

A Bayesian Beta Markov Random Field Calibration of the Term Structure of Implied Risk Neutral Densities

Roberto Casarin^{*}, Fabrizio Leisen[†], German Molina[‡], and Enrique ter Horst[§]

Abstract. We build on the derivative pricing calibration literature, and propose a more general calibration model for implied risk neutral densities. Our model allows for the joint calibration of a set of densities at different maturities and dates through a Bayesian dynamic Beta Markov Random Field. Our approach allows for possible time dependence between densities with the same maturity, and for dependence across maturities at the same point in time. This approach to the risk neutral density calibration problem encompasses model flexibility, parameter parsimony, and, more importantly, information pooling across densities. This proposed methodology can be naturally extended to other areas where multidimensional calibration is needed.

Keywords: Bayesian inference, Beta Markov Random Fields, exchange Metropolis Hastings, risk neutral measure, density calibration, distortion function.

1 Introduction

Linear and nonlinear derivatives are used by investors to hedge risks, as well as speculate on the perceived market uncertainty of underlying assets (Hull, 2014). Furthermore, derivatives convey information about the likelihood that the market assigns to different future outcomes. Therefore, extracting the likelihood or probabilities regarding the different future outcomes through different statistical techniques yields the implied probability density function (pdf), which is used to price the given quoted market derivative prices for every given period (Fackler and King, 1990; Lai, 2011; Ait-Sahalia and Lo, 1998a; Panigirtzoglou and Skiadopoulos, 2004; Bliss and Panigirtzoglou, 2002). These option-implied pdfs give us a forward-looking view of what the market thinks will happen to any given asset together with its respective probabilities (Vergote and Puigvert, 2012; Vesela and Puigvert, 2014).

The economic literature has shown an increasing interest in nonparametric implied risk neutral densities (Fackler and King, 1990; Lai, 2011), both allowing estimation of what the economic agents think about the future, their economic expectations (Bliss and Panigirtzoglou, 2004; Rodriguez and ter Horst, 2008), and also providing superior estimates of such risk-neutral densities (Lai, 2011). Density forecasting is now widely

^{*}University Ca' Foscari of Venice, Italy, r.casarin@unive.it

[†]University of Kent, U.K., fabrizio.leisen@gmail.com

[‡]Idalion Capital Group, U.K. and U.S., german@germanmolina.com

[§]Colegio de Estudios Superiores de Administración-CESA, Colombia & IESA, Venezuela, enriquetterhorst@gmail.com

used in many applied economic contexts, such as nonparametric calibration of implied risk neutral densities used in macroeconomics to generate predictions on inflation and interest rates, among other applications (see a non-comprehensive list in Bhar and Chiarella (2000); Carlson et al. (2005); Vincent-Humphreys and Noss (2012); Vergote and Puigvert (2012); Vesela and Puigvert (2014); Sihvonen and Vähämaa (2014)). The use of densities for predicting quantities of interest is now common in economics and finance (Bhar and Chiarella, 2000; Carlson et al., 2005; Vergote and Puigvert, 2012; Vesela and Puigvert, 2014). In the spirit of West and Crosse (1992) and West (1992) with regards to the inclusion in the forecast of a decision maker's prior from different sources of information, many recent papers (Hall and Mitchell, 2007; Billio et al., 2013; Fawcett et al., 2013; Gneiting and Ranjan, 2013) have focused on the combination and calibration of predictive densities. See also Chapter 16 of West and Harrison (1997) for further initial references.

In order to extract the implied risk neutral densities, our starting point is in line with the nonparametric cubic spline interpolation technique by Bliss and Panigirtzoglou (2002). These authors make use of an important result from Breeden and Litzenberger (1978) where the option-implied risk-neutral density can be inferred from the second partial derivative of the call price function with respect to the strike price. However, the estimation of the second derivative of a call price function can sometimes lead to numerically unstable and inaccurate estimates of the option-implied density. Following the results derived in Malz (1997) and Shimko (1993), Bliss and Panigirtzoglou (2002) argue that numerically stable and accurate results for the implied pdfs are obtained if option prices and strike prices are transformed into implied volatility and delta values prior to their penalized cubic spline interpolation. Full implementation details of the nonparametric estimation of the option-implied density estimation can be found in Bliss and Panigirtzoglou (2002) and Campa et al. (1998). This nonparametric density estimation consists of four steps: transforming the option strike prices into implied volatilities (when those are not directly available); using the market-implied volatilities to calculate the delta values; interpolating the implied volatilities and the deltas calculated in the previous two steps by minimizing the function (26) as defined in Appendix A; and transforming back the interpolated volatility smile from volatility vs delta values to option price vs strike price values. In order to obtain the option-implied, risk-neutral density, we calculate the second derivative of the call function that we computed in the previous four steps (Breeden and Litzenberger, 1978).

As argued in Diebold et al. (1999), many density forecasts of interest do not come from known parametric models such as surveys (Diebold et al., 1999) or from prices of options written at different strike prices (Ait-Sahalia and Lo, 1998b; Soderlind and Svensson, 1997; Campa et al., 1998). Also, modelling the dependence between the risk-neutral and physical measures has been explored by the use of affine and non-affine stochastic processes (Filipovic, 2009; Durham, 2013) as well as the use of copulas (Schonbucher, 2003; Jaworski et al., 2010). While much of the prior work in this area has focused on affine and affine-jump models because of their analytical tractability, Durham (2013) has found that log-volatility models perform dramatically better than affine models. The approach taken in this manuscript is different from the already existing parametric literature on modelling the dependence between risk-neutral and physical densities (Schonbucher, 2003; Filipovic, 2009; Jaworski et al., 2010) in two ways:

first, we model the dependence between the risk-neutral and physical densities by using the beta distortion function (Artzner et al., 1999; Wang et al., 1997; Wang and Young, 1998; Gzyl and Mayoral, 2008), and second, both our risk-neutral and physical densities are nonparametric (Bliss and Panigirtzoglou, 2002), whereas the approach in Durham (2013) and Filipovic (2009) is a parametric one. Our approach on estimating the dependence between the nonparametric risk-neutral (\mathbb{Q}) and physical (\mathbb{P}) densities¹ can be seen as a complement to the already existing parametric literature (Duffie et al., 2000; Schonbucher, 2003; Filipovic, 2009; Jaworski et al., 2010; Durham, 2013). Furthermore, following previous work such as Diebold and Mariano (1995), Christoffersen (1998) and Diebold et al. (1998), we view the forecasts as primitives – our proposed approach will not depend on the method used to produce the risk-neutral density forecasts.

The basic intuition behind our dynamic estimation of the dependence between \mathbb{Q} and \mathbb{P} hinges on the concept of the *Probability Integral Transforms* (PIT) (Fackler and King, 1990; Diebold et al., 1999). Simply put, if we observe n random realizations of a random variable Y and its true density is $f(y)$, then the $\text{PIT}_i = \int_{-\infty}^{y_i} f(y)dy$ for $i = 1, \dots, n$ should be uniformly distributed. If the PIT_i are not uniformly distributed, then by using a beta distortion function (Artzner et al., 1999; Wang et al., 1997; Wang and Young, 1998) to $f(y)$, by modifying their probability, the PIT_i can be distorted to be uniformly distributed (Fackler and King, 1990; Diebold et al., 1999). Gzyl and Mayoral (2008) provide a method for reconstructing distortion functions from the observed prices of risk premiums. Our contribution provides a natural modelling framework for the term structure of the implied nonparametric risk neutral and physical probability distributions, which accounts for the possible dependence between the PITs at different maturities and different dates for a given maturity, whereas previous calibration approaches lack this generality, as they do not take advantage of the different sources of dependence (for example, maturities and/or time). Additionally, our approach potentially allows for modelling other sources of dependence (e.g. cross-asset dependence).

Since the PITs belong to the unit interval, our calibration approach makes use of beta densities as initially suggested by Fackler and King (1990) and Diebold et al. (1999). However, we extend their approach in our paper to the multi-maturity and multi-period setting. In order to account for time and cross-maturity dependence, we propose a random field model with beta densities, which fits well-known features of the data regarding dependencies. We make some general assumptions on the time (lags) and spatial (neighbour system) structure that are needed to obtain a parsimonious model. We provide a Bayesian inference framework, that allows us to account for uncertainty in the density calibration. Moreover, the use of hierarchical prior distributions allows us not only to avoid potential over-fitting due to over-parameterization, but also to achieve different degrees of information pooling across maturities.

Several recent papers focus on the combination and the calibration of the predictive densities (Hall and Mitchell, 2007; Geweke and Amisano, 2011; Billio et al., 2013; Fawcett et al., 2013; Gneiting and Ranjan, 2013). Optimal linear pooling of densities is considered in Hall and Mitchell (2007), as well as Geweke and Amisano (2011), while

¹The notations \mathbb{Q} and \mathbb{P} will refer to the nonparametric risk-neutral and physical densities, respectively, in our manuscript.

more general approaches to density combinations are considered in Billio et al. (2013), Fawcett et al. (2013) and Gneiting and Ranjan (2013). Modelling the time evolution of the optimal combination of predictive densities is one of the challenging issues tackled in these papers. The issue of calibrating densities is considered instead in Gneiting et al. (2005) and Gneiting and Ranjan (2013), who also propose the use of beta densities to achieve a continuous deformation of the predictive density, and to obtain well-calibrated PITs. A well-calibrated PIT is defined as one where the calibration function allows us to obtain a cumulative probability distribution of the observed underlying asset, under the correct target distribution (physical measure), with a resulting uniform histogram (Fackler and King, 1990). Despite the presence in the forecasting and financial literature of similar issues, such as the density calibration and combination, the implied risk neutral calibration literature differs substantially from the forecast calibration literature since the former assumes the calibration model is generating the change of measure needed to obtain the physical measure from the risk neutral. Our paper contributes to this stream of literature through a much-needed extension for capturing key features, since it provides a general approach to the joint calibration of densities, allowing for the pooling of information across different predictive densities (e.g. the risk neutral densities at different maturities).

Finally, as a side note, this paper also contributes more generally to the literature on modelling data on bounded domains (Grunwald et al., 1993; McKenzie, 1985; Wallis, 1987; Ferrari and Cribari-Neto, 2004). Our approach using Bayesian Beta Markov Random Field models, as well as the inferential procedure, are original extensions to the multivariate context of the Bayesian Beta models and inferences recently proposed in the statistical literature. See Branscum et al. (2007) for Bayesian Beta regression, Casarin et al. (2012) for model selection in Bayesian Beta autoregressive models and the references therein, and Kennedy (1994), Goldstein (2000), and Kimmel (2004) for modelling the term structure of interest rates by using Gaussian Markov Random Fields. For further background on Markov Random Fields, see Cressie and Lele (1992). While we build our proposed approach through a financial application, the model and approach proposed here are broader and can be used in other areas where more generic, multidimensional calibration approaches are needed.

