ECONOMIC SCENARIO GENERATOR AND PARAMETER UNCERTAINTY: A BAYESIAN APPROACH

BY

JEAN-FRANÇOIS BÉGIN

ABSTRACT

In this article, we study parameter uncertainty and its actuarial implications in the context of economic scenario generators. To account for this additional source of uncertainty in a consistent manner, we cast Wilkie's four-factor framework into a Bayesian model. The posterior distribution of the model parameters is estimated using Markov chain Monte Carlo methods and is used to perform Bayesian predictions on the future values of the inflation rate, the dividend yield, the dividend index return and the long-term interest rate. According to the US data, parameter uncertainty has a significant impact on the dispersion of the four economic variables of Wilkie's framework. The impact of such parameter uncertainty is then assessed for a portfolio of annuities: the right tail of the loss distribution is significantly heavier when parameters are assumed random and when this uncertainty is estimated in a consistent manner. The risk measures on the loss variable computed with parameter uncertainty are at least 12% larger than their deterministic counterparts.

KEYWORDS

Parameter uncertainty, economic scenario generator, Bayesian inference, Markov Chain Monte Carlo, risk.

1. INTRODUCTION AND LITERATURE REVIEW

Actuarial liabilities are very different from those of other financial institutions, with the investment time horizon being the main distinction between the two groups. In fact, most actuaries—at least life and pension professionals—are concerned mainly with the movement of investment variables over the long run, without being too concerned with very short-term fluctuations. They therefore have a great need to improve their understanding of the current

Astin Bulletin 49(2), 335-372. doi:10.1017/asb.2019.6

economic landscape and to develop a methodology that allows for a deeper understanding of the landscape's future evolution. An economic scenario generator (ESG)—which is also known as an investment model—is the basis for generating simulated asset returns and economic variables. Broadly speaking, it is a parametric model that, based on past observations, captures the joint future behaviour of the economic variables relevant to the application under scrutiny. The notion of joint behaviour is critical to the concept of ESGs: the economic variables considered might be interrelated, and some of that dependence needs to be accounted for in order to understand the future state of the economy as a whole.

Indeed, the future plausible outcomes generated by a parametric ESG rely on parameter estimates. Most studies on ESGs are based on the classical (frequentist) approach and assume that the parameters are known once the estimation procedure is carried out, meaning that all the randomness is due to the stochastic nature of a given model and that parameter uncertainty is not accounted for. Another popular school of statistical inference is the Bayesian paradigm, which was named after Reverend Thomas Bayes. While classical statisticians treat data as random samples and parameters as constant, Bayesian statisticians regard parameters as uncertain. This main philosophical distinction is convenient because it allows researchers to account for parameter uncertainty in a consistent manner.

Uncertainty is indeed a broad concept. In this study, we define uncertainty as a state of limited knowledge in which it is not possible to precisely describe a future outcome. Even though the overall impact of such uncertainty is the failure to forecast the future exactly, the uncertainty arises from different sources. We discuss two different types of uncertainty in this study: (1) process uncertainty and (2) parameter uncertainty. The former refers to the uncertainty due to the stochastic nature of the random process—the actual probabilistic model used—and the latter deals with the uncertainty in the values of the estimated parameters in a model (e.g., Bernardo and Smith, 2001, Chapter 6). A third type of uncertainty—the uncertainty in the model underlying what we can observe—is also often investigated. We leave model uncertainty for future work.

Given that parameter uncertainty *might* have dire consequences on one's risk assessment, an ESG end-user might wonder: what is the impact of parameter uncertainty. In other words, should one care about parameter uncertainty? This study answers this question.

The literature on investment models dates back to the 1980s. One of the first ESGs to be proposed in the actuarial literature is the Wilkie (1986) model. This minimal model is composed of four connected sub-models: a retail price index model, a dividend yield model, a dividend index model and a long-term interest rate model. This model's primary appeal is its ease of implementation. It is also easy to understand, somewhat parsimonious, and well known in the actuarial literature. The model was later extended and updated (Wilkie, 1995) with the inclusion of five additional variables: a wage index, property rentals and yields, short-term interest rates, index-linked stock yields and currency exchange rates.

The Wilkie model has attracted a great deal of attention over the years. Geoghegan *et al.* (1992) review the model and discuss several alternatives without making any specific recommendations. Huber (1997) reviews Wilkie's framework from a statistical perspective; he also makes some relevant suggestions for the future development of ESGs. In a series of papers, Wilkie and his collaborators investigate different assumptions of the classic Wilkie model and consider various extensions (Wilkie *et al.*, 2011; Wilkie and Sahin, 2016, 2017a,b,c, 2018). Finally, in a recent paper, Zhang *et al.* (2018) revisit the Wilkie model and assess the model's performance.

Wilkie's framework is, however, not the only ESG available in the actuarial literature (for an exhaustive review, see the Society of Actuaries' recently published in-depth primer on ESGs, that is, Pedersen *et al.*, 2016). Most of these investment models are based on time series methods and applications of stochastic processes typically used in finance and financial econometrics. For instance, a continuous-time ESG based on Lévy processes is proposed by Chan (1998); threshold autoregressive models are used by Whitten and Thomas (1999) and Chan *et al.* (2004). Chan (2002) generalizes Wilkie's model by allowing for multiple autoregressive and moving average effects across the variables via a vector autoregression moving average (VARMA) model. In a comparative study, Ahlgrim *et al.* (2008) examine two public ESGs, namely the one developed in Ahlgrim *et al.* (2004) and the model recommended by the Life Capital Adequacy Subcommittee of the American Academy of Actuaries.

Although parameter uncertainty in the context of an ESG has not received much attention in the literature, parameter uncertainty is not a new concern for actuaries (Hartman *et al.*, 2017, for a review). The first attempt to understand the role of parameter uncertainty dates back to the early 1980s; Heckman and Meyers (1983) and Meyers and Schenker (1983) both analyse the impact of parameter uncertainty on aggregate losses by adding additional variability to the frequency and severity distribution in order to make the model more sensible. In an extension of Heckman and Meyers (1983), Hayne (1999) shows that "parameter uncertainty [is a more significant issue] to insurers than simple process uncertainty". A full Bayesian approach is proposed by Cairns (2000) to capture parameter—as well as model—uncertainty. In accordance with the rest of the literature, Cairns finds that, when applied to both stochastic interest models and ruin theory, the predictive distributions under parameter uncertainty are more volatile than those derived with point estimates and therefore lead to more conservative decisions. Parameter uncertainty has also interested a number of researchers in health insurance (e.g., Fellingham *et al.*, 2015) and mortality modelling (e.g., Cairns *et al.*, 2006). More similar to our study, Hartman and Heaton (2011) assess the impact of parameter uncertainty in a (regime-switching) stochastic rate of return model. Using standard Bayesian methods, they find that adding parameter uncertainty has a more important effect on the risk measures than does simple process uncertainty. In a similar vein, Hartman and Groendyke (2013) consider a broader class of models, for example, generalized autoregressive conditional heteroskedasticity (GARCH) and stochastic volatility models.

Interestingly, the notion of quantifying parameter uncertainty in the context of investment models has been proposed—but not formally investigated—in the past. Indeed, Section 6 of Cairns (2000) contemplates parameter uncertainty within more complex models. Cairns even suggests that Markov chain Monte Carlo (MCMC) methods could be used to "allow us to deal with more complex models with many parameters, such as the Wilkie model". Additionally, in the conclusion of Wilkie *et al.* (2011), the authors highlight that there is also uncertainty in the estimation of the parameter values.

In this article, we propose a new estimation methodology for the Wilkie model based on the Bayesian paradigm. Specifically, we use the Gibbs sampler along with the Metropolis–Hastings algorithm (Metropolis *et al.*, 1953; Hastings, 1970) to obtain posterior samples of the model parameters. The Bayesian approach used in this study is indeed well suited because parameter uncertainty can be accounted for in a consistent way: the estimated random behaviour of the model parameters is in agreement with the data used. For the sake of robustness, two very different priors are used—a non-informative prior that does not put strong *a priori* beliefs on the parameters, and a subjective prior that expresses definite information about the parameters.

The end product of the MCMC method is a sample of parameters. Our results using post-World War II US data show that the average parameters are similar across the two prior assumptions. Moreover, most of the posterior means are consistent with the point estimates obtained via maximum likelihood estimation (MLE). The MCMC scheme can also recover the shape of the parameters' posterior distribution, which is often very different from the shape of the asymptotic parameter distribution under the MLE.

The MCMC sampler can be used as a way to obtain the so-called funnels of doubt in the case of parameter uncertainty (the term funnel of doubt, introduced by Redington (1952), refers to a plot showing the dispersion of a given process). Using the posterior-predictive density, paths of the four variables under study are generated and compared to those simulated without parameter uncertainty—assuming that the *actual* parameters are set to their posterior means. We find that in the case of Wilkie's ESG, parameter uncertainty adds a great deal of uncertainty to the total risk profile: the funnels of doubt are wider when we consider parameter uncertainty.

Finally, to assess the importance of parameter uncertainty from an actuarial perspective, we consider a portfolio of annuities. The stochastic rate of return of the annuity seller is a function of the variables modelled by the ESG. Through conditional distribution functions, we find that in all the cases considered, the tails of the distributions obtained under parameter uncertainty are heavier than those acquired without. The risk measures—the value at risk (VaR) and the conditional tail expectation (CTE), both at a 95% level—are at least 12% higher when we consider parameter uncertainty. Hence, once we account for parameter uncertainty, the risk profile of this portfolio of annuities changes drastically, which could have dire consequences because the risk is misspecified if one accounts for process uncertainty only. As a matter of fact, most end-users of ESG therefore utilize significantly downward-biased risk measures when they do not consider parameter uncertainty.

The rest of the paper is organized as follows. Section 2 illustrates our research question with a simple example. The Wilkie framework is presented in Section 3. In Section 4, we present the three building blocks of the Bayesian inference. Section 5 discusses the estimation methodology. Empirical results are presented in Section 6, and an application to an annuity portfolio is shown in Section 7. Finally, Section 8 concludes.

2. Illustrative example

In this section, we illustrate the impact of parameter uncertainty by means of a simple example. For the sake of simplicity, we consider only one of the economic variables typically included in ESGs: the dividend index. We assume that the annual (log) rate of return on the dividend index at time t—or the so-called dividend growth—follows a normal distribution such that

$$d(t) \sim \mathcal{N}(\mu, \sigma^2),$$

where μ is the annual mean return and σ is its standard deviation. We assume that σ is known, that is, $\sigma = 0.15$. We further assume that the mean parameter of the Gaussian distribution is uncertain and has the following distribution:

$$\mu \sim \mathcal{N}(0.06, \delta^2).$$

Parameter δ takes four different values in this example; these can be interpreted as four different levels of parameter uncertainty: no parameter uncertainty ($\delta = 0$), low uncertainty ($\delta = 0.005$), intermediate uncertainty ($\delta = 0.05$) and high uncertainty ($\delta = 0.15$).

Using this simple model, a sample of 100,000 dividend index return paths is generated via Monte Carlo simulation. Then, for each path, we cumulate the returns over a 50-year horizon to obtain the total return over this period. Figure 1 presents the funnels of doubt for the four different levels of uncertainty. It seems that the cases with no parameter uncertainty and low parameter uncertainty (upper panels) yield similar results. Indeed, the 2.5th percentiles at a 50-year horizon are almost identical: 0.941 for the case without parameter uncertainty and 0.940 when $\delta = 0.005$. When the uncertainty is moderate, that is, $\delta = 0.05$, the funnel of doubt tends to become larger, although this increase is

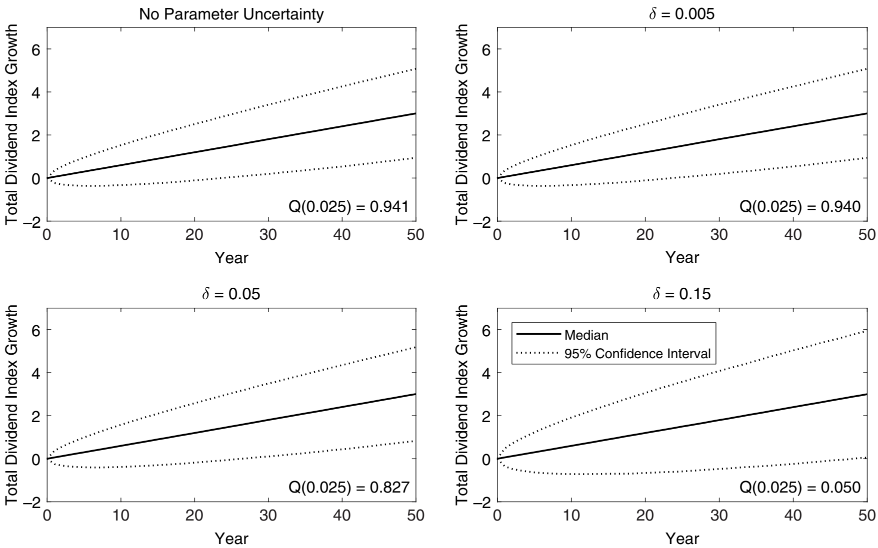

FIGURE 1: Funnels of doubt for the total dividend index growth over a period of 50 years. This figure shows the effect of parameter uncertainty on the dividend index's total (log) rate of return over 50 years, that is, $\sum_{t=1}^{\tau} d(t)$, where τ is the time horizon in years. This example considers four different parameter uncertainty levels: no parameter uncertainty, that is, $\delta = 0$ (upper-left panel), $\delta = 0.005$ (upper-right panel), $\delta = 0.05$ (bottom-left panel) and $\delta = 0.15$ (bottom-right panel). For each path, we cumulate the returns over the 50-year horizon and obtain the median total return (solid line) and the 95% confidence interval (dashed line). The 2.5th percentile of the cumulative return over the 50-year horizon is denoted by $Q(0.025)$ in the figure.

mild. In this case, the 2.5th percentile decreases to 0.827. Finally, when parameter uncertainty is large, that is, $\delta = 0.15$, the funnel of doubt is rather wide, with the 2.5th percentile at 0.050.

