t})$, where $\exp(z_{P,t-1})$ appears to assure a stationary spending-to-output ratio and where $z_{G,t}$ is assumed to follow the AR(1) process: $z_{G,t} = \pi_G z_{G,t-1} + \varepsilon_{G,t}$, $\varepsilon_{G,t} \sim N(0, \sigma_G)$.

For the further analysis, define λ_t as the marginal utility of consumption and define the stochastic discount factor $M_t = \beta\frac{\lambda_t}{\lambda_{t-1}}$. Risk premia arise from investigating the Lucas asset-pricing equation

$$1 = E_t[M_{t+1}R_{t+1}], \tag{28}$$

where R_{t+1} is the one-period return on investing one unit of resources. For investing in capital, for example, let $q(\cdot)$ be the shadow price of a unit of capital and let R_t^k be the return for investing in capital.

The five entries of the shock vector $\varepsilon_t = [\varepsilon_{P,t}, \varepsilon_{L,t}, \varepsilon_{I,t}, \varepsilon_{W,t}, \varepsilon_{G,t}]'$ are assumed to be independent. Equilibrium is defined as usual.

For the numerical analysis, the variables k_t, y_t, c_t, H_t, w_t, w_t^f, x_t, λ_t, and G_t have to be productivity-detrended to solve the model. This is done by dividing each variable by $\exp(z_{P,t-1})$, except capital k_t, which is detrended with $\exp(z_{P,t})$, and λ_t, which is detrended by $\exp(-\eta z_{P,t-1})$. Beside this, l_t, F_t, n_t, q_t, R_t^f, R_t^k, M_t, and d_t are stationary. We use a logarithmic approximation around the stochastic detrended steady state for our computations and solve for the recursive law of motion. In the following discussion, all detrended variables are marked with $\sim$ and the log deviations from the detrended variables are marked with $\wedge$. For details, see the Supplement.

3.2 *Data*

The estimation of the model is based on six time series from 1963:qI to 2008:qII. All data are quarterly and in real terms.[4] For both the unconstrained and constrained estimation, we use the vector of time series $X_t = [\Delta\tilde{y}_t, \Delta\tilde{c}_t, \tilde{n}_t, \tilde{x}_t - \tilde{y}_t, \tilde{R}_t^f, R_t^q]$, where $\Delta\tilde{y}_t$ is the demeaned first difference of log output per capita, $\Delta\tilde{c}_t$ is the demeaned first difference of log consumption per capita, $\tilde{x}_t - \tilde{y}_t$ is the demeaned log investment-to-output ratio, $\tilde{n}_t$ is demeaned hours worked per capita, $\tilde{R}_t^f$ is the demeaned risk-free rate, and $\tilde{R}_t^q$ are the excess returns on an aggregate stock market index. Because there is no equivalent variable in our model, we define log excess returns as

$$R_t^q = -\sigma_{MR^q} + \frac{1}{1-\Upsilon}\left(\hat{R}_t^k - \hat{R}_t^f\right) + \varepsilon_{Q,t}, \tag{29}$$

where Υ is a parameter that can be interpreted as leverage and $\varepsilon_{Q,t}$ is an independent and identically distributed (i.i.d.) error term and assumed to be uncorrelated with the stochastic discount factor. One may wish to simply interpret $\varepsilon_{Q,t}$ as a measurement error. Conceptually, however, the term reflects random variations in dividend policies or random fluctuations in demands by heterogeneous investors, which require far more detailed modeling, but which we view as largely immaterial for the aggregate quantitative dynamics of the model.[5] Similar to Section 2.2, we derive the mean excess returns as the negative covariance between the pricing kernel and the excess returns, $-\sigma_{MR^q}$. While we can observe the Sharpe ratio of the total returns on an aggregate stock market index in the data, we cannot observe a Sharpe ratio for the return on economy-wide capital in the data. Hence, we assume that in our economy all assets are priced along the implied market line and therefore excess returns and capital returns share the same Sharpe ratio. Thus, we use $\omega \cdot \sigma_{R^q}$ as a proxy for the mean excess returns.

3.3 *Choice of the prior*

As in Uhlig (2007), we find it useful to describe our unconstrained prior in terms of an economically meaningful transformation of the model's parameter. Specifically, consider the Frisch elasticity of labor supply, defined as the elasticity of labor supply to frictionless wages by holding the marginal rate of consumption constant,

$$\tau = \frac{dn}{dw^f}\frac{w^f}{n}\bigg|_{\bar{U}_c}. \tag{30}$$

[4] See the Supplement for details on the source and a description of any data used in this paper.

[5] In the Supplement, we present a historical and variance decomposition to investigate the role of $\varepsilon_{Q,t}$ in more detail. Unsurprisingly, this shock explains most of the variation in R_t^q. Per our model assumptions, the pricing kernel does not depend on that shock and therefore the shock does not contribute to the Sharpe ratio. We recheck this model property numerically in the Supplement.

Given our preference assumptions, this yields

$$\tau = \frac{U_n}{\bar{n}\left[U_{nn} - \frac{U_{nc}^2}{U_{cc}}\right]} = \frac{1-\bar{n}}{\bar{n}} \cdot \frac{\eta(1+\alpha)(1-\psi)}{\eta\big(\alpha(1-\nu)+1+\nu\big)-\nu}, \tag{31}$$

where $\alpha = A(1-\psi)^{-\nu}\bar{l}^{-\nu}$. Therefore, rather than specifying a prior for A or ν, we shall specify a prior for τ, and calculate the implied A and ν from τ as well as the other parameters and variables in (31). We assume that the steady-state level of hours worked is $n = 1/3$. From the first-order conditions,

$$\nu = 1 - (1-\psi)\frac{\bar{l}}{1-\bar{l}}\frac{1}{\tau} - \left(2 - \frac{1}{\eta}\right)\frac{1}{(1-\chi)\kappa}, \tag{32}$$

with

$$\kappa = \frac{e^{\varpi}}{1-\theta}\frac{1-\bar{l}}{\bar{l}}\frac{\bar{c}}{\bar{y}}, \tag{33}$$

where the $\bar{c}/\bar{y}$ is the steady-state consumption share of output.[6] Additionally, we can solve for the remaining preference parameters,

$$\alpha = \frac{\kappa\nu(1-\chi)}{1-\psi} - 1, \tag{34}$$

$$A = \alpha(1-\psi)^{\nu}\bar{l}^{\nu}. \tag{35}$$

While the real business cycle literature often assumes a relatively high Frisch elasticity of 2 or more (Prescott (1986), King, Plosser, and Rebelo (1988)), recent papers of Bayesian DSGE model estimation found far smaller values for the Frisch elasticity in a New Keynesian model framework. For example, Justiniano and Primiceri (2008) argue for values between 0.25 and 0.5. These findings are also in line with some microdata-based studies, which argue for small values in a range between 0 and 0.7 (see Pistaferri (2003) and references therein). To that end, we use a prior for the inverse of the Frisch elasticity, which is Gamma distributed with mean 1.0 and a standard deviation of 0.75. This assumption covers the values used in the different strands of the literature.