The paper is organized as follows: Section 2 introduces the density calibration problem and our Bayesian Beta Markov Random Field model for the joint calibration. In Section 3, we discuss the inference difficulties with the proposed model and develop a numerical procedure for posterior computation. In Section 4, we study the efficiency of our estimation procedure through simulation. In Section 5, we provide an application in the Foreign Exchange market, while Section 6 concludes and discusses potential extensions.

2 A dynamic calibration model

2.1 The univariate motivation

The major limitation of using risk-neutral densities is that they do not incorporate risk premium and thus do not correspond to investor's actual forecasts. The principal

reason is that the risk-neutral density extraction relies on the simplifying assumption of an arbitrage free market and the definition of an equivalent martingale measure, which in turn implies that all assets have the same expected return which is equal to the risk-free rate (Delbaen and Schachermayer, 2011; Hull, 2014; Duffie, 2001; Skiadas, 2009). One therefore needs to correct this bias by estimating the real-world densities (RWD) using a statistical calibration such as proposed by Fackler and King (1990) and Diebold et al. (1999). Furthermore, RWD densities put lesser probability weight on bad states than risk-neutral densities do. As suggested by Fackler and King (1990) and Diebold et al. (1999), one can convert the risk-neutral densities to their real-world counterparts by adequately risk adjusting the risk neutral density. Let $f_{t,\tau}^Q(x)$ be the density forecast made at time t for a given random variable $x_{t,\tau}$ that we observe at time $t + \tau$ during $t = 1, \dots, T$ periods. The probability integral transform (PIT) of $x_{t,\tau}$ with respect to $f_{t,\tau}^Q(x)$ is defined as

$$y_{t,\tau} = \int_{-\infty}^{x_{t,\tau}} f_{t,\tau}^Q(x) dx = F_{t,\tau}^Q(x_{t,\tau}). \tag{1}$$

We can therefore see that computing the PIT of the random variable $x_{t,\tau}$ is equivalent to computing the cumulative density function (CDF) defined as $F_{t,\tau}^Q(x)$ of the density used to forecast our random variable $x_{t,\tau}$. Let us write $C(y_{t,\tau})$ as the CDF of the PIT $y_{t,\tau}$. We then have the following relationship:

$$\begin{aligned} C(y_{t,\tau}) &= \mathbb{P}(Y_{t,\tau} \leq y_{t,\tau}) = \mathbb{P}(F_{t,\tau}^Q(x_{t,\tau}) \leq y_{t,\tau}) \\ &= \mathbb{P}(x_{t,\tau} \leq F_{t,\tau}^{(-1)Q}(y_{t,\tau})) = F_{t,\tau}^P(F_{t,\tau}^{(-1)Q}(y_{t,\tau})) \end{aligned} \tag{2}$$

where $F_{t,\tau}^P$ is the CDF of the data generating process for $x_{t,\tau}$. We can therefore see that if the forecasting risk-neutral CDF $F_{t,\tau}^Q$ is correctly calibrated, i.e. $F_{t,\tau}^Q = F_{t,\tau}^P$ then we have that the PITs $y_{t,\tau}$ will be independent and uniformly distributed and $C(y_{t,\tau}) = y_{t,\tau}$. Remembering that \mathbb{P} and \mathbb{Q} denote the physical (Real World) and risk-neutral CDFs, respectively, we have that

$$\begin{aligned} F_{t,\tau}^P(x_{t,\tau}) &= \mathbb{P}(X_{t,\tau} \leq x_{t,\tau}) = \mathbb{P}(F_{t,\tau}^Q(X_{t,\tau}) \leq F_{t,\tau}^Q(x_{t,\tau})) \\ &= \mathbb{P}(Y_{t,\tau} \leq F_{t,\tau}^Q(x_{t,\tau})) = C(F_{t,\tau}^Q(x_{t,\tau})). \end{aligned} \tag{3}$$

Now taking derivatives on both sides of (3), we can see how the physical density relates to the risk-neutral density:

$$f_{t,\tau}^P(x_{t,\tau}) = \frac{\partial C}{\partial F_{t,\tau}^Q(x_{t,\tau})} \frac{\partial F_{t,\tau}^Q(x_{t,\tau})}{\partial x_{t,\tau}} = \frac{\partial C}{\partial F_{t,\tau}^Q(x_{t,\tau})} f_{t,\tau}^Q(x_{t,\tau}) = c(F_{t,\tau}^Q(x_{t,\tau})) f_{t,\tau}^Q(x_{t,\tau}) \tag{4}$$

where $c(F_{t,\tau}^Q(x_{t,\tau})) = \frac{\partial C}{\partial F_{t,\tau}^Q(x_{t,\tau})} = \frac{(F_{t,\tau}^Q(x_{t,\tau}))^{p-1} (1 - (F_{t,\tau}^Q(x_{t,\tau})))^{q-1}}{B(p,q)}$, and $B(p, q)$ is the usual normalization constant of a Beta(p, q) distribution (Fackler and King, 1990). In order to know the distortion function $c(\cdot)$, one can determine both p and q with a regular maximum likelihood estimation (Fackler and King, 1990; Diebold et al., 1999). To test the forecasting ability of the estimated risk-neutral density $f_{t,\tau}^Q(x_{t,\tau})$, we can test the null

hypothesis that $f_{t,\tau}^Q(x_{t,\tau}) = f_{t,\tau}^P(x_{t,\tau})$ under which the PITs $y_{t,\tau}$ are independent and uniformly distributed $U[0, 1]$. In this paper, we follow Bliss and Panigirtzoglou (2004) and apply the Berkowitz test (Berkowitz, 2001) for PITs uniformity and independence (see Section 5). As pointed out in Vergote and Puigvert (2012) and Vesela and Puigvert (2014), using traditional approaches, in order to avoid serial dependence in the PITs that usually arise from overlapping maturities in the option data, it is better to work with monthly maturities or less than longer maturity options due then to a lack of sufficient data. Contrary to their approach, ours will not suffer from this restriction, as detailed in the following section.

2.2 The multivariate case

Let x_{t,τ_i} , $i = 1, \dots, M$, and $t = 1, \dots, T$, be a set of random variables that in our financial application can be treated as the underlying realized forward levels (market-implied estimates of the asset level at maturity), available at time t , for the different future maturities τ_1, \dots, τ_M . Let $F_{t,\tau_i}^Q(x)$ and $F_{t,\tau_i}^P(x)$ denote the risk neutral and the physical cumulative density functions, respectively, and $f_{t,\tau_i}^Q(x)$ and $f_{t,\tau_i}^P(x)$ their probability density functions.

We assume the following joint deformation model:

$$F_t^P(x_{t,\tau_1}, \dots, x_{t,\tau_M}) = C_t(F_{t,\tau_1}^Q(x_{t,\tau_1}), \dots, F_{t,\tau_M}^Q(x_{t,\tau_M})) \quad (5)$$

where $C_t : [0, 1]^M \rightarrow [0, 1]$, $t = 1, \dots, T$, is a sequence of deformation functions. The model can be restated in terms of densities

$$f_t^P(x_{t,\tau_1}, \dots, x_{t,\tau_M}) = c_t(F_{t,\tau_1}^Q(x_{t,\tau_1}), \dots, F_{t,\tau_M}^Q(x_{t,\tau_M})) \prod_{j=1}^M f_{t,\tau_j}^Q(x_{t,\tau_j}) \quad (6)$$

where c_t is the mixed partial derivative of C_t with respect to all the arguments. Let $y_{jt} = F_{t,\tau_j}^Q(x_{t,\tau_j})$, $j = 1, \dots, M$. Then, in order to model the dependence of the prediction densities at different dates, our modelling assumption is a Beta dynamic Markov Random Field (β -MRF). Let $E = [0, 1]$ be the phase space and $S = \{1, \dots, M\}$ the finite set of sites (see Bremaud (1999), Ch. 7) corresponding to the different maturities, then our β -MRF is defined by the following local specification:

$$c_t(y_{1t}, \dots, y_{Mt}) = \frac{1}{Z_t} \prod_{j=1}^M c_{jt}(y_{jt} | y_{N(j)}) \quad (7)$$

where $y_{N(j)} = \{y_{kt}, k \in N(j) \subset S\}$ with $N(j)$ a member of the neighbourhood system N , c_{jt} represents the j th components of the joint calibration function c_t , and Z_t is a normalization function which may depend on the parameter of the calibration model and may be unknown for some β -MRF neighbourhood system specifications.

Modelling the full dependence structure between densities at the different maturities and allowing for time-change in this structure may require a large number of parameters, and consequently lead to over-parameterization (see, e.g. Litterman (1986), Sims

and Zha (1998) and Chib and Greenberg (1995)). The resulting potential over-fitting problem calls for the use of a Bayesian approach to inference, which allows for prior constraints on the model parameters. In order to reduce the number of parameters and to obtain a parsimonious β -MRF model, in this paper we assume a time-invariant topology (S, N) and focus on two special neighbourhood systems. The first one is a Markov model:

$$N(j) = \begin{cases} \emptyset & \text{if } j = 1, \\ \{j - 1\} & \text{if } j \neq 1, \end{cases}$$

and the second one is a proximity model

$$N(j) = \begin{cases} \{2\} & \text{if } j = 1, \\ \{j - 1, j + 1\} & \text{if } j \neq 1, M, \\ \{M - 1\} & \text{if } j = M \end{cases}$$

connecting each density with the two adjacent densities in terms of maturity.

Following the standard practice in the financial calibration literature (e.g. see Fackler and King (1990)) we assume that the j th component of the joint calibration function is the probability density function of a beta distribution. In order to account for possible time dependence in the PITs, we let the parameter of the beta calibration function of the density at maturity τ_j depend on the past values of the PITs for the same maturity. Note that this assumption of dependence only on adjacent densities is well supported in financial applications, where, conditional on the PIT of an adjacent maturity, the PIT is independent of PITs of other maturities (since the times-to-maturity are overlapping), and, conditional on the closest PIT on a given maturity, the PIT is independent of other PITs for that maturity (basic Markovian assumption for the underlying processes). We use the re-parametrization of beta pdfs used in Bayesian mixture models (e.g. see Robert and Rousseau (2002) and Bouguila et al. (2006)) and Bayesian beta autoregressive processes (e.g. see Casarin et al. (2012))

$$c_{jt}(y_{jt}|y_{N(j)}) = B_{jt} y_{jt}^{\mu_{jt}\gamma_{jt}-1} (1 - y_{jt})^{(1-\mu_{jt})\gamma_{jt}-1} \tag{8}$$

with

$$B_{jt} = \frac{\Gamma(\mu_{jt}\gamma_{jt})}{\Gamma(\mu_{jt}\gamma_{jt})\Gamma((1 - \mu_{jt})\gamma_{jt})}$$

and

$$\mu_{jt} = \varphi \left(\alpha_{0j} + \sum_{k=1}^p \alpha_{kj} y_{t-k,j} + \sum_{k \in N(j)} \beta_{kj} y_{t,k} \right), \tag{9}$$

$$\gamma_{jt} = \gamma_j \tag{10}$$

with $\varphi : \mathbb{R} \mapsto [0, 1]$ a twice differentiable strictly monotonic link function. We assume a logistic link function, i.e. $\varphi(x) = \frac{\exp(x)}{1+\exp(x)}$.