This example shows that, to some extent, parameter uncertainty might be inconsequential if it is low. If it is important, however, it can dramatically change the risk profile of the economic variable under scrutiny. The remainder of this article will focus on the economic implications of such parameter uncertainty if it is used in the context of actuarial risk valuation—and when accounted for in a consistent manner.

3. MODELLING FRAMEWORK: THE WILKIE MODEL

The Wilkie (1986) model is the one of the first comprehensive *open access* ESGs to be used in the actuarial academic literature. It is a stochastic asset model that describes the behaviour of various economic factors, such as the dividend yield, the dividend index, the long-term interest rate and the (price) inflation rate (see Figure 2). In that sense, Wilkie's framework is simple to grasp because it includes only four different variables. Since the publication of the full model

https://doi.org/10.1017/asb.2019.6 Published online by Cambridge University Press

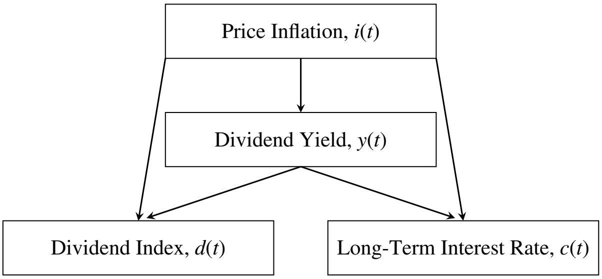

FIGURE 2: The cascade structure of Wilkie's model.
This figure shows the cascade structure adopted by Wilkie (1986). The inflation rate, $i(t)$, is linked to the dividend yield, $y(t)$, the dividend index, $d(t)$ and the long-term interest rate, $c(t)$. Moreover, the dividend index and the long-term interest rate are impacted by the dividend yield. The figure is inspired from the one included in the original Wilkie article (Page 346 of Wilkie, 1986).

in 1986, the framework has been the subject of extensive study and debate (e.g., Geoghegan *et al.*, 1992; Wilkie, 1995; Huber, 1997; Sahin *et al.*, 2008).

The Wilkie model is constructed in a so-called cascade structure, with inflation being the driver of other economic variables. The inflation rate is linked to the dividend yield, the dividend index and the long-term interest rate. Moreover, the dividend index and the long-term interest rate are impacted by the dividend yield.

Indeed, more advanced ESGs are available. The main goal of this paper, however, is to assess the impact of parameter uncertainty, not to discuss the ESG *per se*. Hence, in the following, we will use the Wilkie model as it stands, without commenting on its adequacy.

3.1. Inflation

Let $I(t)$ be the level of the consumer price index (CPI) at time t. For most—if not all—economies, the CPI time series contains a unit root. Therefore, instead of modelling the actual level of the CPI, the continuously compounded rate of change in the CPI is generally used. In Wilkie's setting, the inflation rate—the continuously compounded rate of change—is modelled by an autoregressive model of order one, AR(1), defined as follows:

$$\begin{aligned} i(t) &= \log\left(\frac{I(t)}{I(t-1)}\right) \\ &= \mu_i + a_i\,(i(t-1) - \mu_i) + z_i(t), \quad z_i(t) \sim \mathcal{N}\left(0, \sigma_i^2\right), \end{aligned}$$

where μ_i is the long-run level of the inflation rate, a_i is the parameter that deals with the autoregression and σ_i is the standard deviation of the Gaussian innovations. To make sure that the inflation dynamics are stationary, the autoregressive parameter is constrained to the interval $(-1, 1)$, that is, $|a_i| < 1$.

The autoregressive structure of the inflation rate is common in the actuarial literature; in addition to Wilkie, Sherris *et al.* (1997) and Bégin (2016) use an AR(1) process to model the inflation rate.

3.2. Dividend yield

Let $y(t)$ be the dividend yield in year t. Under Wilkie's assumptions, the log-transformed dividend yield is given by the following equations:

$$\begin{aligned}\log(y(t)) &= \log(\mu_y) + w_y i(t) + y_m(t)\\ &= \mu_y^* + w_y i(t) + y_m(t),\\ y_m(t) &= a_y y_m(t-1) + z_y(t), \quad z_y(t) \sim \mathcal{N}\left(0, \sigma_y^2\right),\end{aligned}$$

where μ_y is the long-run level of the dividend yield, w_y is the loading on the current inflation rate, a_y is the autoregressive factor of the noise process and σ_y is the standard deviation of the innovation terms. Parameter μ_y^* is a transformed version of μ_y; it is the logarithm of the long-run dividend yield. The dividend yield model can be rewritten in an autoregressive form as follows:

$$\log(y(t)) - \mu_y^* - w_y i(t) = a_y \left(\log(y(t)) - \mu_y^* - w_y i(t-1)\right) + z_y(t).$$

Again, for stationary purposes, the autoregressive parameter should satisfy the following condition: $|a_y| < 1$.

In the equations above, the dividend yield at time t is a function of the contemporaneous inflation rate $i(t)$, consistent with the cascade structure in Figure 2. According to Wilkie (1986), the level of the dividend should be (positively) related to the general level of prices in the economy—meaning that w_y should be positive.

3.3. Dividend index

Let $D(t)$ be the level of the dividend index at time t. In Wilkie's setting, the level is described by multiple inputs of current and past inflation rates, and previous innovations in the dividend yield. The dividend growth in period t is given by the following dynamics:

$$\begin{aligned}d(t) &= \log\left(\frac{D(t)}{D(t-1)}\right)\\ &= \mu_d + w_d d_m(t) + (1-w_d) i(t) + y_d z_y(t-1)\\ &\quad + b_d z_d(t-1) + z_d(t), \qquad z_d(t) \sim \mathcal{N}\left(0, \sigma_d^2\right),\\ d_m(t) &= (1-d_d) d_m(t-1) + d_d i(t),\end{aligned}$$

where μ_d is the long-run level of the dividend growth in excess of inflation, $d_m(t)$ captures the exponentially weighted inflation up to time t, d_d governs the impact of current inflation in the weighted average $d_m(t)$, y_d measures the effect of the previous shocks to dividend yield, b_d deals with the moving-average behaviour of the innovations and σ_d is the standard deviation of dividend growth innovations. As in Wilkie (1995) and Zhang *et al.* (2018), we set w_d to 1 and d_d to 0.38.

As shown in Figure 2, the inflation dynamics impact the dividend index. However, since $w_d = 1$, the current inflation rate has only an indirect impact on the dividend growth—it is accounted for only through the weighted inflation estimate, $d_m(t)$. The weighted inflation also implies that past inflation rates have a diminishing effect on the current rate of dividend growth.

The dividend yield impacts the dividend growth according to the cascade structure (see Figure 2). In Wilkie, this is achieved by including the lagged dividend yield innovation. The rationale behind the inclusion of the lagged innovation is that investors can account for unexpected changes in the previous period's dividend yield to forecast changes in the coming year's dividends—which is consistent with the Efficient Market Hypothesis (Huber, 1997).

Finally, a lagged dividend index innovation is included in the dynamics to account for the fact that "companies pay out only part of any additional earnings in dividend in one year, with a further part in the following year" (Wilkie, 1986). We therefore expect the parameter associated with this lagged innovation, b_d, to be positive. At any rate, it must at least be constrained between -1 and 1 to ensure the stationarity of the dividend index dynamics.

3.4. Long-term interest rate

Let $c(t)$ be the long-term interest rate at time t. Following Fisher's (1930) equation, the long-term interest rate is modelled as the sum of inflationary and real components:

$$c(t) = w_c c_m(t) + c_r(t),$$

where w_c is a factor dealing with the impact of the current and past inflation, $c_m(t)$ is the time-t exponentially weighted moving average of inflation and $c_r(t)$ is the real interest rate component at time t. The inflationary and real components are modelled by the following dynamics:

$$\begin{aligned} c_m(t) &= (1 - d_c) c_m(t-1) + d_c i(t), \\ \log((c_r(t)) &= \log(\mu_c) + a_c(\log(c_r(t-1)) - \log(\mu_c)) + y_c z_y(t) + z_c(t) \\ &= \mu_c^* + a_c\left(\log(c_r(t-1)) - \mu_c^*\right) + y_c z_y(t) + z_c(t), \quad z_c(t) \sim \mathcal{N}\left(0, \sigma_c^2\right), \end{aligned} \tag{1}$$

where μ_c is the long-run average of the real interest rate and μ_c^* is its logarithmic equivalent. Parameter a_c is the autoregressive parameter and y_c captures the sensitivity to the dividend yield's current innovations. We constrain the autoregressive parameter to remain within the unit interval. We set d_c to 0.058

and w_c to 1 to be consistent with Wilkie (1995) and Zhang *et al.* (2018); indeed, these two parameters are difficult to estimate (Huber, 1997). The selected values allow for a strictly positive real interest rate, given that $c_m(0)$ is not too high; this step is necessary because of the logarithmic transformation of Equation (1).

The way inflation is incorporated into the long-term interest rate is extremely similar to the case for the dividend index: we assume that the long-term rate is proportional to an exponentially weighted estimate of current and past inflation, $c_m(t)$.

As per Figure 2, the dividend yield must be incorporated into the long-term rate dynamics. In this spirit, Wilkie includes the contemporaneous dividend yield innovation into the (log) long-term interest rate equation. This innovation is multiplied by the parameter y_c.

4. Building blocks of Bayesian inference

Unfortunately, the parameters that govern the model dynamics are unknown: we must estimate them. As explained in the introduction, we want to make a Bayesian inference about Θ—the vector containing all the model parameters—because this paradigm can handle parameter uncertainty in a consistent manner. Bayesian inference allows us to update our initial belief and knowledge as more information becomes available: we wish to learn more about the unknown parameters Θ, and the $X = \{i(t), y(t), d(t), c(t)\}_{t=1}^{T}$ data allow us to do so. In other words, as a result of the Bayesian inference, we can answer the following questions: what are the plausible values of Θ and what is the extent of the uncertainty associated with these estimates?

The Bayesian approach requires not only a model from which we can construct a likelihood function but also an *a priori* distribution for Θ. This distribution, as explained earlier, represents subjective belief and knowledge. In this section, we discuss the three main building blocks of Bayesian inference: the likelihood function, the prior and the posterior distribution of the parameters. The latter is constructed by using the Bayes theorem and by combining the likelihood and the prior density:

$$\underbrace{\pi\left(\Theta|X\right)}_{\text{Posterior}} \propto \underbrace{\mathcal{L}\left(X|\Theta\right)}_{\text{Likelihood}} \underbrace{\pi\left(\Theta\right)}_{\text{Prior}}.$$

4.1. The likelihood

Conditional on the past values of the inflation rate, one can write the density of $i(t)$ as a function of $i(t-1)$:

$$f\left(i(t)\middle|i(t-1), \mu_i, \sigma_i^2, a_i\right) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2}\frac{\left(i(t)-\mu_i-a_i\left(i(t-1)-\mu_i\right)\right)^2}{\sigma_i^2}\right);$$

the likelihood associated with inflation dynamics is thus given by

$$\mathcal{L}\left(\{i(t)\}_{t=1}^{T}\middle|\mu_i, \sigma_i^2, a_i\right) = f\left(i(1)\middle|\mu_i, \sigma_i^2, a_i\right)\prod_{t=2}^{T} f\left(i(t)\middle|i(t-1), \mu_i, \sigma_i^2, a_i\right).$$

The density related to the first term, $i(1)$, is different since we do not know $i(0)$—there is no preceding observation on which to condition. To circumvent this issue, we condition on $i(0) \equiv i_0$ and treat i_0 as an additional parameter in our setting:

$$\mathcal{L}\left(\{i(t)\}_{t=1}^{T}\middle|\Theta_i\right) = \prod_{t=1}^{T} f\left(i(t)\middle|i(t-1), \mu_i, \sigma_i^2, a_i\right), \tag{2}$$

where $\Theta_i = \{\mu_i, \sigma_i^2, a_i, i_0\}$.

We can apply the same strategy to the dividend yield's dynamics: conditional on the current value of the inflation rate and the past dividend yields, we can write the density of $y(t)$ as a function of $y_m(t-1)$ and $i(t)$:

$$\begin{aligned} & f\left(y(t)\middle|y_m(t-1), i(t), \mu_y^*, \sigma_y^2, a_y, w_y\right) \\ & = \frac{1}{\sqrt{2\pi}\sigma_y} \exp\left(-\frac{1}{2}\frac{\left(\log\left(y(t)\right) - \mu_y^* - w_y i(t) - a_y y_m(t-1)\right)^2}{\sigma_y^2}\right). \end{aligned}$$

Again, to cope with the fact that $y_m(0)$ and $z_y(0)$ are unknown, we assume that both are extra parameters, that is, $y_{m,0}$ and $z_{y,0}$, respectively. Thus, the likelihood, conditional on $y_{m,0}$ and $z_{y,0}$, is given by

$$\mathcal{L}\left(\{y(t)\}_{t=1}^{T}\middle|\{i(t)\}_{t=1}^{T}\Theta_y\right) = \prod_{t=1}^{T} f\left(y(t)\middle|y_m(t-1), i(t), \mu_y^*, \sigma_y^2, a_y, w_y\right), \tag{3}$$

where $\Theta_y = \{\mu_y^*, \sigma_y^2, a_y, w_y, y_{m,0}, z_{y,0}\}$.