For explaining business cycle facts and asset-pricing facts simultaneously, we expect that the discount factor β as well as the power utility parameter η play an important role. The business cycle literature often uses values for the discount factor that are slightly smaller than 1 to ensure a positive time preference of the representative agent and steady-state risk-free returns comparable to observed returns. However, from an asset-pricing perspective, discount factors with much smaller values or values greater than 1 are postulated. These opposing assumptions are known as the risk-free rate puzzle (see Weil (1989)). However, Kocherlakota (1990) has shown that values for the discount factor above unity can be in line with positive time preference if the economy

[6]More details regarding steady-state calculation can be found in the Supplement.

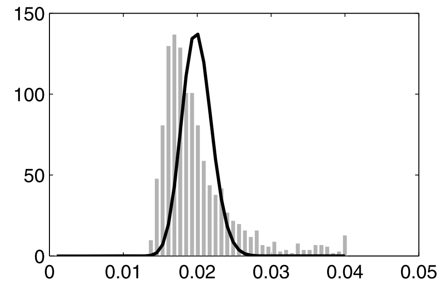


Figure 1. Implied unconstrained probability density function $g(\omega|\bar{X})$ and its approximation by a Gamma distribution (solid line).

is growing. For this reason, we use a prior information for the riskless return to ensure positive time preference and solve recursively for the discount factor:

$$\beta = \exp\bigl(\eta\gamma - \sigma_M^2/2 - \log\bigl(\bar{R}^f\bigr)\bigr). \tag{36}$$

In particular, we assume that the steady-state risk-free real quarterly return, $\log(\bar{R}^f)$, is inverted-gamma distributed with mean 0.005 and standard deviation 0.01. This ensures that the mass of the prior is on positive real annual returns that are smaller than 4%. Finally, we assume that the power utility parameter is uniformly distributed between 1 and 200, which implies most of the prior mass on high values. This is in contrast to the common business cycle literature, which generally assumes small values and therefore uses a quite informative prior, but allows our procedure to consider high degrees of risk aversion when seeking to match asset-pricing facts: this is useful when constraining the prior in the next step. The prior for the remaining deep model parameters are chosen in line with the recent literature. An overview of the priors can be found in the Supplement.

In addition, to the steady-state labor supply of 1/3, we also calibrate the growth rate of the economy γ and the capital share θ. As mentioned in the previous subsection, we calibrate the growth rate to be equal to the observed value of 0.0044 per quarter. The capital share is calibrated to 0.33 as is common. Finally, we assume a steady-state government expenditure-to-output ratio, $\bar{g}/\bar{y}$, of 0.28 and the average wage markup ϖ of 0.05.

To obtain the constraint on ω, we approximate $g(\omega|\bar{X})$ using a Gamma distribution with mean 0.02 and standard deviation 0.002. Figure 1 shows the implied unconstrained probability density function $g(\omega|\bar{X})$ and its approximation by a Gamma distribution.

3.4 *Estimation results*

We estimate the posterior mode of the distribution and employ a random walk Metropolis–Hastings algorithm to approximate the uncertainty distribution of the parameters. We run two chains, each with 300,000 parameter vector draws. The first 75% have been discarded. We provide results for both the unconstrained and the constrained prior. The posterior means for the constrained estimation are $\mu = 0.29$, $\eta = 108.18$, $\chi = 0.84$, $\psi = 0.84$, $\rho_c = 0.66$, $\rho_l = 0.07$, $Y = 0.11$, $\zeta = 7.84$, $\delta = 0.02$, $1/\tau = 7.11$, $\log(\bar{R}^f) \times 100 = 0.32$,

$\pi_G = 0.93$, $\pi_I = 0.70$, $\pi_W = 0.93$, $\pi_L = 0.65$, $\sigma_P \times 100 = 0.90$, $\sigma_I \times 100 = 1.32$, $\sigma_L \times 100 = 0.29$, $\sigma_W \times 100 = 2.11$, $\sigma_G \times 100 = 1.90$, $\sigma_Q \times 100 = 7.38$. In the Supplement, we show detailed posterior statistics, for example, the posterior mean and the highest probability density (HPD) interval, using the range between 5% and 95%. The results indicate that the posterior distributions of all structural parameters are well approximated and different from the prior distribution. In the following discussion, we highlight the key differences between both estimation results.

By comparing the results of the estimation with unconstrained prior and constrained prior, we find the biggest difference for the power utility parameter η. This result is expected, because introducing the constrained prior decisively shifts the marginal prior distribution with respect to η to high values. Additionally, for this class of preferences, the parameter is directly linked to the agents' relative risk aversion (RRA) regarding short-term fluctuations in consumption, which can be calculated as

$$\mathrm{RRA} = \frac{\eta}{1-\chi}. \tag{37}$$

We calculate the relative risk aversion for every draw from the posterior. This calculation of the relative risk aversion regarding short-term fluctuations in consumption has been criticized by Boldrin, Christiano, and Fisher (1997), who argue that relative risk aversion with respect to wealth is more meaningful. Likewise, Swanson (2012) argues that this measure ignores the labor margin. We do not seek to take a stand here on this debate: rather, our results illustrate the intuitive insight that high relative risk aversion for short-term consumption fluctuations is needed to explain stylized asset-pricing facts. In particular, the implied posterior mean of the relative risk aversion with unconstrained prior and constrained prior is 39.2 and 719.25, respectively. For both estimations, we identify similar volatilities of the exogenous shocks. This means both economies face the same economy-wide risk. Since high economy-wide risk is therefore not at the heart of matching the high Sharpe ratio according to these estimates, a high relative risk aversion (in the sense described above) is unavoidable; see also Rudebusch and Swanson (2008) or Lettau and Uhlig (2002). For habit-based DSGE models as presented in the present paper, the elasticity of intertemporal substitution is the inverse of the relative risk aversion. This elasticity can be calculated as 0.026 and 0.0013 for the estimation with unconstrained prior and constrained prior, respectively. Compared to the findings by Hall (1988) or Vissing-Jørgensen (2002), these are very small, but unavoidably so due to the need for a high degree of relative risk aversion. Some recent DSGE models likewise postulate high parameter values for external or internal habit and therefore also imply small elasticities. As shown by Uhlig (2007), wage rigidities can be a helpful ingredient to explain a high risk premium in habit-based DSGE models. Our estimation results, however, show that the degree of wage rigidity is small and similar for both estimation. Instead, the constrained estimation prefers a smaller Frisch elasticity, τ, compared to the unconstrained estimation: the Frisch elasticity of labor supply τ decreases from 0.18 for the unconstrained estimation to 0.14 for the constrained estimation. Both values are in line with findings of the microeconomic literature (see Pistaferri (2003)), but are at odds with some of the macroeconomic literature. We feel comfortable