3 Bayesian inference

Let $\mathbf{x}_t = (x_{t,\tau_1}, \dots, x_{t,\tau_M})$ be a set of observations for different maturities, and $\mathbf{x}_{p+1:T} = (\mathbf{x}_{p+1}, \dots, \mathbf{x}_T)$, then the likelihood of the model can be written as

$$\begin{aligned} L(\mathbf{x}_{p+1:T}|\boldsymbol{\theta}) &= \prod_{t=p+1}^T f_{t,\tau}^P(x_{t,\tau_1}, \dots, x_{t,\tau_M}) \\ &= \prod_{t=p+1}^T \frac{1}{Z_t} \prod_{j=1}^M B_{jt}(\mu_{jt}\gamma_j, (1-\mu_{jt})\gamma_j) \left(F_{t,\tau_j}^Q(x_{t,\tau_j}) \right)^{\mu_{jt}\gamma_j-1} \\ &\quad \times \left(1 - F_{t,\tau_j}^Q(x_{t,\tau_j}) \right)^{(1-\mu_{jt})\gamma_j-1} f_{t,\tau_j}^Q(x_{t,\tau_j}). \end{aligned} \quad (11)$$

Note that this is a pseudo-likelihood, since we assume that the p initial values of the β -MRF are known.

In order to complete the description of our Beta Markov Random Field model, we assume the following hierarchical specification of the prior distribution. For a given j , with $j = 1, \dots, M$, we assume

$$\alpha_{kj} \stackrel{i.i.d.}{\sim} \mathcal{N}(\alpha_j, s_j^2), \quad k = 1, \dots, p, \quad (12)$$

$$\beta_{kj} \stackrel{i.i.d.}{\sim} \mathcal{N}(\beta_j, g_j^2), \quad k = 1, \dots, m_j. \quad (13)$$

For the second level of the hierarchy we assume

$$\gamma_j \stackrel{i.i.d.}{\sim} \mathcal{G}a(\xi_1, \xi_2), \quad j = 1, \dots, M, \quad (14)$$

$$\alpha_j \stackrel{i.i.d.}{\sim} \mathcal{N}(\alpha, s^2), \quad j = 1, \dots, M, \quad (15)$$

$$\beta_j \stackrel{i.i.d.}{\sim} \mathcal{N}(\beta, g^2), \quad j = 1, \dots, M \quad (16)$$

where $m_j = \text{Card}(N(j))$ is the number of elements of $N(j)$ and $\mathcal{G}a(\xi_1, \xi_2)$ denotes the gamma distribution with density

$$f(\gamma|\xi_1, \xi_2) = \frac{1}{\Gamma(\xi_1)} \gamma^{\xi_1-1} \exp\{-\xi_2\gamma\} \xi_2^{\xi_1}.$$

Moreover, in order to design an efficient algorithm for posterior simulation, we reparametrize $\sigma_j = \log(\gamma_j)$, $j = 1, \dots, M$. We will define the parameter vector $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M)$ where $\boldsymbol{\theta}_j = (\boldsymbol{\alpha}_j, \boldsymbol{\beta}_j, \sigma_j, \alpha_j, \beta_j)$, $\boldsymbol{\alpha}_j = (\alpha_{0j}, \alpha_{1j}, \dots, \alpha_{pj})$ and $\boldsymbol{\beta}_j = (\beta_{1j}, \dots, \beta_{m_jj})$. Then the joint probability density function of the prior distribution is

$$\begin{aligned} f(\boldsymbol{\theta}) \propto \exp \left\{ - \sum_{j=1}^M \left(\frac{1}{2s^2}(\boldsymbol{\alpha}_j - \boldsymbol{\alpha})^2 + \frac{1}{2g^2}(\boldsymbol{\beta}_j - \boldsymbol{\beta})^2 + \frac{1}{2}(\boldsymbol{\alpha}_j - \boldsymbol{\mu}_j)' S_j^{-1}(\boldsymbol{\alpha}_j - \boldsymbol{\mu}_j) \right. \right. \\ \left. \left. + \frac{1}{2}(\boldsymbol{\beta}_j - \boldsymbol{\nu}_j)' G_j^{-1}(\boldsymbol{\beta}_j - \boldsymbol{\nu}_j) \right) \right\} \prod_{j=1}^M \exp\{-\xi_1/2 \exp(\sigma_j)\} \exp(\xi_2/2\sigma_j) \end{aligned}$$

where $\boldsymbol{\mu}_j = \alpha_j \boldsymbol{\nu}_{(p+1)}$, $\boldsymbol{\nu}_j = \beta_j \boldsymbol{\nu}_{m_j}$, with $\boldsymbol{\nu}_n$ the n -dimensional unit vector. The prior covariance matrices are $S_j = s_j^2 I_{(p+1)}$ and $G_j = g_j^2 I_{m_j}$, with I_n the n -dimensional identity matrix.

The joint posterior distribution can be written as

$$\begin{aligned} \pi(\boldsymbol{\theta} | \mathbf{x}_{p+1:T}) \propto \exp & \left(- \sum_{t=p+1}^T \log Z_t - \sum_{t=p+1}^T \sum_{j=1}^M \log B_{jt} \right. \\ & \left. + \sum_{t=p+1}^T \sum_{j=1}^M \left(A_{jt} \mu_{jt} + \log(1 - F_{t,\tau_j}^Q(x_{t,\tau_j})) \right) \exp(\sigma_j) \right) f(\boldsymbol{\theta}) \end{aligned} \tag{17}$$

where

$$B_{jt} = B_{jt}(\mu_{jt} \exp(\sigma_j), (1 - \mu_{jt}) \exp(\sigma_j))$$

and

$$A_{jt} = \log(F_{t,\tau_j}^Q(x_{t,\tau_j}) / (1 - F_{t,\tau_j}^Q(x_{t,\tau_j}))).$$

A major problem with this model is that the normalizing constants Z_t , $t = p + 1, \dots, T$, in the likelihood function and in the posterior distribution are unknown and possibly depend on the parameters. Thus, samples from $\pi(\boldsymbol{\theta} | \mathbf{x}_{p+1:T})$ cannot be easily obtained with standard MCMC procedures. For instance, the standard MH algorithm cannot be directly applied because the acceptance probability involves ratios of unknown normalizing constants. In the last two decades, various approximation methods have been proposed in order to circumvent the problem of intractable normalizing constants. Recently, Møller et al. (2006) proposed an auxiliary variable MCMC algorithm, which is a feasible simulation procedure for many models with intractable normalizing constant. The Møller et al. (2006)'s single auxiliary variable method has been successfully improved by Murray et al. (2006). They propose the exchange algorithm, which removes the need to estimate the parameter before sampling begins and has higher acceptance probability than Møller et al. (2006)'s algorithm. Unfortunately, both the single auxiliary variable and the exchange algorithms require exact sampling of the auxiliary variable from its conditional distribution, which can be computationally expensive for many statistical models. An exact simulation algorithm for our β -MRF model is not available, thus in this paper we follow an alternative route and apply the double MH algorithm proposed by Liang (2010). For additional literature intractable normalizing constants, please, see Atchadé et al. (2013). The double MH avoids the exact simulation step by applying an internal MH step to generate the auxiliary variable.

Assume we are interested in simulating the auxiliary variable $\mathbf{z}_{p+1:T}$ from the conditional distribution $L(\mathbf{z}_{p+1:T} | \boldsymbol{\theta}')$. If the sample is generated by iterating n times an MH algorithm with transition kernel $K_{\boldsymbol{\theta}'}(\mathbf{z} | \mathbf{x})$, then the n -step transition probability is

$$P_{\boldsymbol{\theta}'}^n(\mathbf{z}_{p+1:T} | \mathbf{x}_{p+1:T}) = K_{\boldsymbol{\theta}'}(\mathbf{x}_{p+1:T}^1 | \mathbf{x}_{p+1:T}) \cdots K_{\boldsymbol{\theta}'}(\mathbf{z}_{p+1:T} | \mathbf{x}_{p+1:T}^{n-1}),$$

and the acceptance rate of the Murray et al. (2006)'s exchange algorithm writes as

$$\rho(\boldsymbol{\theta}, \boldsymbol{\theta}', \mathbf{z}_{p+1:T} | \mathbf{x}_{p+1:T}) = \frac{f(\boldsymbol{\theta})q(\boldsymbol{\theta}' | \boldsymbol{\theta}, \mathbf{x}_{p+1:T})}{f(\boldsymbol{\theta}')q(\boldsymbol{\theta} | \boldsymbol{\theta}', \mathbf{x}_{p+1:T})} \frac{L(\mathbf{z}_{p+1:T} | \boldsymbol{\theta})}{L(\mathbf{x}_{p+1:T} | \boldsymbol{\theta})} \frac{P_{\boldsymbol{\theta}'}^n(\mathbf{z}_{p+1:T} | \mathbf{x}_{p+1:T})}{P_{\boldsymbol{\theta}'}^n(\mathbf{x}_{p+1:T} | \mathbf{z}_{p+1:T})} \tag{18}$$

If we chose $q(\boldsymbol{\theta}|\boldsymbol{\theta}', \mathbf{x}_{p+1:T})$ as a Metropolis transition kernel then the exchange is an MH step with transition $P_{\boldsymbol{\theta}'}^n(\mathbf{z}_{p+1:T}|\mathbf{x}_{p+1:T})$ and target distribution $L(\mathbf{z}_{p+1:T}|\boldsymbol{\theta})$, and the acceptance probability in (18) becomes

$$\rho(\boldsymbol{\theta}, \boldsymbol{\theta}', \mathbf{z}_{p+1:T}|\mathbf{x}_{p+1:T}) = \frac{L(\mathbf{z}_{p+1:T}|\boldsymbol{\theta}) L(\mathbf{x}_{p+1:T}|\boldsymbol{\theta}')}{L(\mathbf{x}_{p+1:T}|\boldsymbol{\theta}) L(\mathbf{z}_{p+1:T}|\boldsymbol{\theta}')}. \quad (19)$$

Assume the current value of the MH chain is $\boldsymbol{\theta}^{(t)} = \boldsymbol{\theta}$, then the double MH sampler iterates over the following steps:

1. Simulate a new sample $\boldsymbol{\theta}'$ from $\pi(\boldsymbol{\theta})$ using an MH algorithm starting with $\boldsymbol{\theta}$.
2. Generate the auxiliary variable $\mathbf{z}_{p+1:T} \sim P_{\boldsymbol{\theta}'}^n(\mathbf{z}_{p+1:T}|\mathbf{x}_{p+1:T})$ and accept it with probability $\min\{1, \rho(\boldsymbol{\theta}, \boldsymbol{\theta}', \mathbf{z}_{p+1:T}|\mathbf{x}_{p+1:T})\}$ given in (19).
3. Set $\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}'$ if the auxiliary variable is accepted and $\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}$ otherwise.