The density of $d(t)$, conditional on the past value of $d_m(t-1)$, the current value of inflation and the past innovations $z_y(t-1)$ and $z_d(t-1)$, is given by a normal density:

$$\begin{aligned} & f\left(d(t)\middle|d_m(t-1), i(t), z_y(t-1), z_d(t-1), \mu_d, \sigma_d^2, b_d, y_d\right) \\ & = \frac{1}{\sqrt{2\pi}\sigma_d} \exp\left(-\frac{1}{2}\frac{\left(d(t) - \mu_d - 0.62 d_m(t-1) - 0.38 i(t) - y_d z_y(t-1) - b_d z_d(t-1)\right)^2}{\sigma_d^2}\right). \end{aligned}$$

Using the same trick as before—treating the missing information as additional parameters—we can now write the likelihood associated with the dividend index as

$$\begin{aligned} & \mathcal{L}\left(\{d(t)\}_{t=1}^{T}\middle|\{i(t), y(t)\}_{t=1}^{T}, \Theta_y, \Theta_d\right) \\ & = \prod_{t=1}^{T} f\left(d(t)\middle|d_m(t-1), i(t), z_y(t-1), z_d(t-1), \mu_d, \sigma_d^2, b_d, y_d\right), \end{aligned} \tag{4}$$

where $\Theta_d = \{\mu_d, \sigma_d^2, b_d, y_d, d_{m,0}, z_{d,0}\}$, $d_m(0) \equiv d_{m,0}$ and $z_d(0) \equiv d_{m,0}$.

https://doi.org/10.1017/asb.2019.6 Published online by Cambridge University Press

Finally, the (conditional) density of $c(t)$ is given by the following Gaussian probability density function:

$$f\left(c(t)\middle|c(t-1), c_m(t), c_m(t-1), z_y(t), \mu_c^*, \sigma_c^2, a_c, y_c\right)$$
$$= \frac{1}{\sqrt{2\pi}\sigma_c} \exp\left(-\frac{1}{2}\frac{\left(\log\left(c(t)-c_m(t)\right)-\mu_c^*-a_c\left(\log\left(c(t-1)-c_m(t-1)\right)-\mu_c^*\right)-y_c z_y(t)\right)^2}{\sigma_c^2}\right).$$

To obtain a closed-form likelihood function, we will need to include additional parameters in our framework in the following way:

$$\mathcal{L}\left(\{c(t)\}_{t=1}^T\middle|\{i(t), y(t)\}_{t=1}^T, \Theta_y, \Theta_c\right)$$
$$= \prod_{t=1}^T f\left(c(t)\middle|c(t-1), c_m(t), c_m(t-1), z_y(t), \mu_c, \sigma_c^2, a_c, y_c\right), \tag{5}$$

where $\Theta_c = \{\mu_c^*, \sigma_c^2, a_c, y_c, c_{m,0}, c_0\}$.

4.2. The prior

The prior distribution is an important component of the Bayesian inference. It represents the subjective belief and knowledge about the uncertain parameters Θ and allows the researcher to incorporate "non-sample" information in a consistent manner. For instance, parameter constraints can be incorporated readily in the prior distribution. A key issue in selecting this distribution is determining what kind of information goes into the prior. In that respect, we use two different sets of assumptions in this study: a non-informative prior and a subjective one. These two sets correspond to two extremes—either no information at all on the model parameters or fairly precise *a priori* information on them. The non-informative prior is based on a combination of flat as well as very vague densities. On the other hand, the subjective prior is inspired by the parameter estimates obtained by Wilkie (1995) for the US economy. The main reason we use these two sets of prior distributions is to ensure that the prior used does not drive the results and that our conclusions are robust to a change in this assumption. For more details on the *a priori* distributions, see Appendix A.

4.3. The posterior

Given all the above ingredients, we can calculate the conditional probability density of Θ given the data X with Bayes's formula:

$$\pi\left(\Theta|X\right) \propto \mathcal{L}\left(X|\Theta\right)\pi\left(\Theta\right),$$

where the likelihood function, conditional on the parameter $\Theta = \{\Theta_i, \Theta_y, \Theta_d, \Theta_c\}$, is given by

$$\begin{aligned}\mathcal{L}(X|\Theta) = \mathcal{L}\left(\{i(t)\}_{t=1}^{T}\middle|\Theta_i\right) \times \mathcal{L}\left(\{y(t)\}_{t=1}^{T}\middle|\{i(t)\}_{t=1}^{T}, \Theta_y\right) \\ \times \mathcal{L}\left(\{d(t)\}_{t=1}^{T}\middle|\{i(t), y(t)\}_{t=1}^{T}, \Theta_y, \Theta_d\right) \\ \times \mathcal{L}\left(\{c(t)\}_{t=1}^{T}\middle|\{i(t), y(t)\}_{t=1}^{T}, \Theta_y, \Theta_c\right)\end{aligned} \quad (6)$$

and $\pi(\Theta)$ is obtained by multiplying the individual priors given in Table 5.

5. Estimation methodology

The Bayesian inference scheme laid out in the previous section provides a theoretical way to understand parameter uncertainty. However, to actually assess this uncertainty, we would need to compute the posterior distribution, $\pi(\Theta|X)$, as well as the marginal posterior distribution, $\pi(\theta_i|X)$, for each parameter $\theta_i \in \Theta$. It is, however, challenging to do so because this process involves multiple high-dimensional integrals. Various methods exist to numerically recover the marginal *a posteriori* distribution for each model parameter. In this study, we rely on MCMC methods because they allow us to readily generate samples of parameters that are consistent with the posterior distribution. For more details on the estimation methodology, see Appendix B.

6. Empirical results

In this section, we present the data used to estimate the model, report the results for both prior distributions and discuss the Bayesian prediction based on the posterior distribution of the model's parameters.

6.1. Data

The four datasets employed in this study are extracted from the Bloomberg terminal. In this study, we focus on post-World War II data, that is, data for the year 1945 and after. First, to construct the annual inflation rate, we use the monthly non-seasonally adjusted US Consumer Price Index for All Urban Consumers (CPURNSA) series from the end of December 1945 to the end of December 2016. We then convert these monthly observations into the annual inflation rate by taking the (log) return of the CPI index from one December to the next one. This step yields 72 index observations and, thus, 71 inflation rate observations.

Second, the dividend yield in the US economy is proxied by the dividend paid out on the stocks that are part of the S&P 500. It is constructed by taking the sum of the gross dividend payments over a given (calendar) year and dividing it by the value of the index at the end of the year.

Third, the S&P 500 value is used as a proxy for the dividend index. Similar to our process for the CPI, we build our annual returns out of monthly index

https://doi.org/10.1017/asb.2019.6 Published online by Cambridge University Press

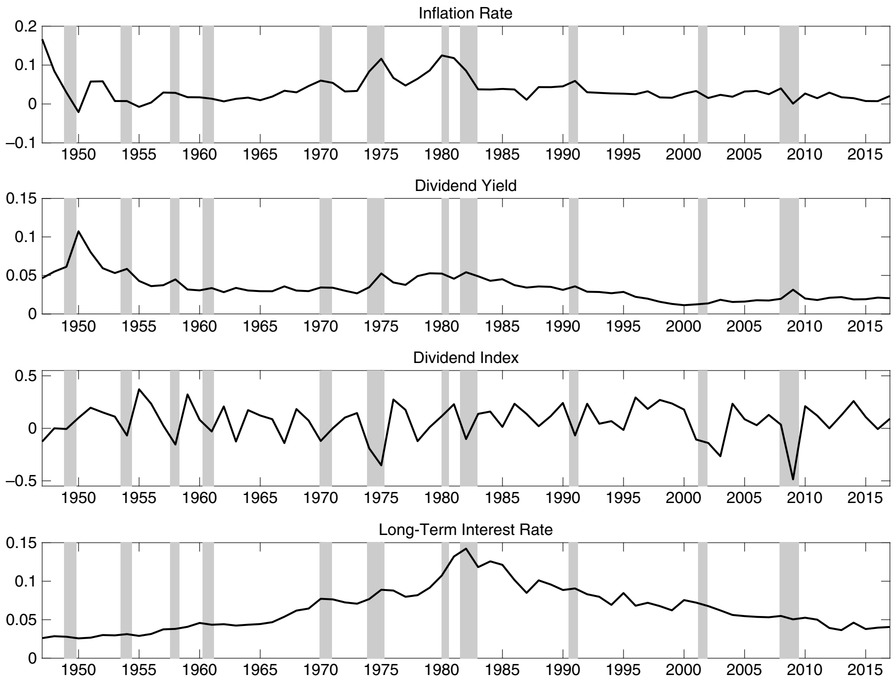

FIGURE 3: Inflation rate, dividend yield, dividend index and long-term interest rate. The upper panel of this figure presents the inflation rate of the annual non-seasonally adjusted US City Average All Items Consumer Price Index for All Urban Consumers (CPURNSA) from the end of December 1945 to the end of December 2016; grey-shaded regions highlight NBER-dated recessions. The second panel reports the dividend yield on the S&P 500 index over the same period. The third panel shows the returns on the dividend index. The lower panel reports the end-of-year Moody's Seasoned AAA Corporate Bond Yield. The four time series were obtained from the Bloomberg terminal.

values by simply log differencing the annual index value time series at the end of December. This step also yields a sample of 71 annual returns.

Finally, the long-term interest rate is proxied by Moody's Seasoned Aaa Corporate Bond Yield. These rates are again obtained via the Bloomberg terminal (MOODCAAA). We use the December series in this study. Figure 3 presents the behaviour of the four time series under study. Since we are considering post-World War II data, we still have 71 observations for this case.

6.2. Estimation results

Using the MCMC methodology explained in Section 5 and in Appendix B, we can numerically obtain the marginal posterior distributions of the model parameters. For each parameter, we obtain a sample of 100,000 values from which we can find the average value and the standard deviation. Table 1 reports the value of such means and standard deviations for both prior cases, that is, non-informative and subjective. For comparison, we also include the

TABLE 1

PARAMETER ESTIMATES FOR THE WILKIE MODEL.

	Non-informative 1945–2017	Subjective 1945–2017	MLE 1945–2017	Wilkie (1995) 1926–1989	ZHS (2018) 1926–2014
Inflation					
μ_q	0.031	0.031	0.041	0.030	0.031
	(0.007)	(0.008)	(0.012)	–	(0.008)
σ_q	0.021	0.021	0.024	0.035	0.034
	(0.002)	(0.002)	(0.002)	–	(0.003)
a_q	0.590	0.643	0.750	0.650	0.573
	(0.082)	(0.061)	(0.055)	–	(0.086)
Dividend Yield					
μ_y	0.033	0.033	0.031	0.043	0.031
	(0.014)	(0.005)	(0.008)	–	(0.010)
σ_y	0.190	0.193	0.172	0.210	0.162
	(0.017)	(0.017)	(0.014)	–	(0.012)
a_y	0.912	0.855	0.927	0.700	0.938
	(0.046)	(0.037)	(0.038)	–	(0.038)
w_y	−0.072	0.524	0.152	0.500	−0.441
	(0.943)	(0.432)	(0.534)	–	(0.469)
Dividend Index					
μ_d	0.040	0.034	0.033	0.015	0.013
	(0.025)	(0.028)	(0.023)	–	(0.007)
σ_d	0.169	0.168	0.157	0.090	0.157
	(0.015)	(0.015)	(0.015)	–	(0.012)
b_d	0.148	0.415	−0.079	0.500	−0.575
	(0.196)	(0.092)	(0.285)	–	(0.108)
y_d	0.145	0.280	0.075	-0.350	0.088
	(0.149)	(0.116)	(0.268)	–	(0.090)
Long-Term Interest Rate					
μ_c	0.031	0.034	0.025	0.026	0.024
	(0.026)	(0.029)	(0.008)	–	(0.010)
σ_c	0.246	0.246	0.235	0.210	0.283
	(0.022)	(0.021)	(0.015)	–	(0.021)
a_c	0.936	0.965	0.917	0.960	0.918
	(0.043)	(0.035)	(0.040)	–	(0.044)
y_c	0.010	0.001	0.028	0.070	0.024
	(0.163)	(0.150)	(0.171)	–	(0.140)

The table reports the posterior means and standard deviations for both priors—non-informative and subjective. The third column reports the maximum likelihood estimates (MLE) and their standard errors (calculated via the observed Fisher information matrix). The rightmost columns report frequentist results obtained by two previous studies: Wilkie (1995) and Zhang *et al.* (2018, ZHS). Values in brackets represent standard deviations of the parameter's posterior distribution (for Bayesian inference) or standard errors (for frequentist methods).

parameter estimates obtained via MLE, the parameters found by Wilkie (1995) and the ones recovered by Zhang *et al.* (2018, hereafter ZHS).

For the inflation dynamics, the average parameters obtained by both priors are similar, although a_q is slightly larger for the subjective case. They are also consistent with those found via the MLE and those estimated by ZHS. The standard deviations and standard errors are also of the same order.

The posterior average for all the dividend yield parameters—either from the non-informative or the subjective case—is consistent with their frequentist counterparts, except for w_y. As explained earlier in Section 3, w_y should be positive because the level of the dividend should be proportional to the general level of prices in the economy. Nonetheless, our Bayesian estimate of w_y in the non-informative case as well as the one found by ZHS seem to be negative on average, whereas those of our subjective case and Wilkie are positive, that is, 0.524 and 0.5, respectively. The variation around these estimators is, however, rather large: 0.943 for the non-informative prior, 0.432 for the subjective prior, 0.534 for the MLE and 0.469 for ZHS. Indeed, Huber (1997) finds that the w_y is sensitive to outliers, which could explain to some extent these high standard deviations and the difference between our Bayesian results and the MLE estimates (ours, Wilkie's and ZHS's).

The means of the dividend index parameter distributions are somewhat similar for both priors. Their standard deviations are also consistent with one another. When compared to frequentist results, the parameters μ_d obtained with the Bayesian approach are consistent with that obtained with MLE, although it is fairly different from Wilkie's and ZHS's estimates. The different sample periods may explain this discrepancy. Parameter σ_d is consistent across the different methods, even though the one obtained by Wilkie is smaller than the others. As explained in Section 3, b_d should be positive in order to account for the fact that companies pay out only part of any additional earnings in a given year. The posterior averages of this parameter are indeed positive for our two priors, which departs from the frequentist results presented in Table 1: ZHS and the estimates obtained via the MLE are negative. The standard deviation for the Bayesian estimates of b_d are different; this hints towards different posterior distribution shapes for the two priors. The two average posterior values for y_d are consistent with those obtained with the MLE and estimated by ZHS, although the sign of y_c contradicts that of Wilkie. Note that the standard errors and deviations of this very parameter are large, which could explain the difference in sign to some extent.