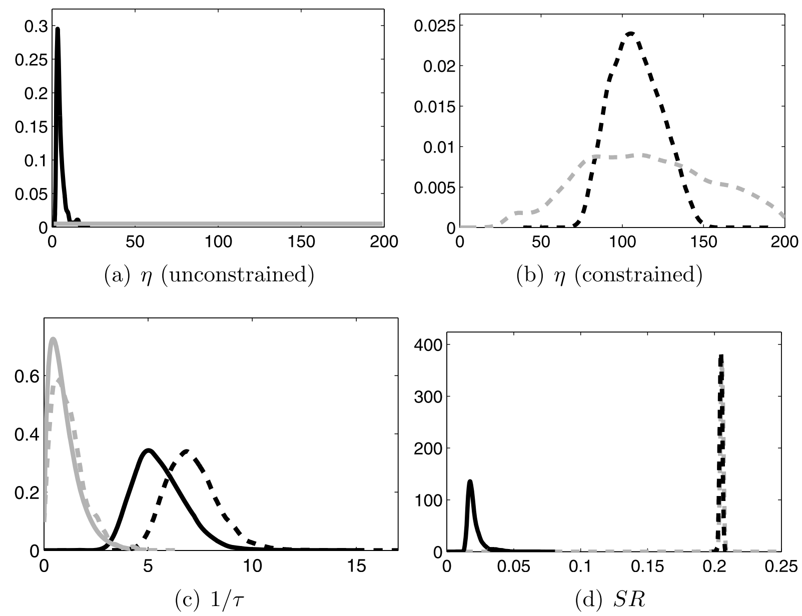


FIGURE 2. Prior and posterior distribution for selected parameters. The solid grey and solid black lines correspond to the unconstrained prior and posterior distribution, while the dashed grey and dashed black lines correspond to the constrained prior and the posterior distribution, respectively.

with these results, however, since the model is nonetheless capable of matching aggregate labor fluctuations as shown in Section 3.6: it is these aggregate observations that are at the heart of the motivation for the large macroeconomic Frisch elasticities used elsewhere. Intuitively, these estimates indicate high labor market rigidities for the estimation with the constrained prior, in line with the insights of Uhlig (2007), but focus on supply elasticities rather than wage elasticities.

Figure 2 shows the unconstrained and constrained prior and posterior distributions for selected key parameters. In particular, Figure 2(a) and 2(b) present the prior and posterior of the unconstrained and constrained estimation, respectively. The constrained estimation of the Sharpe ratio affects the prior for η strongly. While the prior distribution is still diffuse, the prior is also more concentrated around values of 100. In comparison, as illustrated in Figure 2(c), the prior of the inverse of the Frisch elasticity $1/\tau$ is not affected by the constrained estimation. Finally, Figure 2(d) illustrates the mechanism behind our approach. The constrained prior and the corresponding posterior are close to being identical as intended.

3.5 *Implied asset-pricing facts*

Table 1 shows the implied distribution of the first and conditional second moments for both estimations. In general, the estimation with constrained prior delivers asset-pricing facts similar to those observed in the data. The estimation with unconstrained

Table 1. Implied quarterly asset-pricing facts by the estimated models. All values in percent with the exception of the Sharpe ratio.

		Unconstrained Prior			Constrained Prior			
		Posterior			Posterior			
		Mean	5%	95%	Mean	5%	95%	Data
Mean risk-free rate × 100	$\log(\bar{R}^f)$	0.47	0.27	0.65	0.32	0.17	0.46	0.42
S.d. risk-free return	σ_{R^f}	0.39	0.35	0.42	0.38	0.35	0.42	0.54
S.d. return on capital	σ_{R^k}	1.19	0.97	1.41	1.18	0.98	1.37	–
S.d. excess returns	σ_{R^q}	7.86	7.20	8.54	7.59	6.98	8.22	7.27
S.d. pricing kernel	σ_M	4.16	1.90	6.82	92.15	72.13	110.92	–
Risk premium (R^k/R^f)	$-\sigma_{MR^k}$	0.02	0.01	0.03	0.24	0.20	0.28	–
Mean excess returns	$\omega \cdot \sigma_{R^q}$	0.16	0.11	0.23	1.55	1.41	1.69	1.49
Sharpe ratio (R^k) × 100	ω	2.03	1.40	2.70	20.5	20.3	20.6	20.49

prior can only explain the moments of the risk-free rate appropriately. Especially, the Sharpe ratio and the risk premium illustrate the well known difficulty of explaining asset-pricing facts using standard DSGE models.

The conditional second moments of the risk-free rate are similar for both estimations and are comparable to the data. In particular, the mean of risk-free return for the estimation with constrained prior is slightly smaller but still in line with observations. Additionally, the moments of the return of capital for both estimations are comparable with each other. The high risk aversion parameter η delivers the large conditional volatility of the stochastic pricing kernel σ_M, which in turn is needed to obtain the observed Sharpe ratio: this standard deviation is also the maximal Sharpe ratio for any asset and is approximately five times as high as the Sharpe ratio for the return on capital. Put differently, the constrained estimate of our model is in principle compatible with Sharpe ratios for other asset classes, which happen to be up to five times as large as those observed for the stock market, and therefore are compatible in principle with the findings in Scholl and Uhlig (2008) and Piazzesi and Schneider (2012). By contrast, the implied Sharpe ratio and risk premium for the unconstrained estimates is too low by a factor of 10, compared to observations.

3.6 *Implied business cycle facts*

The improvement regarding the asset-pricing implications may worsen the macroeconomic implications and, potentially, dramatically so. Remarkably, the predicted standard deviations and contemporaneous correlations with output for the Hodrick–Prescott- (HP-) filtered model variables do not change much as one moves from the unconstrained to the constrained prior; see Table S4 in the Supplement. That table shows that with both, the standard deviation of output is underpredicted a bit and consumption volatility is overpredicted a bit, compared to the data, but overall, that table suggests a decent match to macroeconomic observations either way.

For a more detailed examination, we compare the cross-correlations of the HP-filtered variables from the two model analyses (unconstrained and constrained) to those