For the first MH step in the double MH, we assume a multivariate random-walk proposal, i.e.

$$\boldsymbol{\theta}^* \sim \mathcal{N}(\boldsymbol{\theta}^{(t)}, \Lambda)$$

where Λ is an n -dimensional positive diagonal matrix with $n = (p+4)M + m + 2$.

Regarding the second MH step, we consider a Gibbs sampler which generates samples from the full conditional distributions of each site.

For $t = p+1, \dots, T$ and for $j = 1, \dots, M$ the Gibbs sampler generates iteratively $z_{t, \tau_j} \sim \pi(z_{t, \tau_j} | \mathbf{z}_{t-p:t-1, \tau_j}, \mathbf{z}_{t+1:t+p, \tau_j}, \mathbf{z}_{t-p:t, \tau_{j-1}}, \mathbf{z}_{t-p:t, \tau_{j+1}}, \boldsymbol{\theta})$, where $\mathbf{z}_{s:t, \tau_j} = (z_{s, \tau_j}, \dots, z_{t, \tau_j})$. The full conditionals are not trivial to simulate from, thus we apply an MH step. See Appendix B for further details on the derivation of the full conditional distributions and on the sampling method.

4 Simulation exercise

The extraction of parametric and nonparametric risk-neutral densities has been important not only for traders in order to use this density to price more exotic derivatives, but also for central bankers as well and policy makers (Campa et al., 1997; Vergote and Puigvert, 2012; Sihvonen and Vähämaa, 2014). Recently, a great deal of interest has grown in predicting both the nonparametric risk-neutral and its physical counterpart simultaneously, as shown in Vesela and Puigvert (2014) for the 3-month Euribor interest rate, using the beta calibration function for fixed expirations of the nonparametric risk-neutral density instead of constant and rolling maturity expirations such as 3,6,9, and 12 months as in Vergote and Puigvert (2012). These constant maturity risk-neutral densities are interpolated in practice from fixed expiration densities as done in Vergote and Puigvert (2012). We do not need to follow their approach, which is constrained by the chosen maturity, since we can model rolling, fixed time to maturity (as opposed to

fixed maturity), option prices, and forwards. In this sense, our approach is more generalizable for over-the-counter markets, where the market centres around (rolling tenor) fixed time-to-maturity points rather than fixed expiry dates. This is usually the case in the over-the-counter markets, including the foreign exchange market.

In this section, we produce several simulation exercises to test the accuracy of our method to produce a calibration function that allows for better assessment of the non-standard features usually encountered in the PIT data, including through mis-estimation of the underlying parameters of the process. This exercise consists of several layers according to the following sequence:

- We produce the simulated data x_{t,τ_j} , $t = 1, \dots, T$ and $j = 1, \dots, M$, under the physical measure, which will be common to all the simulation exercises. The data are asset prices $S_{t+\tau_j}$ simulated from a Geometric Brownian Motion (GBM), i.e. $x_{t,\tau_j} = S_{t+\tau_j}$, under the physical measure for $M = 3$ maturities, that are 3 months (τ_1), 6 months (τ_2) and 12 months (τ_3) for a time interval of 2 years, that is, $T = 504$ trading days, $\mu = 0.20$, $r = 0.05$, $\sigma = 0.15$, $\tau_1 = 0.25$ (years), $\tau_2 = 0.5$ (years) and $\tau_3 = 1$ (year).
- From that data, we estimate the risk neutral measure, assuming that we incorrectly estimate (grossly) the parameters of this risk neutral measure. For this purpose, we assume two potential scenarios that cover the two extremes (note that, in both cases, the values that we are using are different from μ):
 1. (Overestimation of the volatility of the Brownian Motion) We will assume for the calibration exercise that we overestimate the unknown volatility of the physical process and set $\sigma = 0.20$.
 2. (Underestimation of the volatility of the Brownian Motion) We will assume for the calibration exercise that we underestimate the unknown volatility of the physical process and set $\sigma = 0.10$.
- For each of the cases above, and for each of the maturities in the simulation exercise, we compare two curves (see Figure 1). Note that in both cases we use the same playing field for the volatility, to ensure that the comparison is done solely on the calibration benefits:
 1. (NC Curve) This is the non-calibrated curve and corresponds to the empirical CDF of the data $F_{t,\tau_j}^Q(x_{t,\tau_j})$, $t = p + 1, \dots, T$. It simply states the shape of the PITs CDF using the risk neutral data, under the stated value of the volatility.
 2. (C Curve) This is the empirical CDF of the data $F_{t,\tau_j}^P(x_{t,\tau_j})$, $t = p + 1, \dots, T$, where F_{t,τ_j}^P is the marginal distribution of the distribution $F_t^P(x_{t,\tau_1}, \dots, x_{t,\tau_M})$, obtained by using the β -MRF calibration applied to the risk neutral data, under the stated value of the volatility.
- As a reference, we include the 45 degree line, which represents the perfect scenario where the (calibrated) PITs are uniformly distributed.

In order to run this simulation, we assume that the data comes from a standard process in the financial literature (GBM), with S_t , $t \in [0, T]$, to model the price of the underlying as in Black and Scholes (1973) and Merton (1973), i.e.

$$S_t = S_0 + \int_0^t S_u \mu du + \int_0^t S_u \sigma dW(u) \quad (20)$$

where $W(t)$, $t \in [0, T]$, is a Wiener process.

We simulate price sample paths under the physical measure for 3, 6 and 12 months for a time interval of $T = 2$ years, $\mu = 0.20$, $r = 0.05$, $\sigma = 0.15$, $\tau_1 = 0.25$, $\tau_2 = 0.5$, and $\tau_3 = 1$.

We also know analytically the risk-neutral densities of $S_{t+\tau_j}$, $j = 1, 2, 3$, conditional on S_t , which are given by

$$f_{t,\tau_j}^Q(S_{t+\tau_j}) = \frac{1}{S_{t+\tau_j} \sqrt{2\pi\sigma^2\tau_j}} \exp \left[-\frac{[\log(S_{t+\tau_j}/S_t) - (r - 0.5\sigma^2)\tau_j]}{2\sigma^2\tau_j} \right], \quad (21)$$

$j = 1, 2, 3$. Once we observe 3 months later a price level of $S_{t+\tau_j}$ under the historical measure, then we proceed to compute the 3, 6 and 12 months PITs at time t as follows:

$$y_{t,\tau_j} = \int_{-\infty}^{S_{t+\tau_j}} f_{t,\tau_j}^Q(S_{t+\tau_j}) dS_{t+\tau_j}, \quad (22)$$

$j = 1, 2, 3$. The next period, at time $t + 1$, we recompute the PITs in the same way as in (22), obtaining a vector $\mathbf{x}_t = (x_t, x_{t+1}, \dots, x_{t+T})$ where again

$$\mathbf{x}_s = (F_{s,\tau_1}^Q(x_{s,\tau_1}), F_{s,\tau_2}^Q(x_{s,\tau_2}), F_{s,\tau_3}^Q(x_{s,\tau_3})),$$

and where the components of \mathbf{x}_s will be very likely correlated, given the overlapping times to maturity. In our simulation exercise, we assume that a year has 252 trading days (prices) and that 3 (6 and 12) months correspond to 63 (126 and 252) trading days, respectively.

A uniform marginal distribution of the PITs, assuming that they are not autocorrelated, indicates that there is no need for a calibration function. A uniform marginal distribution of the PITs, assuming that they are autocorrelated, does not necessarily say anything about the need for a calibration function. There could be cases where the PITs are extremely autocorrelated, and yet display a perfect uniform histogram leading to the wrong conclusion that both the risk neutral and physical measures are both identical.

The source of autocorrelation of the PITs comes from the rolling nature of the data. For each period t , we obtain a new PIT for each maturity, which is the outcome of the physical process under that given maturity. Since, for a given maturity τ , we will be producing $\tau \times 252$ overlapping periods (with different levels of overlap), these periods will share common contributions to each of those PITs. For example, a 3 month PIT with reference point, and maturity in 65 business days (3 months), will share

64 business days in common with another PIT, with reference point tomorrow, and maturity 65 days from tomorrow. This generates an artificial autocorrelation in the PITs, which is embedded in any overlapping data. Classical approaches include a mere thinning of the data to the frequency of the underlying maturity, to take only non-overlapping periods Vergote and Puigvert (2012). However, this approach is especially penalizing on the longer maturities. For example, for maturities of a year, traditional approaches will only collect one data point per year. Our approach is less restrictive, since it takes into account in the modelling the different sources of correlation (including this autocorrelation) between the PITs through the β -MRF approach. For two given PITs (A, B), for which the data driving them is represented by the combination of the starting points t_A, t_B , and the maturities τ_A, τ_B , the overlapping amounts of information contained in the physical process is the intersection of $[t_A, t_A + \tau_A] \cap [t_B, t_B + \tau_B]$. This information is processed naturally through the β -MRF approach, which takes into account the two causes of autocorrelation (over time and across neighbours).

We apply our Bayesian β -MRF calibration model with the following hyperparameter settings $\alpha = 0$, $\beta = 0$, $s_j^2 = 10$, $g_j^2 = 10$, $s^2 = 100$, and $g = 100$. We apply the proposed MCMC algorithm in order to approximate the posterior quantities of interest. In the MCMC algorithm, we consider 5,000 iterations after convergence (that is detected after about 2,000 burn-in iterations by applying the Geweke (1992) convergence diagnostic test statistics). The scale Λ of the proposal distribution of the MH step for generating θ from q was chosen to achieve average acceptance rates between 0.5 and 0.7 for the two MH algorithms (steps 1 and 2), which is a good sign of efficiency for most MCMC algorithms, as suggested, for example, by Rosenthal (2011). This choice can be done 'on-line' through the runs of the algorithm in the burn-in phase.

With regards to Table 1:

- α_j are the autoregressive parts of the β -MRF (time factor) – representing the time-dependence.
- β_j are the parameters linking the different maturities (maturity factor), representing the cross-maturity dependence.

The autocorrelation over time decreases as the maturity increases. This feature can be seen in the value of the corresponding parameters α . Additionally, β_1 and β_2 represent the correlation parameters of neighbouring maturities before and after, respectively. Hence β_{12} represents the correlation parameters between maturity 1 and maturity 2, while β_{23} represents the correlation parameter between maturity 2 and maturity 3. Furthermore, panels c and d pool across maturities. This pooling produces an interesting practical approach because it assumes the same autoregressive structure over time for the PITs across their maturities, should this be considered a more sensible modelling approach.