Finally, for the long-term interest rate dynamics, the posterior means of the four parameters—μ_c, σ_c, a_c and y_c—are consistent with one another and with their frequentist counterparts. The uncertainty, estimated here as either the standard deviation or the standard error of a parameter, is also consistent across the different methods and for most parameters, although the actual shape of the posterior distributions for parameters μ_c and a_c are different for both prior choices.

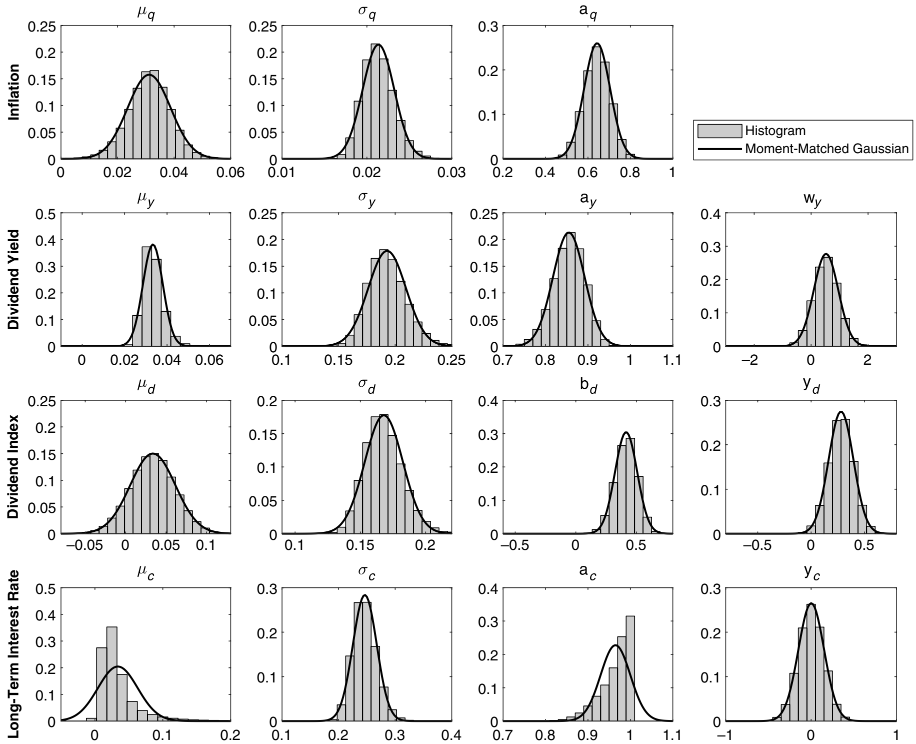

FIGURE 4: Posterior distributions for model parameter: subjective prior.
This figure exhibits the (marginal) posterior distribution for each parameter obtained using the MCMC methodology explained in Section 5 and in Appendix B. For each parameter, we obtain a sample of 100,000 values from which we construct a histogram. In this study, we focus on post-World War II data, that is, data for the year 1945 and after. In addition to determining the histogram, we also obtain the moment-matched Gaussian distribution (solid line). The second column of Table 1 reports the values of such means and standard deviations needed to match the moment of the normal distributions.

Figure 4 plots the density estimates of the parameters for the subjective prior distribution (the results for the non-informative are available in the Supplementary Material). The posterior histograms of the simulated parameters are complemented by moment-matched Gaussian distributions. Even though the posterior standard deviations calculated by our Bayesian approach are consistent with the standard errors obtained by the MLE and in ZHS for most parameters, the actual shape of the posterior distribution might, however, be different from the asymptotic parameter distribution under the MLE framework, for example, the distribution of μ_y, μ_c and a_c does not appear to be normally distributed for the subjective prior. Indeed, parameter uncertainty is accounted for in a more consistent manner by the Bayesian inference developed in Sections 4 and 5 and therefore allows us to understand the actual posterior distribution of these parameters.

In summary, both priors lead to similar posterior distributions for all parameters except w_y, b_d, μ_c and a_c. For the sake of brevity, we will use the subjective prior in the remainder of the paper. For robustness purposes, however, all the figures and tables obtained using the non-informative prior are available in Section SM.B of the Supplementary Material.

6.3. Bayesian prediction

The main advantage of the Bayesian inference method used in this study—in addition to its consistency—is that we can readily make Bayesian predictions using only past data. Specifically, if X represents the past data (used in the estimation) and Y the future values of the four economic values, that is, $Y = \{i(t), y(t), d(t), c(t)\}_{t=T+1}^{T+\tau}$, then we can find the posterior-predictive density of these future observations based on past information:

$$f(Y|X) = \frac{f(Y,X)}{f(X)} = \frac{\int f(Y,X,\Theta)\,d\Theta}{f(X)} = \frac{\int f(Y,X|\Theta)\,\pi(\Theta)\,d\Theta}{f(X)}$$
$$= \frac{\int f(Y|\Theta) f(X|\Theta)\,\pi(\Theta)\,d\Theta}{f(X)} = \int f(Y|\Theta)\,\pi(\Theta|X)\,d\Theta.$$

The Bayesian estimation technique in Section 5 can be thought of as an intermediate step to build the posterior-predictive density: it allows us to obtain parameter sets that are used to generate samples of Y in a consistent manner. To build the Bayesian prediction, each posterior parameter set draw—those kept to build the posterior distribution—is used to draw ten paths of the four processes, making a total of 1,000,000 paths.

In this section, we consider two cases: a first case in which there is parameter uncertainty and a second case without parameter uncertainty, that is, fixed parameters. The uncertainty is characterized by the posterior distribution found previously; in the case without parameter uncertainty, the parameters are fixed to the sample average, for example, $\frac{1}{M}\sum_{j=1}^{M}\mu_q^{(j)}$ for the long-run level of the inflation rate.

To characterize the impact of parameter uncertainty for the four variables of interest, we construct funnels of doubt similar to those presented in Figure 1. These funnels of doubt are constructed on four different variables, each related to one of the processes described in Wilkie's model:

1. The total inflation index growth, defined as $\sum_{t=T+1}^{T+\tau} i(t) = \log\left(\frac{I(T+\tau)}{I(T)}\right)$, where τ is the time horizon.
2. The dividend yield, $y(T+\tau)$.
3. The total dividend index growth, defined as $\sum_{t=T+1}^{T+\tau} d(t) = \log\left(\frac{D(T+\tau)}{D(T)}\right)$.
4. The long-term interest rate, $c(T+\tau)$.

Figure 5 presents the median and the 95% confidence interval for the four variables and for both cases, for example, with (leftmost column) and without parameter uncertainty (rightmost column). The median value is similar in both

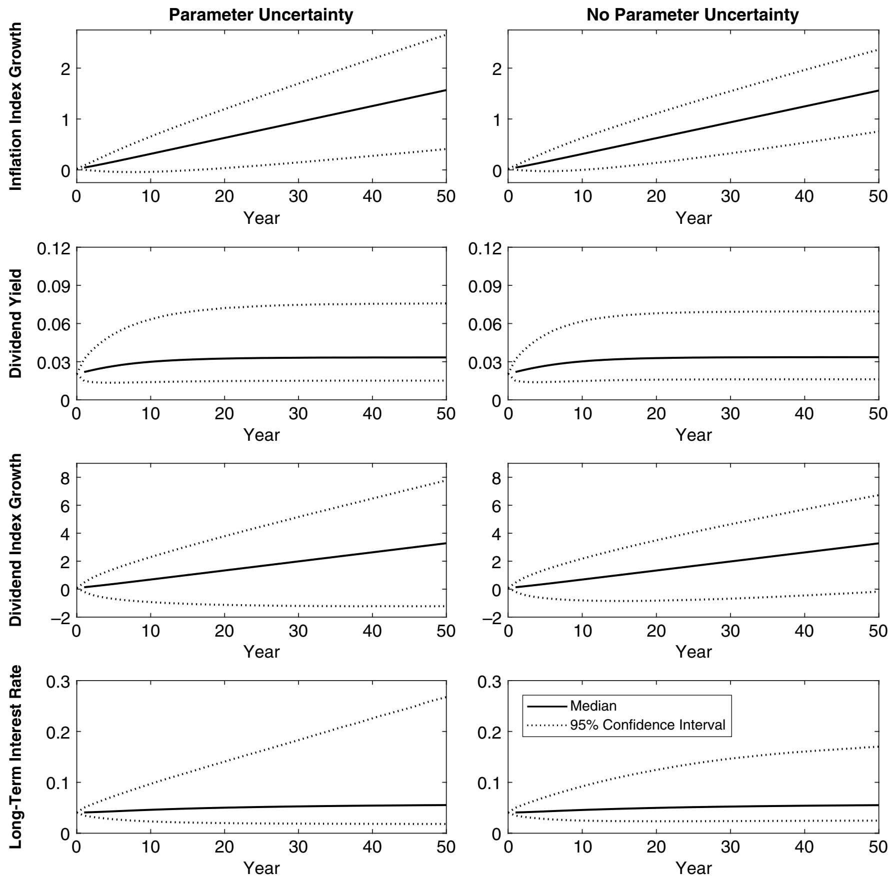

FIGURE 5: Funnels of doubt for the four economic variables: subjective prior.
This figure shows funnels of doubt with (left column) and without (right column) parameter uncertainty, and for each of the four economic variables under study: the total inflation growth, the dividend yield, the total dividend growth and the long-term interest rate. These plots present the median (solid) as well as the 95% confidence interval (dashed line) for each year. The leftmost figures present the funnels of doubt in the case of parameter uncertainty based on the MCMC methodology explained in Section 5 and in Appendix B. For the rightmost figures, we use no parameter uncertainty: instead of using the posterior sets of parameters, we use the average value of each parameter, as given in Table 1.

cases, meaning that the broad behaviour of these economic variables is similar. However, the tails of each of these distributions are thinner in the case without parameter uncertainty. Indeed, the fact that the parameters are uncertain seems to significantly increase the variability: the 95% confidence intervals shown in Figure 5 are wider when considering uncertain parameter. The parameter uncertainty has a larger impact on the total inflation index growth, on the dividend index growth and on the long-term interest rate. On the other hand, the impact is more modest for the dividend yield, albeit there is still a difference between the two cases.

TABLE 2

DESCRIPTIVE STATISTICS AND QUANTILES OF ECONOMIC VARIABLES OVER DIFFERENT TIME HORIZONS: SUBJECTIVE PRIOR.

	Parameter Uncertainty		No Parameter Uncertainty	
	Average	Standard Deviation	Average	Standard Deviation
Total Inflation Index Growth				
5 Years	0.1595	0.0991	0.1593	0.0945
10 Years	0.3134	0.1755	0.3131	0.1600
15 Years	0.4691	0.2375	0.4687	0.2083
25 Years	0.7808	0.3424	0.7803	0.2818
50 Years	1.5598	0.5655	1.5595	0.4109
100 Years	3.1188	0.9703	3.1185	0.5896
Dividend Yield				
5 Years	0.0281	0.0099	0.0283	0.0096
10 Years	0.0322	0.0128	0.0323	0.0122
15 Years	0.0342	0.0143	0.0343	0.0131
25 Years	0.0358	0.0154	0.0356	0.0137
50 Years	0.0364	0.0160	0.0359	0.0138
100 Years	0.0365	0.0161	0.0359	0.0138
Total Dividend Index Growth				
5 Years	0.3701	0.5417	0.3698	0.5202
10 Years	0.6873	0.8191	0.6873	0.7623
15 Years	1.0097	1.0454	1.0101	0.9469
25 Years	1.6594	1.4369	1.6578	1.2342
50 Years	3.2803	2.2852	3.2769	1.7598
100 Years	6.5204	3.8310	6.5182	2.5004
Long-Term Interest Rate				
5 Years	0.0449	0.0116	0.0446	0.0110
10 Years	0.0495	0.0196	0.0489	0.0176
15 Years	0.0537	0.0271	0.0526	0.0229
25 Years	0.0609	0.0415	0.0582	0.0308
50 Years	0.0732	0.0670	0.0654	0.0403
100 Years	0.0856	0.0913	0.0688	0.0440

This table presents descriptive statistics with and without parameter uncertainty and for each of the four economic variables under study: the total inflation index growth, the dividend yield, the total dividend index growth and the long-term interest rate. These statistics are calculated for different time horizons: 5, 10, 15, 25, 50 and 100 years. In the case without parameter uncertainty, instead of using the posterior sets of parameters, we use the average value of each parameter as given in Table 1 to generate 1,000,000 observations for each year.

Table 2 shows two descriptive statistics for each variable under study and for different time horizons. In each case, we compute the average and standard deviation for both cases, for example, with and without parameter uncertainty. In general, the averages are similar in both cases and the standard deviations differ largely. For the total inflation growth, the standard deviation is between

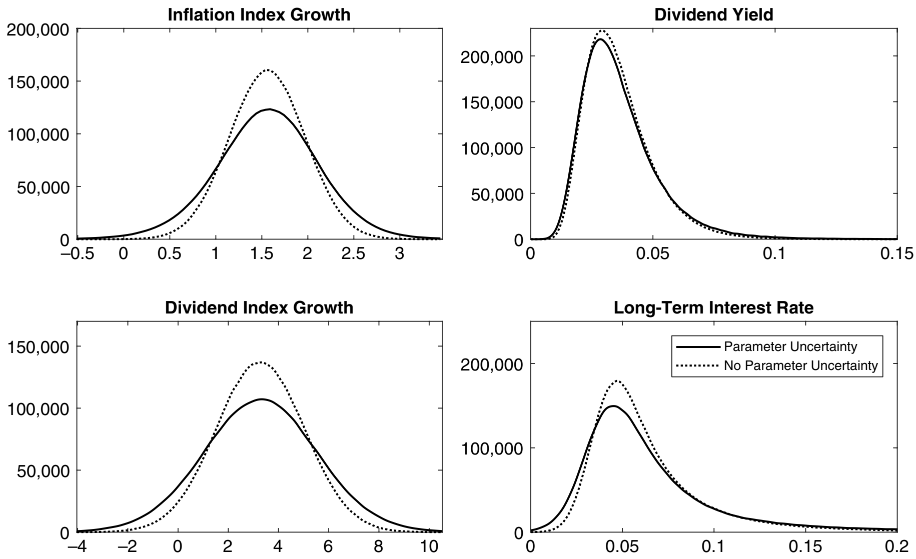

FIGURE 6: Kernel smoothed densities of the economic variables for a period of 50 years: subjective prior. This figure shows densities with (left column) and without (right column) parameter uncertainty, and for each of the four economic variables under study: the total inflation growth, the dividend yield, the total dividend growth and the long-term interest rate. The leftmost figures present the densities in the case of parameter uncertainty based on the MCMC methodology explained in Section 5 and in Appendix B. For the rightmost figures, we use no parameter uncertainty: instead of using the posterior sets of parameters, we use the average value of each parameter, as given in Table 1.