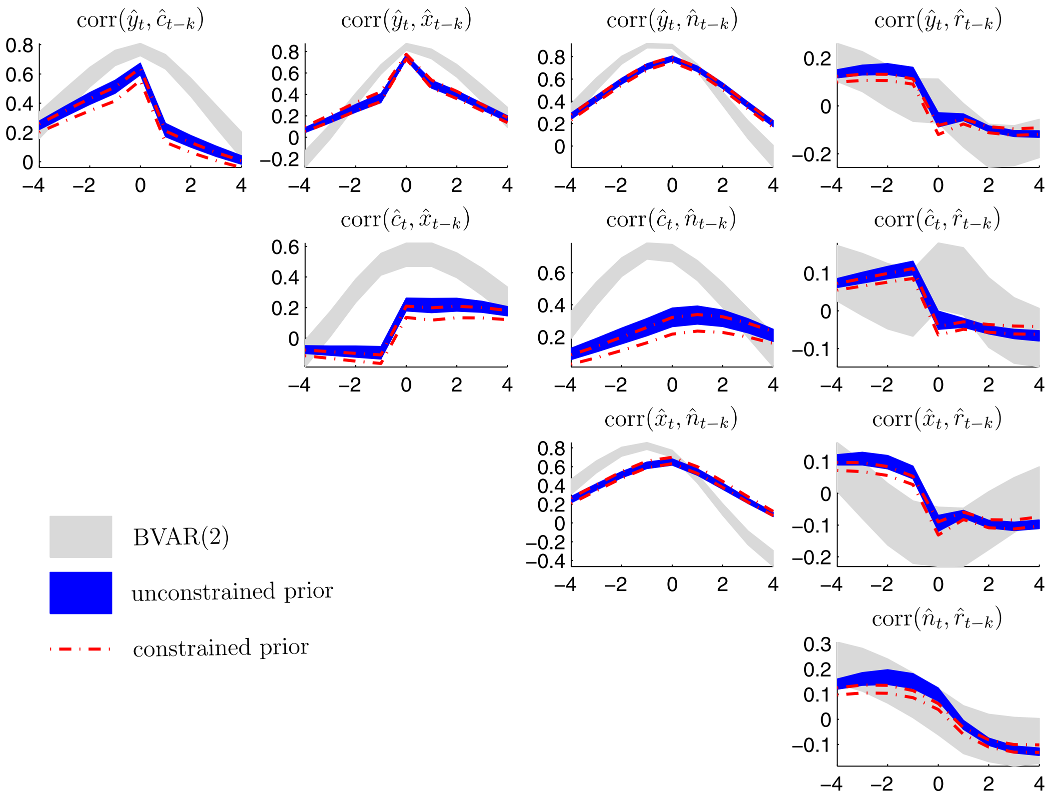


FIGURE 3. Predicted cross-correlations for the DSGEs and the BVAR(2) based on 1200 draws from the corresponding posterior.

generated from a Bayesian vector autoregression (BVAR) with two lags, using HP-filtered data;[7] see Figure 3. Once again, the differences between the unconstrained and constrained predictions are remarkably small to each other, and overall reasonably close to those of the BVAR model, with some obvious differences.

To find a difference between the predictions from the constrained and the unconstrained prior, one needs to look elsewhere. We repeat the exercise above, though this time fitting a BVAR(2) to the variables used in the estimation of the DSGE model, that is, the (log) growth rate of output, the (log) growth rate of consumption, the log investment-to-output ratio, log labor, and the log of the risk-free rate, and comparing the resulting posterior distribution over the unconditional moments to those of the two model analysis. Figure 4 shows the results. While the two model results are fairly similar for the (log) growth rate of output, the (log) growth rate of consumption, log labor, and the log of the risk-free rate, and overall differ from the BVAR(2) results for the (log) growth rate of output and the log of the risk-free rate, differences between the two model analyses

[7]We fit a BVAR(2) to HP-filtered variables. We assume a weak Normal–Wishart prior for the coefficients and the covariance matrix of the BVAR. For the comparison, we draw 1200 parameter vectors from the posterior of the BVAR as well as 1200 parameter vectors from the posterior distributions of both estimated DSGE models.

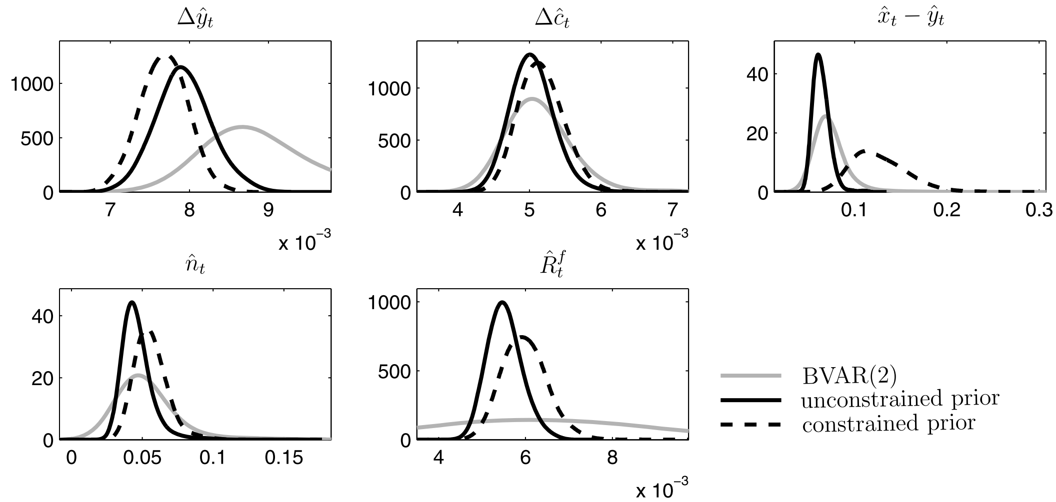


FIGURE 4. Predicted standard deviations of the DSGEs and the BVAR(2) based on 1200 draws from the corresponding posterior.

arise for the log investment-to-output ratio. While the unconstrained estimation predicts values close to those of the BVAR, the constrained estimation predicts a considerably larger standard deviation. One may interpret this as turning short-run shocks into long-run risks. This may point to an interesting connection to the long-run risk literature for Epstein–Zin preferences; see Bansal and Yaron (2004) and Hansen, Heaton, and Li (2008).

3.7 *Model comparison*

Our procedure allows a formal comparison of the constrained and the unconstrained estimate. We use the modified harmonic mean estimator by Geweke (1999) to calculate the marginal data density of each model. The log marginal data density for the unconstrained model is 3439.74 and 3437.86 for the constrained model. This difference in the marginal data density implies posterior probabilities of 0.87 versus 0.13 and a posterior odds ratio of 6.6 in favor of the model estimated with the unconstrained prior. The difference strikes us as remarkably small, and confirms the similarities between both model analyses exhibited in Figures 3 and 4. For comparison, we follow Smets and Wouters (2007) and also calculate the log marginal data density for the BVAR(2) model, which is 3569.77. The 130 log-point difference indicates that the BVAR(2) fits quite a bit better than the model, though this may be rather unsurprising, given the simplicity of the DSGE model. Indeed, we view these numbers as tolerably close, confirming the impressions from Figure 3 that both estimated DSGE models make reasonably similar predictions as the BVAR model about the second moments. In sum, our method has accomplished shifting the model parameters so as to match the Sharpe ratio without giving up too much on the macroeconomic implications, which are tolerable overall to begin with, given the simplicity of the model.

4. Application 2: Smets and Wouters

The model of Smets and Wouters (2007) had become a benchmark medium-scale New Keynesian model prior to the financial crisis in 2008. Versions of it were in widespread use, in particular at central banks. It still often provides a starting point for post-crisis DSGE models. Given the widespread use of this model, we think of it as a natural choice to apply our method. We first investigate a single constraint on the Sharpe ratio and next investigate a two-dimensional constraint on both the Sharpe ratio and the slope of the yield curve.

The Smets–Wouters model features a representative household with nonseparable utility in consumption and leisure and an external consumption habit. The households own the capital stock, choose the level of capital utilization, and pay the investment adjustment costs. Firms produce output with capital and labor. Prices and wages are Calvo sticky. The monetary authority follows a Taylor rule in inflation, output gap, and the change in output gap. Seven disturbances drive the model: the technology shock, the risk premium or equity premium shock, the government expenditure shock, the investment adjustment cost shock, the price and wage markup shocks, and the shock to monetary policy. For the details, the reader is referred to Smets and Wouters (2007).