The results of the calibration exercises are given in Table 1 and Figure 1. The autoregressive coefficient is significant for all maturities. The proximity parameter is significant only for the last maturity. The value of the precision parameter increases

Panel (a) ($\sigma = 0.1$)

	$\tau_j, j = 1$		$\tau_j, j = 2$		$\tau_j, j = 3$	
θ_{ij}	$\hat{\theta}_{ij}$	CI	$\hat{\theta}_{ij}$	CI	$\hat{\theta}_{ij}$	CI
γ_j	1.42	(1.01, 1.51)	2.82	(2.79, 2.93)	13.46	(13.40, 13.64)
α_{0j}	-0.32	(-0.44, -0.25)	-0.55	(-0.64, -0.46)	-1.09	(-1.15, -1.02)
α_{1j}	0.43	(0.32, 0.48)	0.51	(0.35, 0.61)	0.32	(0.23, 0.42)
β_{1j}			0.11	(0.01, 0.21)	0.16	(0.04, 0.26)
β_{2j}	0.18	(0.06, 0.27)	0.03	(0.01, 0.15)		

Panel (b) ($\sigma = 0.2$)

	$\tau_j, j = 1$		$\tau_j, j = 2$		$\tau_j, j = 3$	
θ_{ij}	$\hat{\theta}_{ij}$	CI	$\hat{\theta}_{ij}$	CI	$\hat{\theta}_{ij}$	CI
γ_j	3.75	(3.73, 3.81)	7.01	(6.67, 7.23)	14.03	(13.83, 14.16)
α_{0j}	-0.24	(-0.34, -0.11)	-0.11	(-0.19, -0.03)	-0.23	(-0.29, -0.18)
α_{1j}	0.37	(0.23, 0.47)	0.30	(0.24, 0.41)	0.47	(0.32, 0.58)
β_{1j}			0.37	(0.27, 0.43)	0.05	(-0.09, 0.21)
β_{2j}	0.13	(0.03, 0.21)	-0.02	(-0.09, 0.08)		

Panel (c) ($\sigma = 0.1$)

	$\tau_j, j = 1, 2, 3$	
γ	99.4	(42.7, 171.81)
α_0	0.17	(-5.25, 7.19)
α_1	-0.02	(-8.37, 5.21)
β_1	-0.74	(-6.79, 5.98)
β_2	0.56	(-5.58, 7.55)

Panel (d) ($\sigma = 0.2$)

	$\tau_j, j = 1, 2, 3$	
γ	95.3	(49.57, 159.66)
α_0	1.55	(-4.43, 7.15)
α_1	-0.48	(-5.8, 4.17)
β_1	-0.37	(-5.81, 4.27)
β_2	-0.04	(-7.48, 7.77)

Table 1: Posterior mean ($\hat{\theta}_i$) and 95% credible intervals (CI), for the parameters of the β -MRF. The non-calibrated predictive models with $\sigma = 0.1$ (panels (a) for the hierarchical and (c) for the pooled model) and $\sigma = 0.2$ (panels (b) for the hierarchical and (d) for the pooled model), when the true value of the scale parameter is $\sigma = 0.15$.

with the maturity. Figure 1 shows the non-calibrated and calibrated PITs. Figure 2 shows the predictive density, and the calibrated predictive, for the prices at time $t = 504$ using the implied densities available at time $t - \tau_j$ for different j (rows) and different wrong values of the volatility parameter σ (columns).

We also consider a more parsimonious model, where we assume $\beta_{kj} = \beta_k$ and $\alpha_{kj} = \alpha_k$ for all $j = 1, \dots, M$. The results are also given in Table 1.

5 Foreign exchange market application

We apply our methodology to over-the-counter annualized implied volatilities for options on the EUR/USD (underlying being the currency spot level) for different tenors (one month, two months, and six months), spanning from 01/Jan/2010 until 01/Apr/2013.

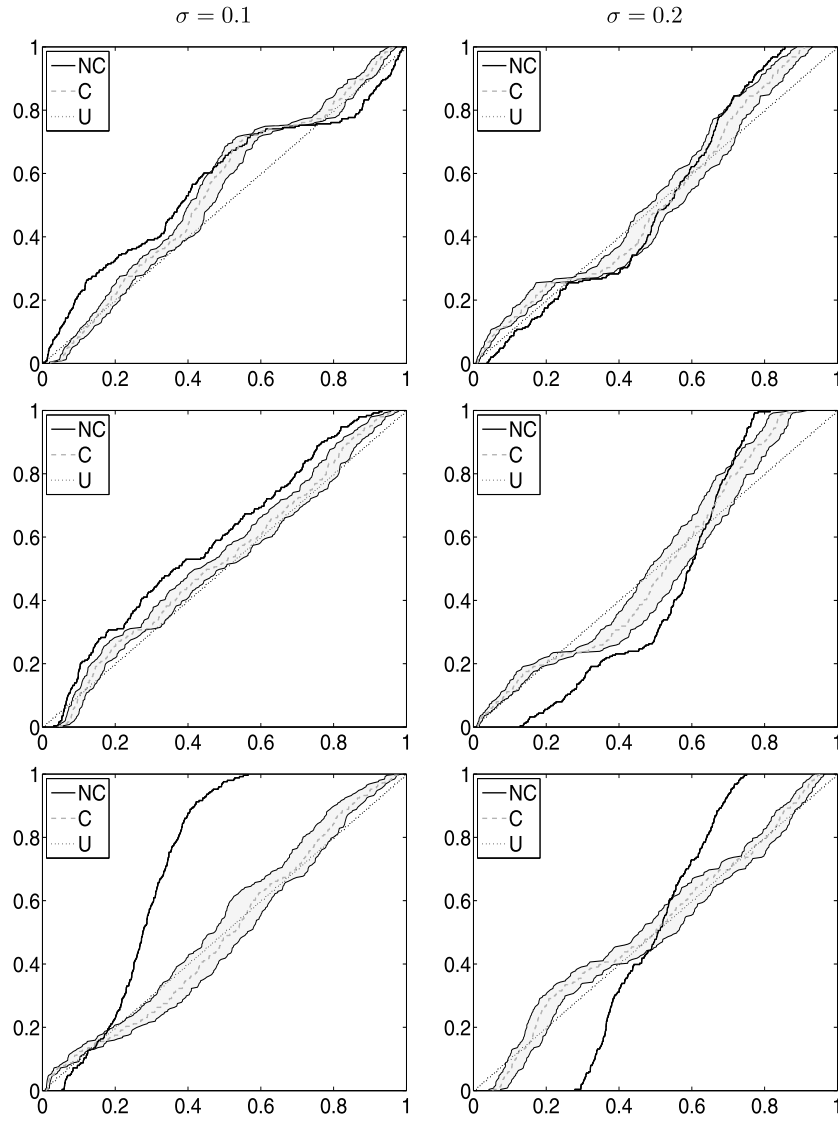


Figure 1: Non-calibrated (dashed line) and calibrated (solid line) risk neutral distribution for different maturities (rows) and volatility levels (columns). In each plot, thin solid lines represent the 95% HPD region.

For the computation of the risk neutral densities, we applied the same procedure consisting of first fitting a spline to the implied volatility for each tenor separately as in Panigirtzoglou and Skiadopoulos (2004); Vergote and Puigvert (2012), in order to transform back to the option price space and take the second derivative to yield the risk

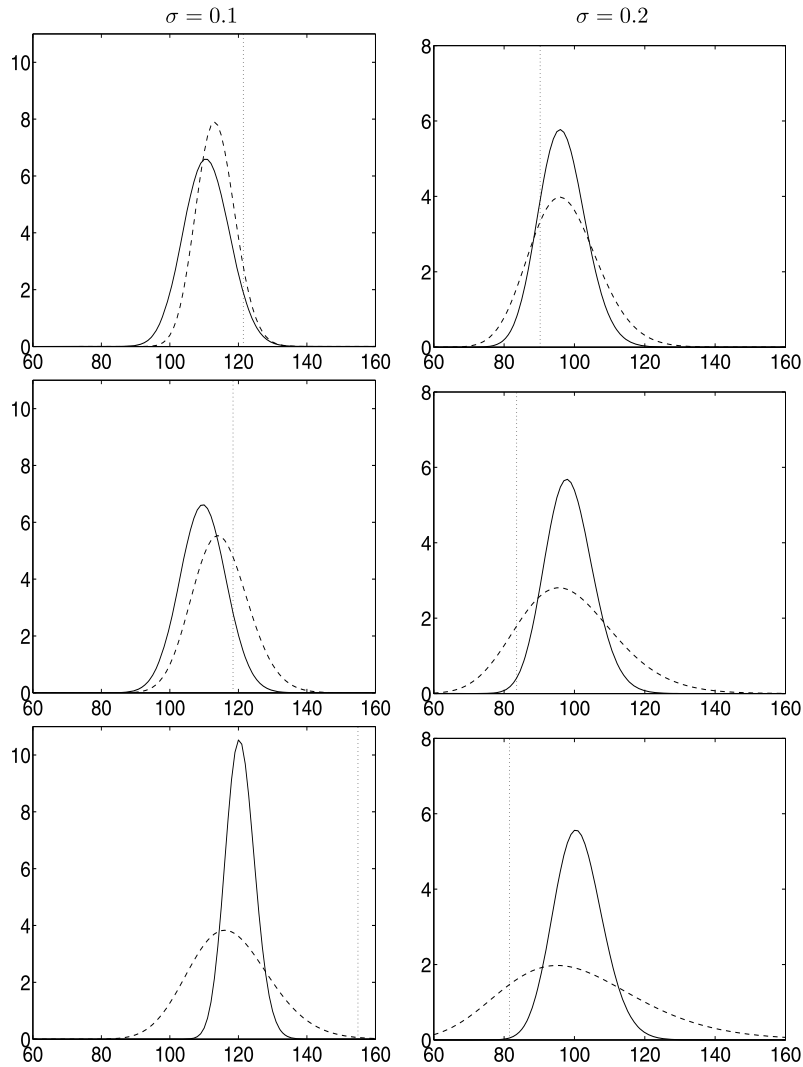


Figure 2: Non-calibrated (dashed line) and calibrated (solid line) risk neutral distribution and price level (vertical dotted line) at last point of the sample, i.e. $t = 504$, for different maturities (rows) and different volatility levels (columns).

neutral density.² For an extensive review on how to extract risk-neutral densities from option prices with Matlab code included, see Fusai and Roncoroni (2000). We apply this methodology, both for the case where we assume that each tenor has its own calibration function, and for the case where we assume that there is a single calibration function

²A more thorough description of how to estimate the risk-neutral density obtained in our work is given in Appendix A.