5% and 65% larger for the time horizons considered in Table 2. The standard deviation is between 3% and 17% larger when we consider parameter uncertainty for the dividend yield. The total dividend index growth's standard deviation is between 4% and 53% larger once parameter uncertainty is accounted for. Finally, for the long-term interest rate, we have standard deviations under the parameter uncertainty assumption that could be up to 108% larger than for the case without uncertainty.

Figure 6 complements Figure 5 and Table 2. It shows kernel smoothed densities for the four economic variables for a period of 50 years, that is, $\tau = 50$. Generally speaking, the densities that allow for parameter uncertainty have fatter tails than those without. The impact of parameter uncertainty is more significant for the dividend index growth, the inflation index growth as well as the long-term interest rate. It is less considerable for the dividend yield, albeit still present.

A conjecture that could explain these results is the cascade structure explained in Section 3. The cascade structure is indeed key in capturing the dependence among the various factors modelled in the Wilkie model. In addition to the obvious linkages among the four economic variables, there are further interactions when parameter uncertainty is allowed for. The fact that some variables sit at the top of the cascade creates the so-called compounded (or *magnified*) parameter uncertainty that makes the whole model even more

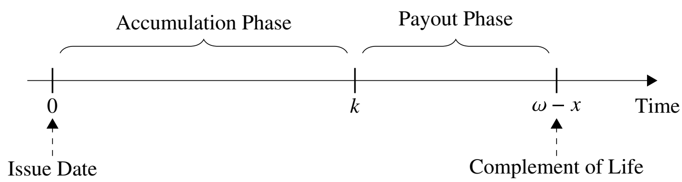

FIGURE 7: A diagram illustrating the two phases of the annuity design.
This figure summarizes the two phases of our annuity design: the accumulation phase (from issue date, 0, to k) and the payout phase (from k to the complement of life, $\omega - x$). The figure also shows the valuation time s, which could be before or after k (in the figure, it is before k).

uncertain. In other words, the uncertainty coming from upper levels of the cascade gets compounded into the lower levels of the structure. Specifically, the parameter uncertainty associated with the inflation and the dividend yield is included in the dividend index and on the long-term interest rate, making the latter even more uncertain.

Overall, parameter uncertainty has an impact on the funnels of doubt and, more broadly, on the dispersion of the economic variables under study. It goes without saying that this uncertainty can have a considerable impact on the tail behaviour of the inflation, the dividend yield, the dividend index and the long-term interest rate future distributions. One question still remains, however: is this uncertainty relevant from an actuarial perspective?

7. Actuarial implications

In this final section, we assess the impact of parameter uncertainty in the case of a life insurer that sells annuities. We consider in this illustration a homogeneous portfolio of identical annuities issued to a group of N policyholders, all aged x. Each policyholder buys a k-year deferred whole life due annuity that pays b at the beginning of each year to each annuitant. To obtain this benefit, the annuitant must make an annual payment of π, the annual level premium, payable at the beginning of each year during the accumulation phase—or, in other words, during the first k years of the contract. If the annuitant dies during the accumulation phase, the annual premiums are refunded, indexed to inflation (see Figure 7).

To account for the time-value of money, a stochastic rate of return is needed. In this study, we assume that the portfolio is rebalanced at the end of each year and that half of the portfolio is invested in the dividend index and the other half is invested in a portfolio of Aaa-rated 20-year bonds:

$$P(t) = \left(\frac{1}{2} e^{d(t)} \left(1 + y(t)\right) + \frac{1}{2} \frac{\left(1 + c(t-1)\right)^{20}}{\left(1 + c(t)\right)^{20}} \right) P(t-1),$$

where $P(t)$ is the investment portfolio value at time t. The total return on the dividend index contains in fact an ex-dividend component as well as a dividend payment. Recall that the dividend yield series is constructed by taking the sum of the gross dividend payments over a given year and dividing it by the value of the index at the end of the year, that is,

$$y(t) = \frac{Y(t)}{D(t)} \quad \Rightarrow \quad y(t)\, D(t) = Y(t),$$

where $Y(t)$ is the gross dividend payments over a given year (i.e., from $t-1$ to t). Therefore, the total return of an investment in the dividend index is given by

$$\begin{aligned}\frac{D(t)+Y(t)}{D(t-1)} &= \frac{D(t)+D(t)y(t)}{D(t-1)} = \frac{D(t)(1+y(t))}{D(t-1)} \\ &= \frac{D(t-1)e^{d(t)}(1+y(t))}{D(t-1)} = e^{d(t)}(1+y(t)).\end{aligned}$$

The (continuously compounded) stochastic rate of return is given by $\delta(t) = \log(P(t)/P(t-1))$.

We assume that the annuity portfolio contains only females aged x whose mortality is random and follows a Gompertz model fitted to the US data. The Gompertz survival function is approximated by $S(x) = \exp\left(-e^{0.0938(x-87.047)}\right)$. These parameters are estimated by Pflaumer (2011) using the 2006 National Vital Statistics Reports life table for females in the US. We further assume that the future lifetimes of the policyholders in the portfolio are independent and identically distributed.

Consider the following life-contingent indicator variables:

$$\begin{aligned} \mathcal{L}_{t,j} &= \mathbf{1}_{\{\text{Policyholder } j \text{ is Alive at the Beginning of Year } t\}} \\ \text{and} \quad \mathcal{D}_{t,j} &= \mathbf{1}_{\{\text{Policyholder } j \text{ Dies in Policy-Year } t\}}. \end{aligned}$$

At the inception of the contract, the (random) net cash outflow at time t is given by

$$C_t = \sum_{j=1}^{N} \left[\pi \ddot{s}_{\overline{t}|\, i}\, \mathcal{D}_{t,j} \mathbf{1}_{\{t<k\}} + b\mathcal{L}_{t,j} \mathbf{1}_{\{t \geq k\}} - \pi \mathcal{L}_{t,j} \mathbf{1}_{\{t<k\}}\right],$$

where $\ddot{s}_{\overline{t}|\, i}$ is a t-year annuity due accumulated at the inflation rate. The time-t net cash outflow is indeed defined as the accumulated premiums if the annuitant dies before the end of the accumulation phase, plus the annuity payment during the payout phase if the annuitant is still alive, minus the premium paid during the accumulation phase if the annuitant is alive.

The loss at issue random variable is therefore given by the discounted value of the (random) net cash outflows:

$$\text{Loss}_x = \sum_{t=0}^{\omega-x} C_t e^{-I(T,T+t)}, \tag{7}$$

https://doi.org/10.1017/asb.2019.6 Published online by Cambridge University Press

where

$$l(T, T+t) = \begin{cases} \sum_{s=T+1}^{T+t} \delta(s) & \text{if } t > 0 \\ 0 & \text{if } t = 0 \end{cases}$$

similar to the notation used by Marceau and Gaillardetz (1999) and Nolde and Parker (2014). The product annual premium π is calculated by using the actuarial equivalence principle, that is, π such that $\mathbb{E}[\text{Loss}_x] = 0$. Different loadings θ on the premium are considered: 0%, 5%, 10% and 20%.

In the example, we specifically consider the case where 100 contracts are sold, that is, $N = 100$. Three different initial ages are considered: 35, 45 and 55 years old. The annual annuity payoff in the payout phase is assumed to be 1000. We select k, the deferred period, such that the payout period starts at 65, regardless of the actual age at issue, x.

To obtain the stochastic rate of return and the inflation rate, we use the same 1,000,000 paths generated in Section 6.3. Then, for each of these economic scenarios (that accounts for parameter uncertainty), life-contingent indicators $\mathcal{L}_{t,j}$ and $\mathcal{D}_{t,j}$ are generated from a multinomial distribution, as mortality is assumed to be random in this study. We do exactly the same thing for the case without parameter uncertainty using the 1,000,000 paths found for this case, as explained in Section 6.3.

Figure 8 shows cumulative distribution function (cdf) for the case with parameter uncertainty (solid line) and without (dashed line). This figure considers four different levels of loading θ and the three different initial ages. The tails of the cdf obtained under parameter uncertainty are significantly heavier than those generated without. At first sight, parameter uncertainty increases the risk and creates heavier tails that result in more extreme scenarios and, if we focus on the right tail, more extreme losses.

Table 3 summarizes some of the descriptive statistics and risk measures of the loss random variable (per annuity sold)—along with the annual premium—for different levels of θ and various ages x. The premium obtained by the equivalence principle, π, is always greater when we account for parameter uncertainty: it is approximately 21% greater when $x = 35$, 12% if $x = 45$ and 6% if $x = 55$. This is also the case for the other risk measures, namely, the 95th quantile (the so-called VaR) and the CTE at 95%: they are between 23% and 51% higher when considering parameter uncertainty.

Therefore, it appears that the parameter uncertainty in Wilkie's ESG changes the risk profile in a significant way. All risk measures are substantially larger when parameter uncertainty is taken into account—with increases above 20% in all cases. Additionally, in the case of this ESG, most risk calculations that do not account for parameter uncertainty are thus downward biased. This conclusion is therefore highly relevant for ESG end-users—and, more generally, actuaries using such models.

This result must remain true for other more advanced ESGs to some extent. Indeed, more advanced ESGs include extra parameters, making the statistical estimation of such models cumbersome—therefore creating more uncertainty around each parameter's point estimates.

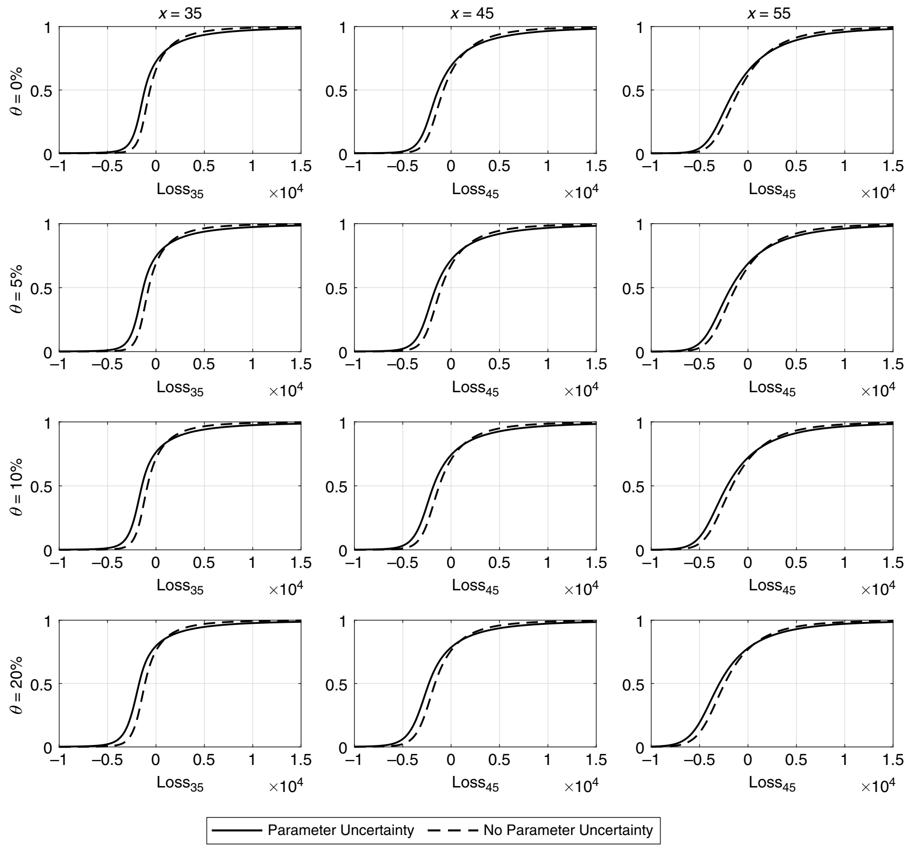

FIGURE 8: Cumulative distribution functions of the loss at issue as a function of different initial ages and levels of loading for a portfolio of 100 annuities: subjective prior.

This figure shows cumulative distribution functions (cdf) for different initial ages (35, 45 and 55 years old) and levels of loading (0%, 5%, 10% and 20%). Two cases are shown in this figure: parameter uncertainty (solid line) and no parameter uncertainty (dashed line). For the parameter uncertainty case, we use the MCMC methodology explained in Section 5. Under no parameter uncertainty, we use the average value of each parameter, as given in Table 1. Associated with each of these paths, we generate the indicator variables $\mathcal{L}_{t,j}$ and $\mathcal{D}_{t,j}$ for 100 lives using the Gompertz model fitted to the US data (Pflaumer, 2011). Then, based on these simulations, we obtain the loss at issue under each scenario to obtain a sample of this random variable.

Finally, to make a parallel with Section 2, the parameter uncertainty considered in this problem is not inconsequential because it greatly increases the risk. The uncertainty uncovered in Section 6 therefore has an important impact on the risk calculation; omitting such uncertainty could (negatively) bias an actuary's risk assessment. Therefore, to answer our research question, parameter uncertainty seems to matter very much.

7.1. Robustness test: Using more data in the estimation

One caveat of this study is the use annual data to perform the Bayesian parameter inference. We did so to make our estimation results comparable to prior

TABLE 3

DESCRIPTIVE STATISTICS AND QUANTILES OF LOSS AT ISSUE FOR A PORTFOLIO OF 100 ANNUITIES: SUBJECTIVE PRIOR.