4.1 *Constraining the sharpe ratio*

To apply our method and to highlight characteristics, we try to stay as close as possible to the original code and data files provided by the authors.[8] Nevertheless, it is necessary to extend and to adjust some issues to make the application feasible and also comparable to the former one. In the Supplement, we provide a list of equations for a complete comparison. The set of log-linearized equations is just slightly extended. In particular, we add the asset-pricing equations (19) and (20), which were substituted out in the original code. Finally, we add equation (29), which defines log excess returns, which goes along with adding the Ψ to the list of parameters and adding ε_Q to the list of exogenous shocks. Next to these changes, we solve the model around a "quasi"-stochastic steady state. More precisely, we correct the deterministic steady state by the conditional second moments from (19) and (20). We call this a quasi-stochastic steady state because we correct just for risk related to the asset-pricing equations but not for risk related to, for example, the price setting equations. This also allows us to use a more diffuse initial prior than Smets and Wouters (2007) did. Similarly to our first application, we are not restricting the discount factor β to be smaller than 1, but rather to ensure a positive time preference via the mean of the real risk-free rate. Hence, we avoid the risk-free rate puzzle by construction and solve recursively for the implied discount factor. In particular, we assume that the expression $100(\bar{R}^f - 1)$ is Gamma distributed with mean 1 and standard deviation 0.5 and we calculate the discount factor recursively as

$$\beta = \frac{1}{\bar{R}^f \exp\bigl(\sigma_m^2/2\bigr)(1+\gamma/100)^{-\sigma_c}},$$

[8] Data and Dynare code are available on http://www.aeaweb.org/aer/data/june07/20041254_data.zip.

where σ_c is the consumption preference parameter in the Smets–Wouters model and γ is the quarterly growth rate of the economy. As in our main application, we assume that the prior of σ_c is uniformly distributed between 1 and 200. The prior distributions of the remaining parameters are the same as in Smets–Wouters. Finally, we extend the original Smets–Wouters data set with the time series for excess returns as described in Section 3.2. To keep our estimation results comparable to the original Smets–Wouters paper, we use the same time episode for estimation. The Sharpe ratio in the data for this time is around 0.15, and thus is somewhat smaller than the longer time episode used for estimation in Section 3. To avoid drastic changes to its macroeconomic performance as well as to avoid numerical maximization problems, we soften our Sharpe ratio constraint by imposing a mean of 0.075 instead of 0.15 as in the data. That way, the model will only go half the way of accounting for the observed risk premium: our interest will be in how much the parameters and the macroeconomic implications change.

Figure 5 shows the unconstrained as well as constrained prior, and the posterior distributions for selected key parameters.[9] There are similarities to the findings from our first application. As should be expected, the distribution for the power utility parameter σ_c is now centered around a value corresponding to high relative risk aversion. This goes along with a tighter constrained prior distribution for the higher values of σ_c. The Frisch elasticity of labor supply $1/\sigma_l$ decreases and nominal rigidities increase when imposing the constraint. These effects are similar to those observed in Section 3. Interestingly and also similar to the first application, the parameter for habit formation h becomes smaller in both constrained estimations. While this parameter is often mentioned as a key parameter to explain asset-pricing facts, our estimation results suggest a smaller importance, especially compared to the curvature parameters, that is, σ_c. The estimated standard deviations of the preference shock becomes smaller in the constrained model, because of the higher volatility of the pricing kernel due to, for example, higher risk aversion. There are some further, but minor differences, concerning the remaining parameters. The technology shock becomes more persistent in the constrained model (see Table S6 in the Supplement), which is again an interesting connection to the long-run risk literature.

Table 2 shows the implied asset-pricing facts of the model for both estimation approaches. We are successful in matching the Sharpe ratio constraint: as a reminder, it was set at half of the observed value for this episode. Moreover, the conditional second moments of the risk-free rate are similar for both estimations and comparable to the data. Additionally, the moments of the return of capital for both estimations are comparable with each other. As before, the high risk aversion parameter σ_c delivers the large conditional volatility of the stochastic pricing kernel σ_M, which in turn is needed to obtain the Sharpe ratio constraint.

However, the posterior odds ratio is $3.7e^{24}$ to 1 in favor of the model estimated with the unconstrained prior, compared with 6.6 for the previous application; see Section 3.7. The Bayesian posterior odds ratio therefore makes the constrained model look noncredible from the unconstrained perspective. To illustrate the difficulties of the Smets–Wouters model to jointly explain both asset-pricing and macroeconomic facts, we plot

[9]We present detailed estimations results for all parameters in the Supplement.