Panel (a) (Original data sample)

	$\tau_j, j = 1$		$\tau_j, j = 2$		$\tau_j, j = 3$	
θ_{ij}	$\hat{\theta}_{ij}$	CI	$\hat{\theta}_{ij}$	CI	$\hat{\theta}_{ij}$	CI
γ_j	2.24	(2.14, 3.01)	2.78	(2.76, 2.98)	3.55	(3.42, 3.97)
α_{0j}	-0.04	(-0.16, 0.09)	-0.06	(-0.21, 0.07)	-0.08	(-0.24, 0.04)
α_{1j}	0.15	(0.06, 0.24)	0.31	(0.14, 0.45)	0.31	(0.21, 0.52)
β_{1j}			0.14	(0.04, 0.28)	0.13	(-0.01, 0.26)
β_{2j}	0.17	(0.04, 0.28)	-0.01	(-0.2, 0.01)		

Panel (b) (Thinned data sample)

	$\tau_j, j = 1$		$\tau_j, j = 2$		$\tau_j, j = 3$	
θ_{ij}	$\hat{\theta}_{ij}$	CI	$\hat{\theta}_{ij}$	CI	$\hat{\theta}_{ij}$	CI
γ_j	2.63	(2.52, 2.78)	2.57	(2.34, 2.71)	3.52	(3.48, 3.62)
α_{0j}	0.03	(-0.14, 0.23)	0.04	(-0.15, 0.18)	0.02	(-0.17, 0.21)
α_{1j}	0.05	(-0.25, 0.21)	0.10	(-0.04, 0.33)	0.11	(-0.02, 0.23)
β_{1j}			0.06	(0.01, 0.26)	0.09	(-0.01, 0.34)
β_{2j}	0.07	(-0.10, 0.32)	0.03	(-0.14, 0.23)		

Table 2: Posterior mean ($\hat{\theta}_i$) and 95% credible intervals (CI), for the parameters of the β -MRF. The non-calibrated and β -MRF calibrated predictive pits empirical distribution function for original data (panel (a)) and thinned data (panel (b)), with thinning factor 100/15.

that works across several tensors by setting $\beta_{kj} = 0$ in the specification of μ_{jt} .

Figure 3 shows the time series (left column) and the histograms (right column) of the different PIT series.

We further display below the risk neutral densities estimated on the last day of the sample, 01/Apr/2013, for the different maturities as well as their physical densities, computed by applying the calibration function to each of the risk-neutral densities. We apply our β -MRF calibration model with the prior and MCMC setting used in the simulated experiments (see previous section).

The results are given in Figure 4. As it results from panel (a) in Table 2, we found evidence of autocorrelation component (coefficient α_{1j}) and of dependence across neighbouring maturities (coefficients β_{ij}). From panel (b) of the same figure, we can see that the value of the autoregressive coefficient decreases when thinning (thinning factor 100/15) is applied to the PITs time series in order to reduce the dependence between the samples.

In order to test whether the physical density before and after the estimation of the distortion (calibration) Beta function is adequate or not, we follow Bliss and Panigirtzoglou (2004) and apply the Berkowitz test. The Berkowitz test (Berkowitz, 2001) consists in transforming the PITs to normal random variables through the following approach:

$$z_{t,\tau} = \Phi^{-1}(y_{t,\tau}) \quad (23)$$

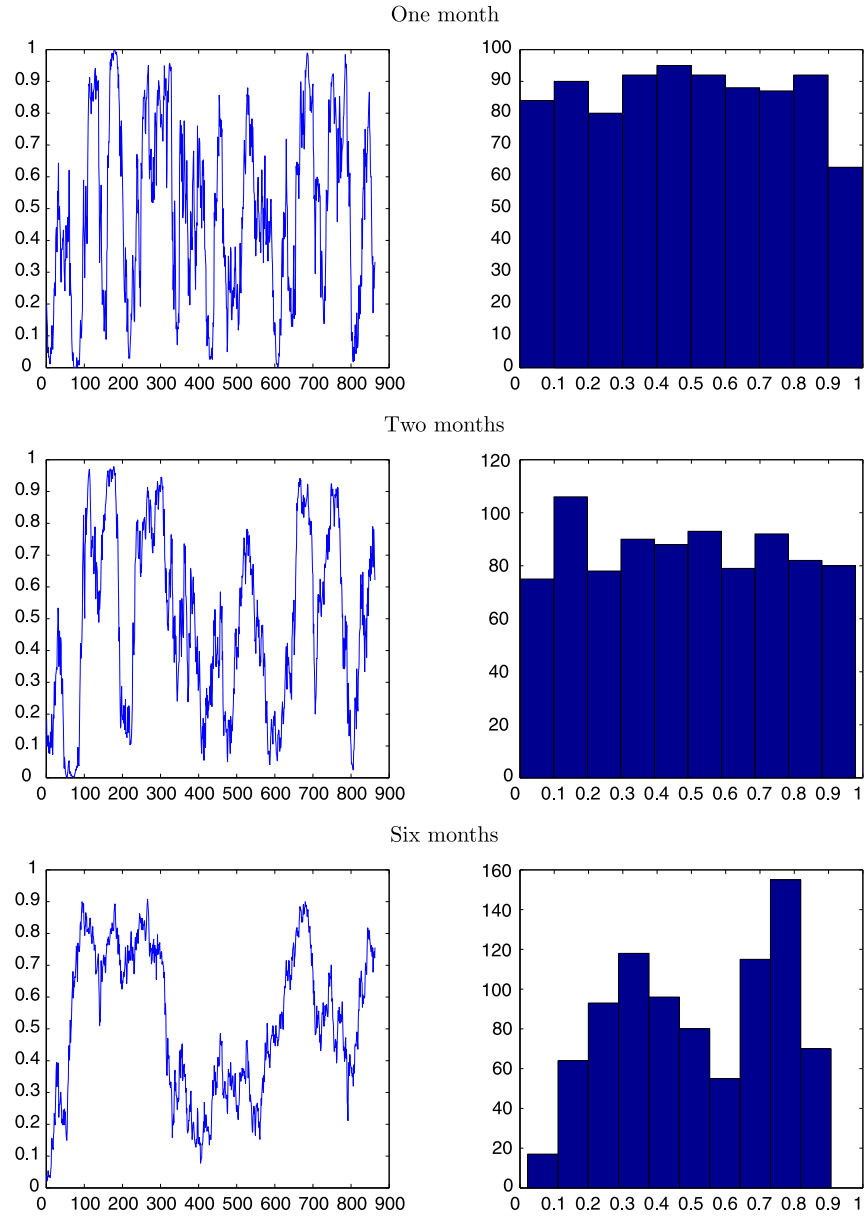


Figure 3: PIT time series (first column) and unnormalized histogram of the non-calibrated PITs (second column).

where Φ^{-1} is the inverse CDF of a standard normally distributed random variable. Under the null hypothesis the $z_{t,\tau}$ have to be iid standard normally distributed random variables, meaning that when estimating the model $z_{t,\tau} - \nu_\tau = \rho_\tau(z_{t-1,\tau} - \nu_\tau) + \epsilon_{t,\tau}$,

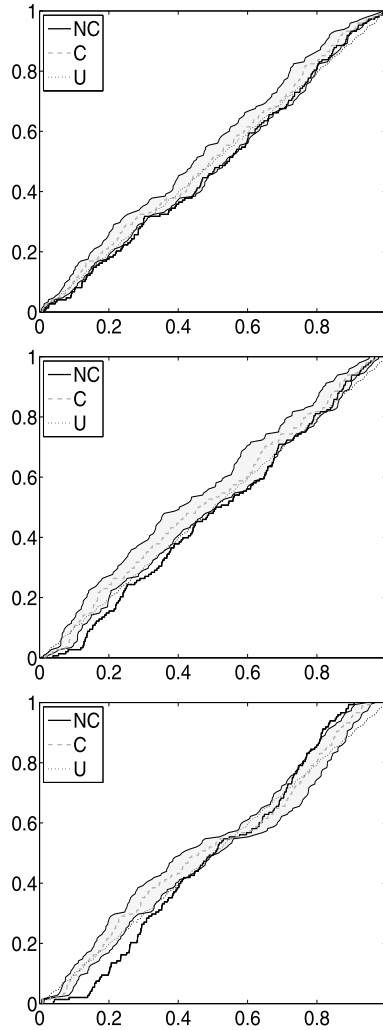


Figure 4: Non-calibrated (solid line), calibrated (dotted line), and β -MRF calibrated (thinner dashed line) risk neutral distributions for the three different maturities (rows): one, two and six months. In each plot, thinner solid lines represent the 95% HPD region.

with $\epsilon_{t,\tau} \sim \mathcal{N}(0, \zeta_\tau^2)$, we would obtain $\rho_\tau = 0$, $\nu_\tau = 0$ and $\zeta_\tau^2 = 1$. The likelihood of this autoregressive normal model is written as $L(\nu_\tau, \zeta_\tau, \rho_\tau)$, and we use the statistic

$$LR_3 = -2 \left(L(0, 1, 0) - L(\hat{\nu}_\tau, \hat{\zeta}_\tau, \hat{\rho}_\tau) \right) \tag{24}$$

to test the null hypothesis of independence and normality of the $z_{t,\tau}$ (Berkowitz, 2001). We report the p -values for different maturities under the null hypothesis. The LR_3 statistic has a $\chi^2(3)$, that is, a chi-square distribution with 3 degrees of freedom, under

Maturity	LR ₃			LR ₁		
	NC	β -C	β -MRF-C	NC	β -C	β -MRF-C
1 month	0.003	0.092	0.153	0.001	0.023	0.063
3 months	0.001	0.017	0.157	0.001	0.001	0.056
6 months	0.001	0.001	0.223	0.001	0.001	0.074

Table 3: Berkowitz tests (p -value) for uniformity and independence, LR₃ and for independence, LR₁, for different maturities (rows), and different PITs series (columns): physical distribution PITs or non-calibrated PITs (NC), beta calibrated PITs (β -C) and β -MRF calibrated PITs (β -MRF-C)

the null where $\rho = 0$, $\nu = 0$ and $\text{Var}(\epsilon_t) = 1$. Moreover, as suggested in (Berkowitz, 2001), we also test the independence assumption separately by evaluating the statistics

$$LR_1 = -2(L(\hat{\nu}_\tau, \hat{\zeta}_\tau, 0) - L(\hat{\nu}_\tau, \hat{\zeta}_\tau, \hat{\rho})) \quad (25)$$

which follows a $\chi^2(1)$ under the null hypothesis where $\rho = 0$.

As suggested by the graphical evidence of Figure 3, when performing the Berkowitz test for the physical density, the null hypothesis stating that $\rho_\tau = 0$, $\nu_\tau = 0$ and $\zeta_\tau = 1$ for all different maturities was rejected with p -values equal to 0 (see columns NC in Table 3).