	Parameter Uncertainty				No Parameter Uncertainty			
	Premium $\pi(1+\theta)$	Standard Deviation	$Q(0.95)$	CTE(0.95)	Premium $\pi(1+\theta)$	Standard Deviation	$Q(0.95)$	CTE(0.95)
$\boldsymbol{x = 35}$								
$\theta = 0\%$	230.52	5280.95	6381.56	15469.82	190.53	2577.60	4264.38	10706.27
$\theta = 5\%$	242.05	5249.63	6110.80	15136.55	200.05	2563.02	4063.73	10390.29
$\theta = 10\%$	253.57	5219.77	5836.43	14805.24	209.58	2549.73	3866.88	10080.81
$\theta = 20\%$	276.62	5167.47	5298.01	14149.86	228.63	2527.11	3470.26	9449.84
$\boldsymbol{x = 45}$								
$\theta = 0\%$	435.14	5582.00	7822.73	16800.36	386.99	3277.64	5658.41	12721.60
$\theta = 5\%$	456.90	5537.88	7430.29	16341.56	406.33	3254.45	5339.50	12281.79
$\theta = 10\%$	478.65	5498.38	7031.16	15887.05	425.68	3233.31	5024.19	11832.35
$\theta = 20\%$	522.17	5430.86	6256.86	14989.01	464.38	3197.37	4396.81	10950.82
$\boldsymbol{x = 55}$								
$\theta = 0\%$	1046.70	5556.75	8968.00	16882.15	983.41	3898.45	7020.53	13860.30
$\theta = 5\%$	1099.04	5517.61	8425.98	16292.12	1032.58	3870.52	6535.73	13280.38
$\theta = 10\%$	1151.37	5480.61	7885.42	15704.85	1081.75	3845.26	6052.60	12697.51
$\theta = 20\%$	1256.04	5413.87	6827.80	14538.73	1180.10	3802.94	5096.78	11559.27

This table summarizes descriptive statistics and risk measures of the loss at issue (per annuity sold) for different initial ages (35, 45 and 55 years old) and levels of loading (0%, 5%, 10% and 20%). Two cases are used in this table: parameter uncertainty (leftmost columns) and no parameter uncertainty (rightmost columns). For the parameter uncertainty case, we use the MCMC methodology explained in Section 5 and in Appendix B. Under no parameter uncertainty, we use the average value of each parameter, as given in Table 1. Associated with each of these paths, we generate the indicator variables $\mathcal{L}_{t,j}$ and $\mathcal{D}_{t,j}$ for 100 lives using the Gompertz model fitted to the US data (Pflaumer, 2011). Then, based on these simulations, we obtain the loss at issue under each scenario to determine a sample of this random variable. The table shows, for both cases, the annual premium paid by the annuitant, the standard deviation of the losses at issue, the 95th quantile of the loss distribution $Q(0.95)$ and the CTE at the 95% level.

studies in the literature (e.g. Wilkie, 1995; Zhang *et al.*, 2018). Yet, this fairly reduces the amount of data available in the estimation state—recall that only 71 data points were available for each time series—and this could have a not-so-marginal impact on the uncertainty of parameters: having a modest number of observations makes the parameter more uncertain, broadly speaking.

For the sake of robustness, we assess once more the impact of parameter uncertainty on a portfolio of annuities, but this time we consider economic time series at a higher frequency: we use quarterly data instead of annual time series. Still using the subjective prior—for which the parameters of the *a priori* distributions of Table 5 were adjusted to account for the fact that we have quarterly data—we rerun the estimation procedure applied in Section 6 and generate samples of the future values of the economic variables.

TABLE 4

DESCRIPTIVE STATISTICS AND QUANTILES OF LOSS AT ISSUE FOR A PORTFOLIO OF 100 ANNUITIES: SUBJECTIVE PRIOR AND QUARTERLY DATA.

	Parameter Uncertainty				No Parameter Uncertainty			
	Premium $\pi(1+\theta)$	Standard Deviation	$Q(0.95)$	CTE(0.95)	Premium $\pi(1+\theta)$	Standard Deviation	$Q(0.95)$	CTE(0.95)
$\boldsymbol{x=35}$								
$\theta=0\%$	195.44	2003.19	3279.06	6148.01	178.77	1255.06	2270.86	4861.76
$\theta=5\%$	205.21	1986.53	3087.56	5935.86	187.71	1246.53	2113.11	4654.87
$\theta=10\%$	214.98	1970.65	2898.02	5724.64	196.65	1238.83	1955.04	4449.76
$\theta=20\%$	234.53	1941.15	2520.52	5304.88	214.53	1225.97	1640.70	4039.07
$\boldsymbol{x=45}$								
$\theta=0\%$	381.29	2368.74	4117.58	7174.87	360.81	1683.00	3060.71	6024.51
$\theta=5\%$	400.36	2345.36	3823.61	6854.50	378.85	1669.17	2804.92	5711.88
$\theta=10\%$	419.42	2323.32	3530.60	6535.79	396.89	1656.67	2546.89	5399.41
$\theta=20\%$	457.55	2282.98	2950.29	5903.04	432.98	1635.80	2037.18	4778.13
$\boldsymbol{x=55}$								
$\theta=0\%$	945.05	2713.26	4928.75	7894.87	917.82	2146.70	3936.93	7025.32
$\theta=5\%$	992.31	2688.28	4486.60	7426.84	963.71	2129.99	3522.03	6563.80
$\theta=10\%$	1039.56	2665.04	4044.98	6961.13	1009.60	2115.15	3109.28	6099.49
$\theta=20\%$	1134.07	2623.67	3166.10	6035.91	1101.38	2091.24	2292.35	5178.84

This table summarizes descriptive statistics and risk measures of the loss at issue (per annuity sold) for different initial ages (35, 45 and 55 years old) and levels of loading (0%, 5%, 10% and 20%) using quarterly data. For more details, please refer to the caption of Table 3.

Similar to Table 3, Table 4 summarizes some of the descriptive statistics and risk measures of the loss at inception random variable (per annuity sold). Overall, the premiums are still greater when we allow for parameter uncertainty, although the difference is slightly less sizeable: the premium is approximately 9% greater when $x=35$, 6% if $x=45$ and 3% if $x=55$. The VaR measures and the CTEs at 95% using quarterly data are generally smaller than those using annual data, although the relative increases are larger in the quarterly case. Specifically, the 95th quantile (CTE estimates) is 48% (29%) greater when $x=35$, 39% (21%) when $x=45$ and 30% (14%) when $x=55$. Therefore, it appears that even if we use more data—in our case, four times more—the uncertainty on the parameters remains and still has dire consequences on the risk profile of the portfolio of annuities.

When comparing the results of Table 4 to those of Table 3, we observe a substantial twofold decrease in most of the risk measures (i.e., standard deviation, VaR and CTE). Even though the average value of the parameters is not that different in the quarterly case (vis-à-vis the annual case), small differences in the parameters can compound over time and create discrepancies in the long haul. Specifically, in our case, the overall long-run risk in the dividend

https://doi.org/10.1017/asb.2019.6 Published online by Cambridge University Press

index and the long-term interest rate—two crucial variables in determining the portfolio return—is lower when using quarterly data.

For the leftmost columns of Tables 3 and 4, the uncertainty on the parameters can be quite different when using different frequencies, which could also explain partly the differences mentioned above (e.g., when appropriately scaled, the standard deviation of μ_q is reduced by about half, from 0.0076 with annual data to 0.0043 with quarterly observations). Indeed, having more data points yields less uncertainty on the parameters, which translates into a loss distribution that is less dispersed. Yet, the impact of parameter uncertainty does not vanish: even if you increase the sampling frequency, there is still uncertainty on the parameters.

One slight note of caution, however, for those who would be tempted to increase considerably the sampling frequency: even though increasing the frequency makes sense from a statistical perspective, it could be misguided. ESGs are mainly used for generating scenarios over the long term and using *high*-frequency series may capture features that are unrelated to the future evolution of the economic variables under scrutiny. For instance, increasing the frequency could raise the process uncertainty by capturing *noise*—similar in essence to the microstructure noise in intraday returns, as described in Zhang *et al.* (2005), among others—instead of the underlying risk. Hence, higher sampling frequencies do not necessarily lead to more confidence. Finding the adequate frequency to use given a specific application is an open question and is left for future research.

8. Concluding remarks

A new estimation methodology for the Wilkie model was proposed based on the Bayesian paradigm. Using MCMC methods, we obtained posterior distributions of the model parameters; we did so for two different *a priori* distributions: a non-informative prior and a subjective one. Using post-World War II US data, we found the posterior densities for each of the model parameters, consistent with the available data. We also found funnels of doubt for the four economic quantities under scrutiny using the posterior-predictive density: parameter uncertainty adds a great deal of uncertainty to the total risk profile, and the funnels of doubt are wider when we consider parameter uncertainty.

In an additional application, we considered a portfolio of annuities in order to assess the relevance of parameter uncertainty from an actuarial perspective. In all the cases considered, we found that the distribution of the loss at issue that accounted for parameter uncertainty had heavier tails than that without uncertainty: the risk profile of this portfolio of annuities changes drastically when parameter uncertainty is accounted for consistently. Not accounting for this uncertainty could therefore have dire consequences for an insurer's risk assessment.

The bulk of the results obtained in this study should remain accurate for more advanced ESGs because models that are less parsimonious than Wilkie's model will lead to parameters that are more difficult to identify—and, therefore, to posterior distributions that are less precise, generally speaking. Yet, other—more complex—ESGs might be more difficult to cast into a Bayesian model. A rather straightforward proposition in these cases is to use the MLE-based standard errors as proxies for the posterior standard deviations, as well as the asymptotic normality to approximate the distribution of said parameters. Notice, however, that the distribution of some parameters is far from being normal (i.e., Figure 4). Thus, using the asymptotic distribution might distort the uncertainty quantification. For the Wilkie model, the MLE-based approximation yields biased results when compared to our Bayesian method (these untabulated results are available upon request). The end-product of such an approximation should, therefore, be used with precaution. The Bayesian methodology put forward in this study is indeed a good starting point to assess parameter uncertainty in more complex ESGs, although it might complicate the risk assessment methodology considerably.

Finally, one source of uncertainty that we have not assessed is model risk; even though this risk seems to be of lesser importance than parameter uncertainty (Cairns, 2000), it would be interesting to quantify its impact in the case of ESGs. This question is also left for future research.

Supplementary Material

To view supplementary material for this article, please visit http://dx.doi.org/10.1017/asb.2019.6.

Acknowledgments

For their useful comments, the author would like to thank Daniel Bauer (the editor), two anonymous referees, Gary Parker, Tim Swartz and seminar participants at the 2018 Actuarial Research Conference. This research was in part supported by the Society of Actuaries, the Natural Sciences and Engineering Research Council of Canada, and Simon Fraser University.

References

Ahlgrim, K.C., D'Arcy, S.P. and Gorvett, R.W. (2004) *Modeling of economic series coordinated with interest rate scenarios.* Technical report, Casualty Actuarial Society and the Society of Actuaries.

Ahlgrim, K.C., S.P. D'Arcy, and Gorvett, R.W. (2008) A comparison of actuarial financial scenario generators. *Variance* **2**, 111–134.

Bernardo, J.M. and Smith, A.F. (2001) *Bayesian Theory*. New York, NY, USA: Wiley.

Bégin, J.-F. (2016) Deflation risk and implications for life insurers. *Risks,* **4**, 46.

Cairns, A.J. (2000) A discussion of parameter and model uncertainty in insurance. *Insurance: Mathematics and Economics,* **27**, 313–330.

CAIRNS, A.J., BLAKE, D. and DOWD, K. (2006) A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. *Journal of Risk and Insurance,* **73**, 687–718.
CHAN, T. (1998) Some applications of Lévy processes to stochastic investment models for actuarial use. *ASTIN Bulletin,* **28**, 77–93.
CHAN, W.-S. (2002) Stochastic investment modelling: a multiple time-series approach. *British Actuarial Journal,* **8**, 545–591.
CHAN, W.-S., WONG, A.C. and TONG, H. (2004) Some nonlinear threshold autoregressive time series models for actuarial use. *North American Actuarial Journal,* **8**, 37–61.
FELLINGHAM, G.W., KOTTAS, A. and HARTMAN, B.M. (2015) Bayesian nonparametric predictive modeling of group health claims. *Insurance: Mathematics and Economics,* **60**, 1–10.
FISHER, I. (1930) *The Theory of Interest*. New York, NY, USA: The Macmillan Co.
GELMAN, A. and RUBIN, D.B. (1992) Inference from iterative simulation using multiple sequences. *Statistical Science,* **7**, 457–472.
GEOGHEGAN, T., CLARKSON, R., FELDMAN, K., GREEN, S., KITTS, A., LAVECKY, J., ROSS, F., SMITH, W. and TOUTOUNCHI, A. (1992) Report on the Wilkie stochastic investment model. *Journal of the Institute of Actuaries,* **119**, 173–228.
GILL, J. and FREEMAN, J.R. (2013) Dynamic elicited priors for updating covert networks. *Network Science,* **1**, 68–94.
HARTMAN, B.M. and GROENDYKE, C. (2013) Model selection and averaging in financial risk management. *North American Actuarial Journal,* **17**, 216–228.
HARTMAN, B.M. and HEATON, M.J. (2011) Accounting for regime and parameter uncertainty in regime-switching models. *Insurance: Mathematics and Economics,* **49**, 429–437.
HARTMAN, B.M., RICHARDSON, R. and BATEMAN, R. (2017) *Parameter uncertainty*. Technical report, Casualty Actuarial Society, Canadian Institute of Actuaries, and Society of Actuaries Report.
HASTINGS, W.K. (1970) Monte Carlo sampling methods using Markov chains and their applications. *Biometrika,* **57**, 97–109.
HAYNE, R.M. (1999) Modeling parameter uncertainty in cash flow projections. In *CAS Forum*, pp. 133–151.
HECKMAN, P.E. and MEYERS, G.G. (1983) The calculation of aggregate loss distributions from claim severity and claim count distributions. *Proceedings of the Casualty Actuarial Society*, **70**, 22–61.
HUBER, P. (1997) A review of Wilkie's stochastic asset model. *British Actuarial Journal,* **3**, 181–210.
MARCEAU, E. and GAILLARDETZ, P. (1999) On life insurance reserves in a stochastic mortality and interest rates environment. *Insurance: Mathematics and Economics,* **25**, 261–280.
METROPOLIS, N., ROSENBLUTH, A.W., ROSENBLUTH, M.N., TELLER, A.H. and TELLER, E. (1953) Equation of state calculations by fast computing machines. *Journal of Chemical Physics,* **21**, 1087–1092.
MEYERS, G. and SCHENKER, N. (1983) Parameter uncertainty in the collective risk model. *PCAS LXX,* p. 111.
NOLDE, N. and PARKER, G. (2014) Stochastic analysis of life insurance surplus. *Insurance: Mathematics and Economics,* **56**, 1–13.
O'HAGAN, A. (1998) Eliciting expert beliefs in substantial practical applications. *Journal of the Royal Statistical Society: Series D (The Statistician),* **47**, 21–35.
PEDERSEN, H., CAMPBELL, M.P., CHRISTIANSEN, S.L., COX, S.H., FINN, D., GRIFFIN, K., HOOKER, N., LIGHTWOOD, M., SONLIN, S.M. and SUCHAR, C. (2016) *Economic scenario generators: a practical guide*. Technical report, Society of Actuaries.
PFLAUMER, P. (2011) Methods for estimating selected life table parameters using the Gompertz distribution. In *Proceedings of the Joint Statistical Meetings, Social Statistics Section, Miami Beach*, pp. 733–747.
REDINGTON, F.M. (1952) Review of the principles of life-office valuations. *Journal of the Institute of Actuaries,* **78**, 286–340.
SAHIN, S., CAIRNS, A. and KLEINOW, T. (2008) Revisiting the Wilkie investment model. In *18th International AFIR Colloquium, Rome*, pp. 1–24.