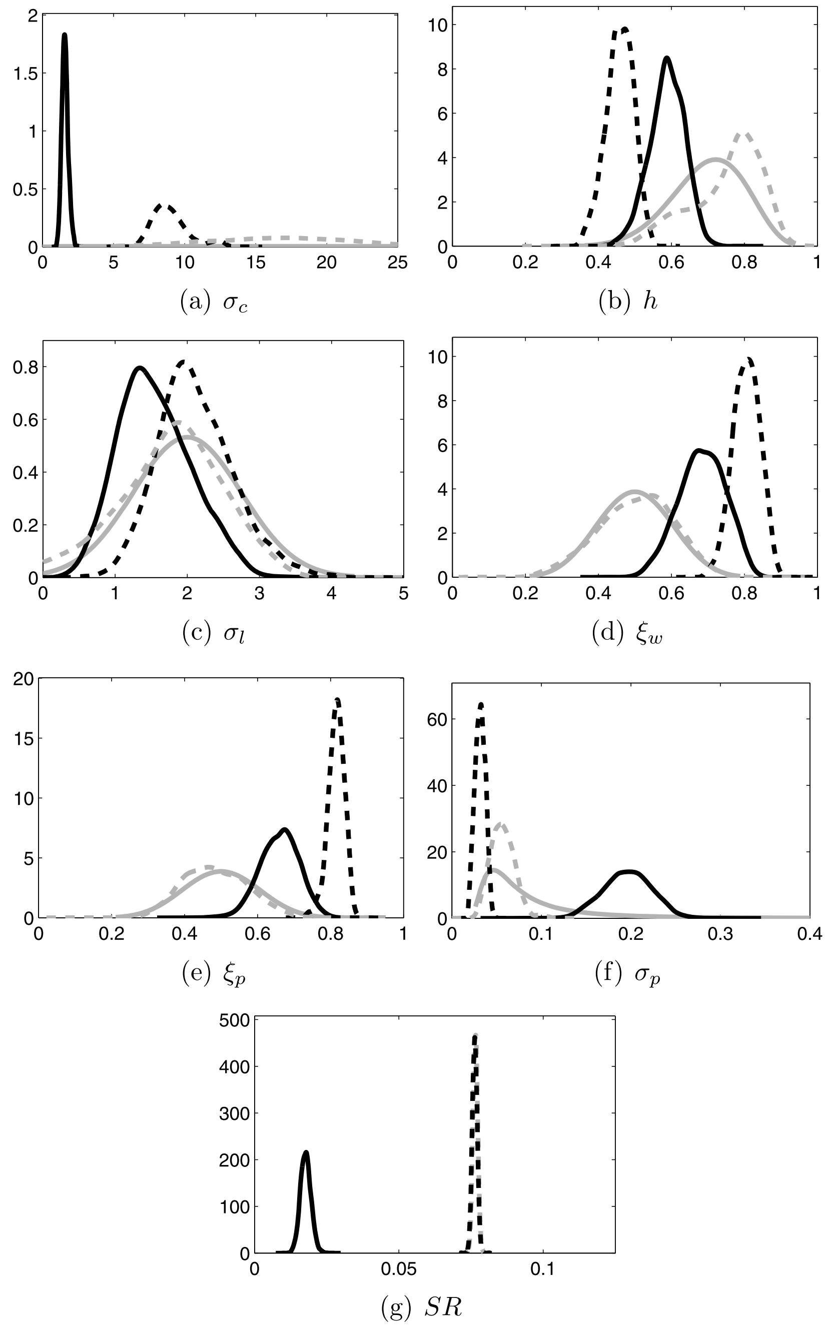


FIGURE 5. Prior and posterior distribution for selected parameters of the Smets and Wouters (2007) model. The solid grey and black lines correspond to the unconstrained prior and posterior distribution, respectively. The dashed grey and black lines correspond to the constrained prior and posterior distribution, respectively.

Table 2. Implied quarterly asset-pricing facts by the estimated Smets–Wouters models. All values are in percent with the exception of the Sharpe ratio.

		Unconstrained Prior			Constrained Prior		
		Posterior			Posterior		
		Mean	5%	95%	Mean	5%	95%
S.d. risk-free return	σ_{R^f}	0.24	0.22	0.26	0.26	0.23	0.28
S.d. return on capital	σ_{R^k}	1.76	1.52	2.01	1.83	1.56	2.13
S.d. return on capital	σ_{R^q}	8.12	7.35	8.77	8.18	7.49	8.99
S.d. pricing kernel	σ_M	2.13	1.80	2.48	12.73	10.87	14.50
Risk premium	$-\sigma_{MR^k}$	0.03	0.02	0.04	0.13	0.12	0.16
Sharpe ratio $(R^k) \times 100$	ω	1.77	1.48	2.08	7.63	7.49	7.77

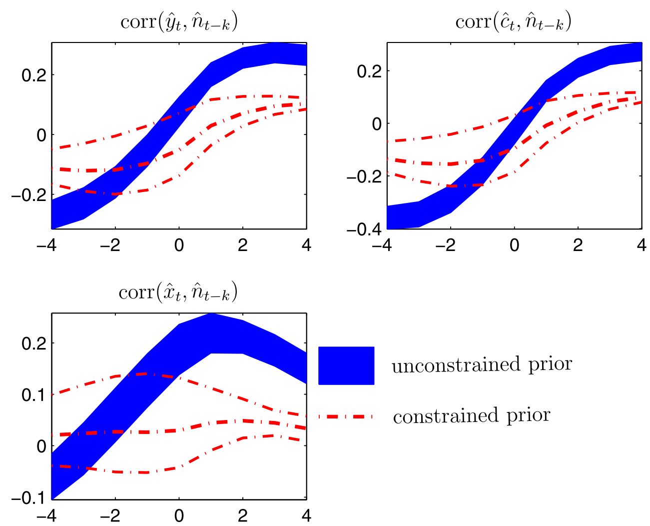


Figure 6. Predicted cross-correlations (HP-filtered) of the DSGEs based on 1200 draws from the corresponding posterior.

selected predictions of second moments for both estimations in Figure 6. In particular, we show the cross-correlation of output, consumption, and investment with inflation, respectively, based on 1200 parameter vector draws from the posterior distribution. While one prominent result of the original Smets–Wouters model is the success in accounting for the cross-correlations between output and inflation, our results show that this characteristic diminishes by using the constrained prior.

The application to the Smets–Wouters model furthermore shows the limitations of our procedure: while it tries hard to push this model to explain both asset-pricing facts and macroeconomic facts, larger gaps between the quantitative model results and the observations emerge. Indeed, trying to push the Smets–Wouters model to fully explain the observed Sharpe ratio rather than just half of it runs into the boundary problem for the constrained prior, as described more generally in Section 2.2. We therefore followed the general recommendation of Section 2.2 and pursued the more modest goal of achieving half the observed Sharpe ratio.

4.2 *A two-dimensional constraint*

As an illustration, we show how to impose a two-dimensional constraint and to calculate its implications. For the Smets–Wouters (SW) model just presented, a reasonable extension is to constrain both the bond premium as well as the Sharpe ratio. Like the Sharpe ratio, the upward sloping yield curve is a stylized fact on asset markets that is difficult to account for in standard DSGE models. In this regard, a continuously growing literature has started with the work by Backus, Gregory, and Zin (1989) and den Haan (1995), who have postulated this so-called bond-premium puzzle.

We calculate the 10-year nominal zero-coupon bond premium over a one-quarter nominal risk-free return. The return of a zero-coupon bond with k periods maturity left is analogously priced via the Euler equation

$$1 = E_t[M_{t,t+k}]R_t^{(k)}, \tag{38}$$

where $M_{t,t+k}$ is the nominal pricing kernel and R_t^k is the safe return after k periods in $t + k$. Then the quarterly bond premium, bp_t, is calculated as $\sqrt[k]{R_t^{(k)}} - R_t^{(1)}$, which in steady state can be simplified to

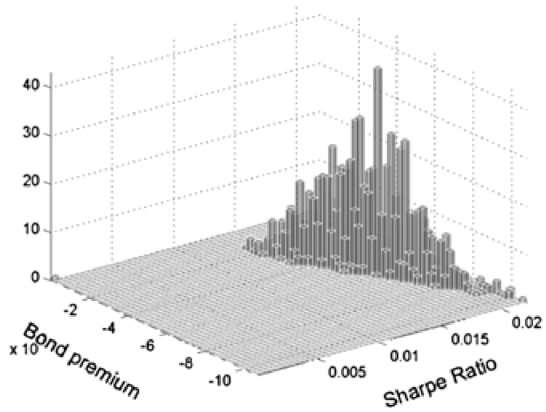
$$bp = \frac{1}{k} \cdot \frac{\sigma_M^{(k)2}}{2} - \frac{\sigma_M^{(1)2}}{2}.$$

For the unconstrained SW model we find a posterior mean for bp of approximately -0.85 basis points, which implies a slightly downward sloping yield curve. This finding is in line with the literature investigating a similar class of models. Given our application, let $\omega = [sr, bp]$ be a vector including the Sharpe ratio sr and bond premium bp in percent. Taking draws from the unconstrained posterior, Figure 7(a) illustrates the corresponding joint histogram. We approximate the joint posterior $g(\omega|X)$ with a multivariate normal distribution, $g(\omega) \sim (\bar{\mu}_M, \sum_M)$ as illustrated in Figure 7(b). To approximate the distribution $f(\omega)$, we use the bond premium data by Gürkaynak, Sack, and Wright (2007) and the excess returns of the Standard and Poors (S&P) 500 as described before.[10] In particular, we estimate a (vector autoregression) VAR(1) model consisting of the normalized excess returns and the demeaned bond premium data to calculate $f(\omega) \sim (\bar{\mu}_D, \sum_D)$.[11] Moreover, to ensure a proper prior, we impose some $f(\omega)$, which is considerably tighter than would be justified on estimation uncertainty of the joint Sharpe ratio and bond premium process alone; see Figure 7(b).