Also, we compare the physical density PITs with the beta calibrated PITs (see columns β -C in Table 3), which correspond to the physical density PITs calibrated with a beta distortion. It can be obtained as a restricted version of our β -MRF model by setting $\alpha_{kj} = 0$ and $\beta_{kj} = 0$ for all k and j . The β -C PITs are rejected as good forecast for the 3 month and 6 month maturities. However, when we examine the p -values of the LR₁ tests for autocorrelation we reject the null hypothesis that the underlying PITs are uncorrelated, suggesting that the rejection of LR₃ is potentially due to the lack of independence in the PITs. For the β -MRF-C PITs the absence of autocorrelation (see LR₁ p -values) cannot be rejected, and the hypothesis that they provide good forecasts cannot be rejected for all maturities by comfortable margins.

6 Conclusion

This paper, which builds on the methodology by Casarin et al. (2012), provides a new modelling framework using both the derivative implied volatilities and synchronized spot and forwards for the term structure of the implied probability, which accounts for the possible dependence between PITs at different maturities, and different dates for a given maturity. This approach allows for borrowing of information between the different tenors, for both the risk-neutral and the physical measures. We also provide an inferential Bayesian framework that allows the inclusion of uncertainty in the density calibration functions, normally a factor overlooked in the literature, and therefore also in the physical densities.

Modelling the time evolution of the predictive densities and the relationship between densities from many sources is a challenging issue. For example, in traditional

approaches, when reconstructing the calibration function there cannot be any overlapping time intervals so that the PITs are independent in order to estimate the beta calibration function as explained in Fackler and King (1990) and later used in Vergote and Puigvert (2012), and Vesela and Puigvert (2014). Using independent PITs has the drawback of requiring large amounts of data (even more for longer maturities), not always available for new assets or assets that do not trade frequently, to have a reliable calibration function. Our methodology takes advantage of using all sources of information available, without thinning of the source, since the induced correlation is incorporated in the modelling of the dependent PITs.

Future research will include adapting and comparing our methodology to other assets, such as stocks, commodities, fixed income indices, and other exchange-traded markets, which trade on fixed expiry contracts rather than rolling constant term contracts, by interpolating the risk neutral densities for different constant maturities from fixed expiry contracts as done in (Vergote and Puigvert, 2012), as well as applications of this method outside the financial world.

Appendix A: Risk neutral density estimation

Our application consists of the daily implied annualized volatility on the Eurodollar currency for different (constant maturity, rather than fixed day of expiry) expirations. The full data consists of the closing snapshots around the end of the London business day for spot, forwards and implied volatilities for the period of the 01/Jan/2010 to the 01/Apr/2013. We follow Campa et al. (1998); Vergote and Puigvert (2012) and fit a cubic smoothing spline to the volatility smile. We work in the implied volatility space which provides a natural interpolation setting for the full surface, especially as most of the higher-quality data comes from the more-traded out-of-the-money options. By fitting the implied volatility (sigma-delta) instead of the option prices directly, one is able to circumvent the latter problem of the noise in the options data, as in Shimko (1993); Hutchinson et al. (1994); Malz (1997); Ait-Sahalia and Lo (1998a); Engle and Rosenberg (2000); Bliss and Panigirtzoglou (2002). Additionally, this is the natural approach in over-the-counter options markets, where option prices are often quoted by market-makers in volatility space, rather than price space.

Using the same notation as in Vergote and Puigvert (2012), the optimal cubic smoothing spline of the implied volatility is the one that minimizes the following function:

$$\min \lambda \sum_{i=1}^n \omega_i (\sigma_i - \widehat{\sigma(\Theta)}_i)^2 + (1 - \lambda) \int_0^1 g''(\delta, \Theta) d\delta \quad (26)$$

where δ is the partial derivative of the Black and Scholes option call price with respect to the underlying, σ_i , $\widehat{\sigma(\Theta)}$, and $\omega_i = \frac{\nu_i}{\sum_i \nu_i}$ are the observed implied volatility, fitted volatility, and weight of observation i , together with its Greek Black and Scholes vega ν , respectively. Furthermore, Θ represents the matrix of polynomial parameters of the cubic spline while $g(\cdot, \cdot)$ is the cubic spline function. An arbitrary value of $\lambda = 0.99$ is chosen as standard in the literature, as seen, for example, in Vergote and Puigvert

(2012), Bliss and Panigirtzoglou (2004) or Bliss and Panigirtzoglou (2002). It is worthwhile noting that the Black–Scholes formula is used solely to convert the option prices into/from their implied volatilities, in order to make the smoothing more effectively. This approach does not imply that we are assuming the Black and Scholes pricing formula is the correct one, but is only a way to make the smoothing more effective in the interpolation.³

Appendix B: Computational details

In order to generate $\mathbf{z}_{p+1:T}$ from $L(\mathbf{z}_{p+1:T}|\boldsymbol{\theta})$ we apply an MH within Gibbs algorithm. The full conditional distributions of the Gibbs sampler are derived by using the Markov property of our dynamic random field with respect to the chosen neighbourhood system. Let $B_{js} = B_{js}(\mu_{js}\gamma_j, (1 - \mu_{js})\gamma_j)$, then the full conditional distribution of the j th site, for $t = p + 1, \dots, T - p$ is

$$\begin{aligned}
& \pi(z_{t,\tau_j} | \mathbf{z}_{t-p:t-1,\tau_j}, \mathbf{z}_{t+1:t+p,\tau_j}, \mathbf{z}_{t-p:t,\tau_{j-1}}, \mathbf{z}_{t-p:t,\tau_{j+1}}, \boldsymbol{\theta}) \\
& \propto L(\mathbf{z}_{p+1:T}|\boldsymbol{\theta}) \\
& \propto \prod_{s=p+1}^T \prod_{j=1}^M B_{js} \left(F_{s,\tau_j}^Q(z_{s,\tau_j}) \right)^{\mu_{js}\gamma_j-1} \left(1 - F_{s,\tau_j}^Q(z_{s,\tau_j}) \right)^{(1-\mu_{js})\gamma_j-1} f_{s,\tau_j}^Q(z_{s,\tau_j}) \\
& \propto \prod_{s=t}^{t+p} \prod_{j=1}^M B_{js} \left(F_{s,\tau_j}^Q(z_{s,\tau_j}) \right)^{\mu_{js}\gamma_j-1} \left(1 - F_{s,\tau_j}^Q(z_{s,\tau_j}) \right)^{(1-\mu_{js})\gamma_j-1} f_{s,\tau_j}^Q(z_{s,\tau_j}) \\
& \propto \prod_{s=t}^{t+p} \prod_{i \in N(j)} B_{is} \left(F_{s,\tau_i}^Q(z_{s,\tau_i}) \right)^{\mu_{is}\gamma_i-1} \left(1 - F_{s,\tau_i}^Q(z_{s,\tau_i}) \right)^{(1-\mu_{is})\gamma_i-1} f_{s,\tau_i}^Q(z_{s,\tau_i}) \\
& \quad \times \prod_{l \notin N(j), l \neq j} B_{ls} \left(F_{s,\tau_l}^Q(z_{s,\tau_l}) \right)^{\mu_{ls}\gamma_l-1} \left(1 - F_{s,\tau_l}^Q(z_{s,\tau_l}) \right)^{(1-\mu_{ls})\gamma_l-1} f_{s,\tau_l}^Q(z_{s,\tau_l}) \\
& \quad \times B_{js} \left(F_{s,\tau_j}^Q(z_{s,\tau_j}) \right)^{\mu_{js}\gamma_j-1} \left(1 - F_{s,\tau_j}^Q(z_{s,\tau_j}) \right)^{(1-\mu_{js})\gamma_j-1} f_{s,\tau_j}^Q(z_{s,\tau_j}).
\end{aligned} \tag{27}$$

Conditionally on the remaining sites, the full conditional distribution of the z_{t,τ_j} is a function of the sites in the neighbourhood of j , i.e.

$$\begin{aligned}
& \pi(z_{t,\tau_j} | \mathbf{z}_{t-p:t-1,\tau_j}, \mathbf{z}_{t+1:t+p,\tau_j}, \mathbf{z}_{t-p:t,\tau_{j-1}}, \mathbf{z}_{t-p:t,\tau_{j+1}}, \boldsymbol{\theta}) \\
& \propto \prod_{s=t-p}^{t+p} \prod_{i \in N(j)} B_{is} \left(F_{s,\tau_i}^Q(z_{s,\tau_i}) \right)^{\mu_{is}\gamma_i-1} \left(1 - F_{s,\tau_i}^Q(z_{s,\tau_i}) \right)^{(1-\mu_{is})\gamma_i-1} f_{s,\tau_i}^Q(z_{s,\tau_i}) \\
& \quad \times B_{js} \left(F_{s,\tau_j}^Q(z_{s,\tau_j}) \right)^{\mu_{js}\gamma_j-1} \left(1 - F_{s,\tau_j}^Q(z_{s,\tau_j}) \right)^{(1-\mu_{js})\gamma_j-1} f_{s,\tau_j}^Q(z_{s,\tau_j}) \\
& \propto B_{jt} \left(F_{t,\tau_j}^Q(z_{t,\tau_j}) \right)^{\mu_{jt} \exp(\sigma_j)-1} \left(1 - F_{t,\tau_j}^Q(z_{t,\tau_j}) \right)^{(1-\mu_{jt}) \exp(\sigma_j)-1} f_{t,\tau_j}^Q(z_{t,\tau_j})
\end{aligned}$$

³The function `csaps` was used to perform cubic smoothing spline interpolation in Matlab.

$$\begin{aligned}
 & \times \prod_{k=1}^p B_{j,t+k} \left(F_{t+k,\tau_j}^Q(z_{t+k,\tau_j}) \right)^{\mu_{j,t+k} \exp(\sigma_j) - 1} \left(1 - F_{t+k,\tau_j}^Q(z_{t+k,\tau_j}) \right)^{(1-\mu_{j,t+k}) \exp(\sigma_j) - 1} \\
 & \times \prod_{i \in N(j)} B_{it} \left(F_{t,\tau_i}^Q(z_{t,\tau_i}) \right)^{\mu_{it} \exp(\sigma_i) - 1} \left(1 - F_{t,\tau_i}^Q(z_{t,\tau_i}) \right)^{(1-\mu_{it}) \exp(\sigma_i) - 1}.
 \end{aligned} \tag{28}$$