SHERRIS, M., TEDESCO, L. and ZEHNWIRTH, B. (1997) Stochastic investment models: unit roots, cointegration, state space and GARCH models. In *Actuarial Research Clearing House*, Vol. 1, pp. 95–144.
UPTON, G. and COOK, I. (2014) *A Dictionary of Statistics*. New York: Oxford University Press.
WHITTEN, S. and THOMAS, R.G. (1999) A non-linear stochastic asset model for actuarial use. *British Actuarial Journal,* **5**, 919–953.
WILKIE, A., SAHIN, S., CAIRNS, A. and KLEINOW, T. (2011) Yet more on a stochastic economic model, part 1: updating and refitting, 1995 to 2009. *Annals of Actuarial Science,* **5**, 53–99.
WILKIE, A.D. (1986) A stochastic investment model for actuarial use. *Transactions of the Faculty of Actuaries,* **39**, 341–403.
WILKIE, A.D. (1995) More on a stochastic asset model for actuarial use. *British Actuarial Journal,* **1**, 777–964.
WILKIE, A.D. and SAHIN, S. (2016) Yet more on a stochastic economic model: part 2: initial conditions, select periods and neutralising parameters. *Annals of Actuarial Science,* **10**, 1–51.
WILKIE, A.D. and SAHIN, S. (2017a) Yet more on a stochastic economic model: part 3A: stochastic Interpolation: Brownian and Ornstein–Uhlenbeck (OU) Bridges. *Annals of Actuarial Science,* **11**, 74–99.
WILKIE, A.D. and SAHIN, S. (2017b) Yet more on a stochastic economic model: part 3B: stochastic bridging for retail prices and wages. *Annals of Actuarial Science,* **11**, 100–127.
WILKIE, A.D. and SAHIN, S. (2017c) Yet more on a stochastic economic model: part 3C: stochastic bridging for share yields and dividends and interest rates. *Annals of Actuarial Science,* **11**, 128–163.
WILKIE, A.D. and SAHIN, S. (2018) Yet more on a stochastic economic model: part 4: a model for share earnings, dividends, and prices. *Annals of Actuarial Science,* **12**, 67–105.
ZHANG, L., MYKLAND, P.A. and AÏT-SAHALIA, Y. (2005) A tale of two time scales. *Journal of the American Statistical Association,* **100**, 1394–1411.
Zhang, S., Hardy, M. and Saunders, D. (2018) Updating Wilkie's economic scenario generator for U.S. applications. *North American Actuarial Journal*, **22**, 600–622.

JEAN-FRANÇOIS BÉGIN
Department of Statistics and Actuarial Science
Simon Fraser University
8888 University Drive, Burnaby
British Columbia, V5A 1S6, Canada
E-mail: jbegin@sfu.ca

APPENDIX A. PRIOR DISTRIBUTIONS

A.1. Non-informative prior

When there is insufficient information about the value of a parameter or if the end-user does not have strong prior beliefs, then a non-informative prior can be used. They are also called vague or diffuse prior in the literature. Typically, in their simplest forms, these priors are rectangular distributions (i.e., flat) over the feasible set of parameter values (Upton and Cook, 2014).

A flat prior is considered improper (i.e., does not exist) if the interval is infinite, which would be the case for the model parameters that are unbounded. An improper prior is not necessarily a bad choice as long as the posterior distribution is proper. If it leads to an improper posterior, however, then no Bayesian inference can be achieved.

To verify that a posterior distribution is, in fact, proper, we would need to derive its density in closed form which is not always possible. Therefore, on the one hand, we use a flat (improper) prior for cases where we can obtain the closed-form solution. On the other hand, however, we use a very diffuse prior if we cannot find the density in closed form, that is, a distribution with a very large dispersion, albeit proper.

Specifically, for the inflation dynamics, both parameters μ_i and i_0 have flat priors, that is, $\mu_i \sim \mathcal{U}(-\tau, \tau)$ and $i_0 \sim \mathcal{U}(-\tau, \tau)$ with $\tau \to \infty$. The autoregressive parameter a_i is constrained between -1 and 1, so a uniform distribution over this interval is used as the prior distribution for this parameter: $a_i \sim \mathcal{U}(-1, 1)$. We also use a flat prior for $\log(\sigma_i)$; in other words, $\pi(\sigma_i) \propto 1/\sigma_i$, so that the $\log(\sigma_i)$ is uniformly distributed on the real line. It is possible to show that all these posterior distributions are in fact proper (see Section 5 and Appendix B for more details).

The rationale to select non-informative priors for long-term interest rate parameters remains almost the same. The mean parameter μ_c^* has a flat prior. The autoregressive parameter a_c has a uniform prior over $(-1, 1)$. Finally, the $\log(\sigma_c)$ are uniformly distributed on the real line. The posterior distributions of parameters $c_{m,0}$ and c_0 cannot be found in closed form if we use a flat prior for these parameters; therefore, for these two parameters, we use a very diffuse prior, that is, $\mathcal{N}(0, 10^2)$.

The same very diffuse prior is used for the majority of dividend yield and dividend index parameters: μ_y^*, w_y, $y_{m,0}$, $z_{y,0}$, μ_d, y_d, $d_{m,0}$ and $z_{d,0}$ are normally distributed (with a mean of zero and a standard deviation of 10). The only exceptions are the scale, the autoregressive and the moving-average parameters. The scale priors are based on exponential distributions for which the rate parameter is very large, that is, 10. The autoregressive and the moving-average parameters are flat over the interval $(-1, 1)$, which is proper.

The leftmost column of Table 5 summarizes the different non-informative prior distributions used in this study.

A.2. Subjective prior

A subjective prior expresses definite and specific information about the parameters. In this study, the subjective prior we use is inspired by the parameter estimates obtained by Wilkie (1995) for the US economy. The prior is centred around the parameters estimated in Wilkie's 1995 update. For the *a priori* dispersion, we use the following rules of thumb: initial values, mean, scale, autoregressive and moving-average parameters have a standard deviation of 0.1, and other parameters have a dispersion of 0.5. The rightmost column of Table 5 provides details on the various subjective prior distributions we use.

Generally speaking, we assume that all parameters are *a priori* normally distributed, except for the scale, the autoregressive and the moving-average parameters. The scale parameters are given by moment-matched gamma distributions. The autoregressive and the moving-average parameters, a_i, a_y, b_d and a_c, are given by moment-matched (general) beta distributions defined over the interval $(-1, 1)$, in the spirit of Gill and Freeman (2013). The beta distribution is used because of its flexibility (O'Hagan, 1998). The probability density function of this general beta distribution we use is given by

$$f(y|\alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{(y+1)^{\alpha-1}(1-y)^{\beta-1}}{2^{\alpha+\beta-1}},$$

where α and β are shape parameters.

TABLE 5

SUMMARY OF PRIOR DISTRIBUTIONS.

Non-informative			Subjective		
Inflation			**Inflation**		
μ_i	Uniform	$\mathcal{U}(-\tau,\tau)$, $\tau \to \infty$	μ_i	Normal	$\mathcal{N}(0.03, 0.1^2)$
$\log(\sigma_i)$	Uniform	$\mathcal{U}(-\tau,\tau)$, $\tau \to \infty$	σ_i	Gamma	$\mathcal{G}amma(0.123, 0.286)$
a_i	Uniform	$\mathcal{U}(-1, 1)$	a_i	Beta	$\mathcal{B}eta(46.819, 9.931)$
i_0	Uniform	$\mathcal{U}(-\tau,\tau)$, $\tau \to \infty$	i_0	Normal	$\mathcal{N}(0.03, 0.1^2)$
Dividend Yield			**Dividend Yield**		
μ_y^*	Normal	$\mathcal{N}(0, 10^2)$	μ_y^*	Normal	$\mathcal{N}(\log(0.043), 1.028^2)$
σ_y	Exponential	Exp(10)	σ_y	Gamma	$\mathcal{G}amma(4.41, 0.048)$
a_y	Uniform	$\mathcal{U}(-1, 1)$	a_y	Beta	$\mathcal{B}eta(42.5, 7.5)$
w_y	Normal	$\mathcal{N}(0, 10^2)$	w_y	Normal	$\mathcal{N}(0.5, 0.5^2)$
$y_{m,0}$	Normal	$\mathcal{N}(0, 10^2)$	$y_{m,0}$	Normal	$\mathcal{N}(0, 0.1^2)$
$z_{y,0}$	Normal	$\mathcal{N}(0, 10^2)$	$z_{y,0}$	Normal	$\mathcal{N}(0.021, 0.1^2)$
Dividend Index			**Dividend Index**		
μ_d	Normal	$\mathcal{N}(0, 10^2)$	μ_d	Normal	$\mathcal{N}(0.0155, 0.1^2)$
σ_d	Exponential	Exp(10)	σ_d	Gamma	$\mathcal{G}amma(0.81, 0.111)$
b_d	Uniform	$\mathcal{U}(-1, 1)$	b_d	Beta	$\mathcal{B}eta(55.5, 18.5)$
y_d	Normal	$\mathcal{N}(0, 10^2)$	y_d	Normal	$\mathcal{N}(-0.35, 0.5^2)$
$d_{m,0}$	Normal	$\mathcal{N}(0, 10^2)$	$d_{m,0}$	Normal	$\mathcal{N}(0.03, 0.1^2)$
$z_{d,0}$	Normal	$\mathcal{N}(0, 10^2)$	$z_{d,0}$	Normal	$\mathcal{N}(0, 0.1^2)$
Long-Term Interest Rate			**Long-Term Interest Rate**		
μ_c^*	Uniform	$\mathcal{U}(-\tau,\tau)$, $\tau \to \infty$	μ_c^*	Normal	$\mathcal{N}(\log(0.0265), 1.208^2)$
$\log(\sigma_c)$	Uniform	$\mathcal{U}(-\tau,\tau)$, $\tau \to \infty$	σ_c	Gamma	$\mathcal{G}amma(4.41, 0.048)$
a_c	Uniform	$\mathcal{U}(-1, 1)$	a_c	Beta	$\mathcal{B}eta(6.703, 0.137)$
y_c	Uniform	$\mathcal{U}(-\tau,\tau)$, $\tau \to \infty$	y_c	Normal	$\mathcal{N}(0.07, 0.5^2)$
$c_{m,0}$	Normal	$\mathcal{N}(0, 10^2)$	$c_{m,0}$	Normal	$\mathcal{N}(0.0265, 0.1^2)$
c_0	Normal	$\mathcal{N}(0, 10^2)$	c_0	Normal	$\mathcal{N}(0.0265, 0.1^2)$

This table summarizes the two sets of priors: non-informative (leftmost column) and subjective (rightmost column).

Finally, when we use the transformed version of the mean parameters (i.e., μ_y^* and μ_c^*), we moment-matched the standard deviation of the normal prior such that the actual untransformed parameter has a standard deviation of 0.1.

APPENDIX B. MORE ON THE ESTIMATION METHODOLOGY

The main goal of this appendix is to present the specific methodology used to obtain a sequence of Monte Carlo samples, $\{\Theta^{(j)}\}_{h=1}^{M}$, where $\Theta^{(j)}$ is the jth simulated parameter set.

We apply a Gibbs sampler to obtain samples of the model's parameters. Gibbs sampling is a technique that generates a sequence of random variables—in our case, the model

https://doi.org/10.1017/asb.2019.6 Published online by Cambridge University Press

parameters—by using multiple (small-dimensional) sampling instead of sampling directly from a multidimensional probability distribution. Specifically, parameters are sampled one at a time from their full conditional distributions—the distribution of a given parameter, θ_i, conditional on X and all the other parameters. Therefore, instead of sampling directly from the multivariate distribution we wish to approximate—which involves multiple integrals—we deal with multiple one-dimensional distributions. Beginning from an initial parameter set $\Theta^{(0)}$, we can generate the subsequent set, $\Theta^{(1)}$, and do the same thing for the following sets in an iterative way. In other words, we generate the ith parameter at step j from the following distribution:

$$\theta_i^{(j)} \sim \pi\left(\cdot \middle| X, \Theta_{1:i-1}^{(j)}, \Theta_{i+1:n}^{(j-1)}\right),$$

where $\Theta_{a:b}^{(j)}$ corresponds to model parameters, from the ath up to the bth, at the jth step. We therefore use information from both the $j-1$th and jth sets to obtain $\theta_i^{(j)}$. As the Gibbs sampler iterates, it produces a good approximation of the marginal posterior distribution for the said parameter, $\pi(\theta_i|X)$, and for every parameter in Θ.