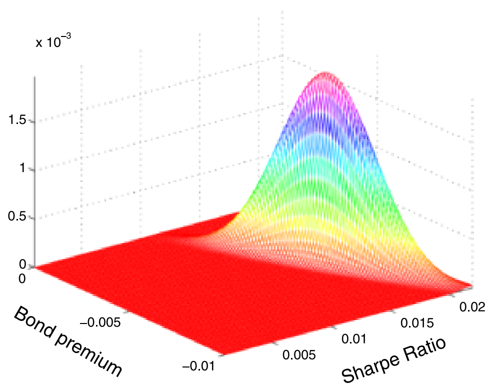
The figure furthermore illustrates the difference between the small negative bond premium implied by the unconstrained SW model and the bond premium of 47 basis points in the data. Our dual constraint estimation procedure only manages to push

[10] Due to limited data availability, we use data for the 10-year nominal zero-coupon yields from 1971:Q1 onward.

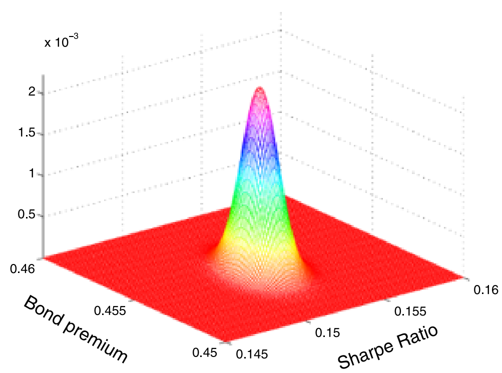
[11] See, for example, Lüthepohl (2007, p. 84) for estimation and asymptotic properties of the sample mean for a mean-adjusted VAR process.



(a) Predicted joint posterior$g(\omega|X)$



(b) Approximated joint posterior $g(\omega|X)$



(c) Joint distribution $f(\omega)$

FIGURE 7. Examples for potential distributions regarding application of a multidimensional constraint for a 10-year nominal bond premium and a Sharpe ratio.

the bond premium into slightly positive territory, with a posterior mode of 0.62 basis points, and thus rather far from the desired target of 47 basis points. Further experimentation did not result in a full resolution of accounting for both the bond premium and the Sharpe ratio, as desired. Our procedure runs into the boundary problem described more generally in Section 2.2. This finding illustrates the general remarks in Section 2.2 that our econometric approach is not a cure-all. It may well be that adding features like long-run inflation risk (see Rudebusch and Swanson (2012) and the discussion there) or Epstein–Zin preferences can provide a helpful way out if one wishes to pursue a modification of the Smets–Wouters model, which also accounts for asset market features, but these matters are beyond the scope of the present paper.

5. CONCLUSION

We have presented a novel Bayesian method for estimating dynamic stochastic general equilibrium (DSGE) models subject to constraining the posterior distribution of the implied Sharpe ratio. After presenting our methodology generally, we have applied it to two models. The first is a DSGE model with habit formation in consumption and leisure, real wage rigidities, and capital adjustment costs, provided by Uhlig (2007) and extended to feature additional shocks. We show that the estimation subject to the constraint on the Sharpe ratio produces a quantitative model with both reasonable asset-pricing as well as business-cycle implications and with tolerable posterior odds compared to the unconstrained model, thus offering more hope than the somewhat more pessimistic message in Rudebusch and Swanson (2008, 2012) regarding habit-formation models. The particular parameterization delivered by our procedure might have been hard to find otherwise.

The second application is to the Smets and Wouters (2007) model, extended with the appropriate asset-pricing equations. Now, the constrained estimator yields a parameterization, which is highly unlikely from the perspective of the unconstrained estimator. The discrepancies in macroeconomic implications are now more substantial. The results here are akin to a glass half-full or half-empty: one can either tolerate the compromise found and celebrate it as progress toward also accounting for asset market features or point to a more fundamental inability of the model to account for them. It would seem more reasonable to think about useful modifications of that model that could help with the latter.

The broader lesson here is this. If one wishes to construct a DSGE model for the purpose of, say, analyzing monetary policy or fiscal policy, and wishes that model to account for asset-pricing features such as the Sharpe ratio, it is worth introducing features that it will help to do so. Our procedure is certainly not a cure-all. Rather, our methodology is useful for finding a good parameterization when the model has a chance to succeed in principle and for finding the least bad parameterization when it does not.

References

Abel, A. B. (1990), "Asset prices und habit formation and catching up with the Joneses." *American Economic Review*, 80, 38–42. [258, 268]

Backus, D. K., A. W. Gregory, and S. E. Zin (1989), "Risk premiums in the term structure: Evidence from artificial economies." *Journal of Monetary Economics*, 24, 371–399. [282]

Bansal, R. and A. Yaron (2004), "Risks for the long run: A potential resolution of asset pricing puzzles." *Journal of Finance*, 59, 1481–1509. [259, 277]

Barro, R. J. (2006), "Rare disasters and asset markets in the twentieth century." *Quarterly Journal of Economics*, 121, 823–866. [258]

Blanchard, O. and J. Galí (2007), "Real wage rigidities and the new Keynesian model." *Journal of Money, Credit and Banking*, 39, 35–65. [269]

Boldrin, M., L. J. Christiano, and J. D. M. Fisher (1997), "Habit persistence and asset returns in an exchange economy." *Macroeconomic Dynamics*, 1, 312–332. [273]

Boldrin, M., L. J. Christiano, and J. D. M. Fisher (2001), "Habit persistence, asset returns, and the business cycle." *American Economic Review*, 91, 149–166. [258]

Campbell, J. Y. (1994), "Inspecting the mechanism: An analytical approach to the stochastic growth model." *Journal of Monetary Economics*, 33, 463–506. [264]

Campbell, J. Y. and J. H. Cochrane (1999), "Force of habit: A consumption-based explanation of aggregate stock market behavior." *Journal of Political Economy*, 107, 205–251. [258]

Campbell, J. Y. and J. H. Cochrane (2000), "Explaining the poor performance of consumption-based asset pricing models." *Journal of Finance*, 55, 2863–2878. [265]

Campbell, J. Y., C. Pflueger, and L. M. Viceira (2015), "Monetary policy drivers of bond and equity risks." Working Paper 14-031, Harvard Business School Finance. [258]

Canton, E. (2002), "Business cycles in a two-sector model of endogenous growth." *Economic Theory*, 19, 477–492. [268]

Christiano, L. J., M. Trabandt, and K. Walentin (2011), "Introducing financial frictions and unemployment into a small open economy model." *Journal of Economic Dynamics and Control*, 35, 1999–2041. [258, 259, 263]

Coeurdacier, N., H. Rey, and P. Winant (2011), "The risky steady state." *American Economic Review*, 101, 398–401. [267]

de Groot, O. (2013), "Computing the risky steady state of DSGE models." *Economics Letters*, 120, 566–569. [267]

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. [258, 259]

den Haan, W. J. (1995), "The term structure of interest rates in real and monetary economies." *Journal of Economic Dynamics and Control*, 19, 909–940. [282]

Epstein, L. G. and S. E. Zin (1989), "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework." *Econometrica*, 57, 937–969. [258]

Epstein, L. G. and S. E. Zin (1991), "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis." *Journal of Political Economy*, 99, 263–286. [258]

Gabaix, X. (2012), "Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance." *Quarterly Journal of Economics*, 127, 645–700. [258]

Geweke, J. (1999), "Using simulation methods for Bayesian econometric models: Inference, development, and communication." *Econometric Reviews*, 18, 1–73. [267, 277]

Gourio, F. (2012), "Disaster risk and business cycles." *American Economic Review*, 102, 2734–2766. [258]

Gürkaynak, R. S., B. Sack, and J. H. Wright (2007), "The U.S. treasury yield curve: 1961 to the present." *Journal of Monetary Economics*, 54, 2291–2304. [282]

Guvenen, F. (2009), "A parsimonious macroeconomic model for asset pricing." *Econometrica*, 77, 1711–1750. [258]

Hall, R. E. (1988), "Intertemporal substitution in consumption." *Journal of Political Economy*, 96, 339–357. [273]

Hall, R. E. (2005), "Employment fluctuations with equilibrium wage stickiness." *American Economic Review*, 95, 50–65. [269]

Hansen, L. P., J. C. Heaton, and N. Li (2008), "Consumption strikes back? Measuring long-run risk." *Journal of Political Economy*, 116, 260–302. [259, 277]

Hansen, L. P. and R. Jagannathan (1997), "Assessing specification errors in stochastic discount factor models." *Journal of Finance*, 52, 557–590. [257, 265]

Jermann, U. J. (1998), "Asset pricing in production economies." *Journal of Monetary Economics*, 41, 257–275. [269]

Juillard, M. (2011), "Local approximation of DSGE models around the risky steady state." Working Paper 0087, Department of Communication, University of Teramo. [267]

Justiniano, A. and G. E. Primiceri (2008), "The time-varying volatility of macroeconomic fluctuations." *American Economic Review*, 98, 604–641. [271]

King, R. G., C. I. Plosser, and S. T. Rebelo (1988), "Production, growth and business cycles: I. The basic neoclassical model." *Journal of Monetary Economics*, 21, 195–232. [271]

Kliem, M. (2009), *Essays on Asset Pricing and the Macroeconomy.* Dissertation, Humbold Universität zu Berlin. [258]

Kocherlakota, N. R. (1990), "On the 'discount' factor in growth economies." *Journal of Monetary Economics*, 25, 43–47. [271]

Lettau, M. and H. Uhlig (2002), "The sharpe ratio and preferences: A parametric approach." *Macroeconomic Dynamics*, 6, 242–265. [264, 265, 273]

Ljungqvist, L. and H. Uhlig (2000), "Tax policy and aggregate demand management under catching up with the Joneses." *American Economic Review*, 90, 356–366. [258]

Ljungqvist, L. and H. Uhlig (2015), "Comment on the Campbell–Cochrane habit model." *Journal of Political Economy*, 123, 1201–1213. [258]

Lüthepohl, H. (2007), *New Introduction to Multiple Time Series Analysis*. Springer, Berlin. [282]

McCloskey, D. N. (1983), "The rhetoric of economics." *Journal of Economic Literature*, 21, 481–517. [263]

Piazzesi, M. and M. Schneider (2007), "Equilibrium yield curves." In *NBER Macroeconomics Annual 2006*, Vol. 21, 389–472, MIT Press, Cambridge, MA. [258]

Piazzesi, M. and M. Schneider (2012), "Inflation and the price of real assets." Working paper, Stanford University. [275]

Pistaferri, L. (2003), "Anticipated and unanticipated wage changes, wage risk, and intertemporal labor supply." *Journal of Labor Economics*, 21, 729–754. [271, 273]

Prescott, E. C. (1986), "Theory ahead of business-cycle measurement." *Carnegie–Rochester Conference Series on Public Policy*, 25, 11–44. [271]

Rudebusch, G. D. and E. T. Swanson (2008), "Examining the bond premium puzzle with a DSGE model." *Journal of Monetary Economics*, 55, S111–S126. [273, 284]

Rudebusch, G. D. and E. T. Swanson (2012), "The bond premium in a DSGE model with long-run real and nominal risks." *American Economic Journal: Macroeconomics*, 4, 105–143. [258, 283, 284]

Schmitt-Grohe, S. and M. Uribe (2004), "Solving dynamic general equilibrium models using a second-order approximation to the policy function." *Journal of Economic Dynamics and Control*, 28, 755–775. [265]

Scholl, A. and H. Uhlig (2008), "New evidence on the puzzles: Results from agnostic identification on monetary policy and exchange rates." *Journal of International Economics*, 76, 1–13. [275]

Shimer, R. (2005), "The cyclical behavior of equilibrium unemployment and vacancies." *American Economic Review*, 95, 25–49. [269]

Smets, F. and R. Wouters (2007), "Shocks and frictions in US business cycles: A Bayesian DSGE approach." *American Economic Review*, 97, 586–606. [258, 260, 277, 278, 280, 284]

Swanson, E. T. (2012), "Risk aversion and the labor margin in dynamic equilibrium models." *American Economic Review*, 102, 1663–1691. [273]

Tallarini, T. D. (2000), "Risk-sensitive real business cycles." *Journal of Monetary Economics*, 45, 507–532. [258]

Uhlig, H. (1999), "A toolkit for analysing nonlinear dynamic stochastic models easily." In *Computational Methods for the Study of Dynamic Economies*, ed. by R. Marimon and A. Scott, 30–61, Chapter 3, Oxford University Press, Oxford. [264]

Uhlig, H. (2007), "Explaining asset prices with external habits and wage rigidities in a DSGE model." *American Economic Review*, 97, 239–243. [258, 259, 260, 268, 269, 270, 273, 274, 284]

Vissing-Jørgensen, A. (2002), "Limited asset market participation and the elasticity of intertemporal substitution." *Journal of Political Economy*, 110, 325–353. [273]

Weil, P. (1989), "The equity premium puzzle and the risk-free rate puzzle." *Journal of Monetary Economics*, 24, 401–421. [259, 271]

Co-editor Frank Schorfheide handled this manuscript.

Submitted October, 2013. Final version accepted November, 2015.