Fortunately, we can find the *a posteriori* full conditional distribution for some of the model parameters when the non-informative prior is used—which greatly simplifies our implementation of the Gibbs sampler. However, in the subjective case and for some parameters in the non-informative case, it is difficult to sample directly from their full conditional distribution. Therefore, for these, we use the Metropolis–Hastings algorithm in lieu of direct steps in the Gibbs sampler. For each iteration j, the Metropolis–Hastings method allows us to obtain the sampled parameters by generating a candidate from some proposal distributions that are as yet undefined. Then, the candidate is either accepted or rejected. Two main ingredients are thus needed to apply Metropolis-within-Gibbs steps: (1) a proposal density and (2) the (marginal) posterior density from which we wish to sample. The proposal density g can be almost anything as long as it is easy to generate a sample from it. As for the *a posteriori* full conditional density from which we want to sample, we can compute it readily from Equation (6) up to a normalizing constant.

Using these two building blocks, we can now apply Metropolis-within-Gibbs steps, as described in the following algorithm.

Algorithm 1. Metropolis–Hastings Algorithm

At step j and for $\theta_i \in \Theta$, we do the following steps.

1. *Generate a candidate state θ_i^* from $g(\,\cdot \mid \theta_i^{(j-1)})$. This proposal density should be based on the previous value of θ_i in the Markov chain, that is, $\theta_i^{(j-1)}$.*
2. *Calculate the acceptance probability*

$$\alpha\left(\theta_i^*, \theta_i^{(j-1)}\right) = \min\left(1, \frac{\pi\left(\theta_i^* \middle| X, \Theta_{1:i-1}^{(j)}, \Theta_{i+1:n}^{(j-1)}\right)}{\pi\left(\theta_i^{(j-1)} \middle| X, \Theta_{1:i-1}^{(j)}, \Theta_{i+1:n}^{(j-1)}\right)} \frac{g(\theta_i^{(j-1)}|\theta_i^*)}{g(\theta_i^* \mid \theta_i^{(j-1)})}\right). \tag{8}$$

3. *Accept or reject this new sample using a uniform random number $u \sim \mathcal{U}(0,1)$, that is, if $u \leq \alpha\left(\theta_i^*, \theta_i^{(j-1)}\right)$, accept the new state and set $\theta_i^{(j)} = \theta_i^*$, or if $u > \alpha\left(\theta_i^*, \theta_i^{(j-1)}\right)$, reject the new sate and copy the old state forward.*

Because of its specific structure, our MCMC sampler can be divided into four smaller blocks—one for each sub-model of the ESG. These four blocks are described below.

B.1. Inflation

For the subjective case, we will use the Metropolis-within-Gibbs methods to sample μ_i, σ_i^2, a_i and i_0, as explained above. For the non-informative case, we can find the full conditional posterior distribution of certain parameters.

Proposition 1. *Full Conditional Posterior Distribution for Inflation Parameters (Non-informative Case)*
The full conditional a posteriori distributions of the inflation parameters are given by

$$\mu_i \big| a_i, \sigma_i, \{i(t)\}_{t=0}^T \sim \mathcal{N}\left(\frac{\sum_{t=1}^T i(t) - a_i i(t-1)}{T(1-a_i)}, \frac{\sigma_i^2}{T(1-a_i)^2}\right),$$

$$\sigma_i^2 \big| \mu_i, a_i, \{i(t)\}_{t=0}^T \sim \mathrm{IG}\left(\frac{T}{2}+1, \frac{1}{2}\sum_{t=1}^T \left(i(t) - \mu_i - a_i\left(i(t-1) - \mu_i\right)\right)^2\right),$$

$$a_i \big| \mu_i, \sigma_i, \{i(t)\}_{t=0}^T \sim \mathcal{N}_{(-1,1)}\left(\frac{\sum_{t=1}^T \left(i(t-1) - \mu_i\right)\left(i(t) - \mu_i\right)}{\sum_{t=1}^T \left(i(t-1) - \mu_i\right)^2}, \frac{\sigma_i^2}{\sum_{t=1}^T \left(i(t-1) - \mu_i\right)^2}\right),$$

$$i_0 \big| \mu_i, \sigma_i, a_i, \{i(t)\}_{t=1}^T \sim \mathcal{N}\left(\frac{i(1) - \mu_i(1-a_i)}{a_i}, \frac{\sigma_i^2}{a_i^2}\right),$$

where $\mathcal{N}_{(-1,1)}(\mu, \sigma^2)$ *is a Gaussian distribution truncated to* $(-1, 1)$ *with a mean parameter of* μ *and a variance parameter of* σ^2, *and* $\mathrm{IG}(\alpha, \beta)$ *is an inverse Gamma distribution with a shape parameter of* α *and a scale parameter of* β.

Proof. See Section SM.A.1 of the Supplementary Material. ■

B.2. Dividend yield

For both the non-informative and the subjective prior distributions, we need to use the Metropolis-within-Gibbs method to obtain *a posteriori* samples of the dividend yield parameters. Indeed, because the dividend yield innovations $z_y(t)$ are also involved in the dividend index and the long-term interest rate dynamics, the full conditional posterior distribution of Θ_y is cumbersome and prevents us from directly using the Gibbs sampler.

B.3. Dividend index

Finding full conditional posterior distributions for the parameters involved in the dividend index dynamics is also troublesome since the time t density involved in Equation (4) depends on $z_d(t-1)$, which in turn depends on the other dividend index parameters. Therefore, the parameters μ_d, σ_d^2, b_d, y_d, $d_{m,0}$ and $z_{d,0}$ are all sampled via Metropolis-within-Gibbs steps.

B.4. Long-term interest rate

We use the Metropolis-within-Gibbs method to sample the long-term interest rate parameter Θ_c for the subjective case. For the non-informative prior, a few full conditional distributions can be obtained analytically.

Proposition 2. *Full Conditional Posterior Distribution for Long-Term Interest Rate Parameters (Non-informative Case)*
The full conditional posterior distributions of the long-term interest rate parameters are given by

$$\mu_c^* \big| \ \sigma_c^2, a_c, y_c, c_{m,0}, c_0, \{i(t), y(t), c(t)\}_{t=0}^T$$
$$\sim \mathcal{N}\left(\frac{\sum_{t=1}^T \left(\log\left(c(t)-c_m(t)\right) - a_c \log\left(c(t-1)-c_m(t-1)\right) - y_c z_y(t)\right)}{T(1-a_c)}, \frac{\sigma_d^2}{T(1-a_c)^2}\right),$$

$$\sigma_c^2 \big| \ \mu_c^*, a_c, y_c, c_{m,0}, c_0, \{i(t), y(t), c(t)\}_{t=0}^T$$
$$\sim \mathrm{IG}\left(\frac{T}{2}+1, \frac{1}{2}\sum_{t=1}^T \left(\log\left(c(t)-c_m(t)\right) - \mu_c^* - a_c \left(\log\left(c(t-1)-c_m(t-1)\right) - \mu_c^*\right) - y_c z_y(t)\right)^2\right),$$

$$a_c \big| \ \mu_c^*, \sigma_c^2, y_c, c_{m,0}, c_0, \{i(t), y(t), c(t)\}_{t=0}^T$$
$$\sim \mathcal{N}_{(-1,1)}\left(\frac{\sum_{t=1}^T \left(\log\left(c(t-1)-c_m(t-1)\right) - \mu_c^*\right)\left(\log\left(c(t)-c_m(t)\right) - \mu_c^* - y_c z_y(t)\right)}{\sum_{t=1}^T \left(\log\left(c(t-1)-c_m(t-1)\right) - \mu_c^*\right)^2}, \frac{\sigma_c^2}{\sum_{t=1}^T \left(\log\left(c(t-1)-c_m(t-1)\right) - \mu_c^*\right)^2}\right),$$

$$y_c \big| \ \mu_c^*, \sigma_c^2, a_c, c_{m,0}, c_0, \{i(t), y(t), c(t)\}_{t=0}^T$$
$$\sim \mathcal{N}\left(\frac{\sum_{t=1}^T z_y(t)\left(\log\left(c(t)-c_m(t)\right) - \mu_c^* - a_c\left(\log\left(c(t-1)-c_m(t-1)\right) - \mu_c^*\right)\right)}{\sum_{t=1}^T z_y^2(t)}, \frac{\sigma_c^2}{\sum_{t=1}^T z_y^2(t)}\right).$$

The extra parameters $c_{m,0}$ and c_0 are generated using Metropolis–Hastings steps, as explained above.

Proof. See Section SM.A.2 of the Supplementary Material. ■

B.5. Implementation and convergence issues

In the Metropolis-within-Gibbs steps, a proposal density is needed. In this study, we use the Gaussian distribution or, when the parameter is constrained, a truncated Gaussian distribution. The location parameter of these normal distributions is set to the current value of the parameter in the Markov chain, and their standard deviation is chosen as half of the standard error on the parameter estimate obtained via MLE (see Table 1).

Because the truncated Gaussian distribution is no longer symmetric, the Metropolis–Hastings acceptance ratios need to be changed. The acceptance probability is still given by Equation (8), although the last part of the latter is no longer equal to 1; for nonsymmetric proposals, the probability density function of the normal distribution needs to be scaled to account for the fact that some parameters must be positive. This scaling factor is a function of the normal cdf.

https://doi.org/10.1017/asb.2019.6 Published online by Cambridge University Press

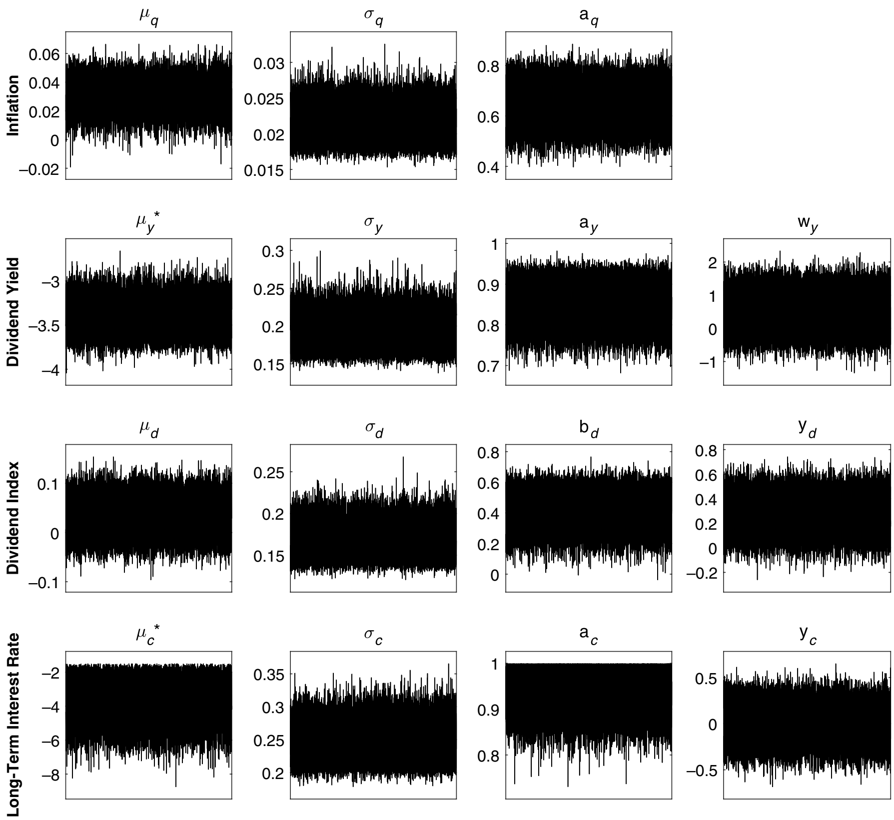

FIGURE 9: Trace plots for model parameters: subjective prior.
This figure exhibits the trace plot of the draws from the posterior distribution for each parameter obtained using the MCMC methodology explained in Section 5. For each parameter, we obtain a sample of 100,000 values. In this study, we focus on post-World War II data, that is, data for the year 1945 and after.

The proposal distribution used in this study is reasonable: the acceptance rate ranges between 20% and 73%. On average, it is about 49% (42%) across the various parameters for the non-informative (subjective) prior.

To address potential slow convergence issues, we use a long Markov chain, that is, $M = 510,000$ observations. The first 10,000 observations are considered the burn-in period and are therefore removed from the sample. Thereafter, every fifth simulation is recorded for posterior analysis to cope with potential highly dependent samples (i.e., thinning). This process yields a final Markov chain of size 100,000 that can be used for empirical purposes.

In most cases, the Gibbs sampler is convergent for problems with a continuous target density and a non-disjoint support. However, it is very difficult to verify this convergence from a theoretical standpoint in high-dimensional problems. As for many practical applications of MCMC, the convergence can therefore be assessed on the basis of the empirical analysis of the output of sampled parameters, for example, using different starting values and providing output for which the empirical distributions are indistinguishable. As a robustness test, we use different starting values and find the same posterior distributions.

Using these different chains, we can evaluate empirically the convergence of our sampler. The Gelman and Rubin (1992) diagnostic is a neat way to do that. This diagnostic test, based on the potential scale reduction factor (PSRF), assesses the convergence by comparing the estimated between-chains and within-chain variances for each model parameter. Simply put, if the chains under scrutiny have converged to the target posterior distribution, then the PSRF should be close to 1. Indeed, for all the parameters and for the two priors used in this study, all the PSRFs range between 1 and 1.0001, meaning that the chains seem to have converged to their target posterior distribution.

Finally, to assess the mixing of our chains, we present trace plots. Figure 9 shows the trace plots for all the model parameters under scrutiny. Generally speaking, we want to try to avoid flat bits or too many consecutive steps in one direction. Figure 9 shows the opposite: the trace plots exhibit a *hairy caterpillar* behaviour, meaning that the mixing seems adequate. The trace plots for the non-informative case are similar, qualitatively speaking; they are available in Section SM.B of the Supplementary Material.