Similarly, it is easy to check that the full conditional distribution for the j th site, for $t = T - p + 1, \dots, T$ is

$$\begin{aligned}
 & \pi(z_{t,\tau_j} | \mathbf{z}_{t-p:t-1,\tau_j}, \mathbf{z}_{t+1:t-T,\tau_j}, \mathbf{z}_{t-p:t,\tau_{j-1}}, \mathbf{z}_{t-p:t,\tau_{j+1}}, \boldsymbol{\theta}) \\
 & \propto L(\mathbf{z}_{p+1:T} | \boldsymbol{\theta}) \\
 & \propto B_{jt} \left(F_{t,\tau_j}^Q(z_{t,\tau_j}) \right)^{\mu_{jt} \exp(\sigma_j) - 1} \left(1 - F_{t,\tau_j}^Q(z_{t,\tau_j}) \right)^{(1-\mu_{jt}) \exp(\sigma_j) - 1} f_{t,\tau_j}^Q(z_{t,\tau_j}) \\
 & \times \prod_{k=1}^{t-T} B_{j,t+k} \left(F_{t+k,\tau_j}^Q(z_{t+k,\tau_j}) \right)^{\mu_{j,t+k} \exp(\sigma_j) - 1} \left(1 - F_{t+k,\tau_j}^Q(z_{t+k,\tau_j}) \right)^{(1-\mu_{j,t+k}) \exp(\sigma_j) - 1} \\
 & \times \prod_{i \in N(j)} B_{it} \left(F_{t,\tau_i}^Q(z_{t,\tau_i}) \right)^{\mu_{it} \exp(\sigma_i) - 1} \left(1 - F_{t,\tau_i}^Q(z_{t,\tau_i}) \right)^{(1-\mu_{it}) \exp(\sigma_i) - 1}
 \end{aligned} \tag{29}$$

with $\prod_{k=1}^0 = 1$. The full conditionals are not trivial to simulate from, thus we apply an MH step with proposal distribution

$$q(z | \mathbf{z}_{t-p:t-1,\tau_j}, z_{t,\tau_{j-1}}, z_{t,\tau_{j+1}}, \boldsymbol{\theta}) \propto \left(F_{t,\tau_j}^Q(z) \right)^{\mu_{jt} \gamma_{jt} - 1} \left(1 - F_{t,\tau_j}^Q(z) \right)^{(1-\mu_{jt}) \gamma_{jt} - 1} f_{t,\tau_j}^Q(z)$$

which can be simulated exactly as follows: $y^* \sim \mathcal{B}e(\mu_{jt} \exp(\sigma_j), (1 - \mu_{jt}) \exp(\sigma_j))$ and $z^* = F_{t,\tau_j}^{Q,-1}(y^*)$. This choice of the proposal distribution leads to a simplification of the logarithmic acceptance probability. For the j th site, with $t = p + 1, \dots, T - p$ the acceptance probability is

$$\begin{aligned}
 \log \rho &= \sum_{i \in N(j)} (\log B_{it} - \log B_{it}^* + A_{it}(\mu_{it}^* - \mu_{it}) \exp(\sigma_i)) \\
 &+ \sum_{k=1}^p (\log B_{jt+k} - \log B_{jt+k}^{**} + A_{jt+k}(\mu_{jt+k}^{**} - \mu_{jt+k}) \exp(\sigma_j))
 \end{aligned}$$

with $B_{it}^* = B_{it}(\mu_{it}^* \exp(\sigma_i), (1 - \mu_{it}^*) \exp(\sigma_i))$ and

$$\mu_{it}^* = \varphi \left(\alpha_{0i} + \sum_{k=1}^p \alpha_{kj} y_{t-k,i} + \sum_{k \in N(i), k \neq j} \beta_{ki} y_{t,k} + \beta_{ji} y^* \right)$$

which follows by the definition of neighbourhood system, that is, if $i \in N(j)$ then $j \in N(i)$, and with $B_{jt+k}^{**} = B_{jt+k}(\mu_{jt+k}^{**} \exp(\sigma_j), (1 - \mu_{jt+k}^{**}) \exp(\sigma_j))$ and

$$\mu_{jt+k}^{**} = \varphi \left(\alpha_{0j} + \sum_{l=1, l \neq k}^p \alpha_{lj} y_{t+k-l,i} + \sum_{k \in N(j)} \beta_{ki} y_{t,k} + \alpha_{kj} y^* \right).$$

With regards to the MH step for the full conditional distributions of the j th site, for $t = T - p + 1, \dots, T$ the calculation of the acceptance probability is similar to the one given here above and is omitted for the sake of conciseness of the paper.

References

- Ait-Sahalia, Y. and Lo, A. (1998a). “Nonparametric estimation of state-price densities implicit in financial asset prices.” *Journal of Finance*, 53(2): 449–547. 791, 811
- Ait-Sahalia, Y. and Lo, A. (1998b). “Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices.” *Journal of Finance*, 53: 499–547. doi: <http://dx.doi.org/10.1111/0022-1082.215228>. 792
- Artzner, P., Delbaen, F., Eber, J., and Heath, D. (1999). “Coherent Measures of Risk.” *Mathematical Finance*, 9(3): 203–228. MR1850791. doi: <http://dx.doi.org/10.1111/1467-9965.00068>. 793
- Atchadé, Y. F., Lartillot, N., and Robert, C. (2013). “Bayesian computation for statistical models with intractable normalizing constants.” *Brazilian Journal of Probability and Statistics*, 27(4): 416–436. MR3105037. doi: <http://dx.doi.org/10.1214/11-BJPS174>. 799
- Berkowitz, J. (2001). “Testing density forecasts with applications to risk management.” *Journal of Business and Economic Statistics*, 19: 465–474. MR1947080. doi: <http://dx.doi.org/10.1198/07350010152596718>. 796, 807, 809, 810
- Bhar, R. and Chiarella, C. (2000). “Expectations of monetary policy in Australia implied by the probability distribution of interest rate derivatives.” *European Journal of Finance*, 6: 113–125. doi: <http://dx.doi.org/10.1080/13518470050020798>. 792
- Billio, M., Casarin, R., Ravazzolo, F., and Van Dijk, H. (2013). “Time-varying Combinations of Predictive Densities using Nonlinear Filtering.” *Journal of Econometrics*, 177: 213–232. MR3118557. doi: <http://dx.doi.org/10.1016/j.jeconom.2013.04.009>. 792, 793, 794
- Black, F. and Scholes, M. (1973). “The Pricing of Options and Corporate Liabilities.” *Journal of Political Economy*, 81: 637–654. doi: <http://dx.doi.org/10.1086/260062>. 802
- Bliss, R. and Panigirtzoglou, N. (2002). “Testing the stability of implied probability density functions.” *Journal of Banking and Finance*, 26(2): 381–422. doi: [http://dx.doi.org/10.1016/S0378-4266\(01\)00227-8](http://dx.doi.org/10.1016/S0378-4266(01)00227-8). 791, 792, 793, 811, 812
- Bliss, R. and Panigirtzoglou, N. (2004). “Option-implied risk aversion estimates.” *Journal of Finance*, 59: 407–446. doi: <http://dx.doi.org/10.1111/j.1540-6261.2004.00637.x>. 791, 796, 807, 812
- Bouguila, N., Ziou, D., and Monga, E. (2006). “Practical Bayesian estimation of a finite beta mixture through Gibbs sampling and its applications.” *Statistics*

- and Computing*, 16(2): 215–225. MR2227397. doi: <http://dx.doi.org/10.1007/s11222-006-8451-7>. 797
- Branscum, A. J., Johnson, W., and Thurmond, M. (2007). “Bayesian beta regression: Applications to household expenditure data and genetic distance between foot-and-mouth disease viruses.” *Australian & New Zealand Journal of Statistics*, 49: 287–301. MR2405396. doi: <http://dx.doi.org/10.1111/j.1467-842X.2007.00481.x>. 794
- Breeden, D. and Litzenberger, R. (1978). “Prices of state-contingent claims implicit in option prices.” *Journal of Banking and Finance*, 51(4): 621–651. 792
- Bremaud, P. (1999). *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*. Springer. MR1689633. doi: <http://dx.doi.org/10.1007/978-1-4757-3124-8>. 796
- Campa, J., Chang, K., and Reider, R. (1998). “Implied exchange rate distributions: evidence from OTC option markets.” *Journal of International Money and Finance*, 17(1): 117–160. doi: [http://dx.doi.org/10.1016/S0261-5606\(97\)00054-5](http://dx.doi.org/10.1016/S0261-5606(97)00054-5). 792, 811
- Campa, J., Chang, P., and Reider, R. (1997). “ERM bandwidths for EMU and after: evidence from foreign exchange options.” *Economic Policy*, 24: 55–89. 800
- Carlson, J. B., Craig, B., and Melick, W. R. (2005). “Recovering market expectations of FOMC rate changes with options on federal funds futures.” *Journal of Futures Markets*, 25: 1203–1242. doi: <http://dx.doi.org/10.1002/fut.20187>. 792
- Casarin, R., Valle, L. D., and Leisen, F. (2012). “Bayesian Model Selection for Beta Autoregressive Processes.” *Bayesian Analysis*, 7(1): 1–26. MR2934956. doi: <http://dx.doi.org/10.1214/12-BA713>. 794, 797, 810
- Chib, S. and Greenberg, E. (1995). “Hierarchical analysis of SUR models with extensions to correlated serial errors and time-varying parameter models.” *Journal of Econometrics*, 68: 339–360. doi: [http://dx.doi.org/10.1016/0304-4076\(94\)01653-H](http://dx.doi.org/10.1016/0304-4076(94)01653-H). 797
- Christoffersen, P. F. (1998). “Evaluating Interval Forecasts.” *International Economic Review*, 39: 841–862. MR1661906. doi: <http://dx.doi.org/10.2307/2527341>. 793
- Cressie, N. and Lele, S. (1992). “New Models for Markov Random Fields.” *Journal of Applied Probability*, 29(4): 877–884. MR1188543. doi: <http://dx.doi.org/10.2307/3214720>. 794
- Delbaen, F. and Schachermayer, W. (2011). *The Mathematics of Arbitrage*. Springer Finance. MR2200584. 795
- Diebold, F., Hahn, J., and Tay, A. (1999). “Multivariate density forecast evaluation and calibration in Financial risk management: high-frequency returns on foreign exchange.” *Review of Economics and Statistics*, 81: 661–673. doi: <http://dx.doi.org/10.1162/003465399558526>. 792, 793, 795
- Diebold, F. X., Tay, A., and Wallis, K. (1999). “Evaluating Density Forecasts of Inflation: The Survey of Professional Forecasters.” In: R. Engle and H. White (eds.), *Festschrift in Honor of C. W. J. Granger*, 76–90. Oxford: Oxford University Press, 1999. 792

- Diebold, F. X., Gunther, T., and Tay, A. (1998). “Evaluating Density Forecasts, with Applications to Financial Risk Management.” *International Economic Review*, 39: 863–883. doi: <http://dx.doi.org/10.2307/2527342>. 793
- Diebold, F. X. and Mariano, R. S. (1995). “Comparing Predictive Accuracy.” *Journal of Business and Economic Statistics*, 13: 253–263. 793
- Duffie, D. (2001). *Dynamic Asset Pricing*. Princeton University Press, Princeton and Oxford. 795
- Duffie, D., Pan, J., and Singleton, K. (2000). “Transform Analysis and Asset Pricing for Affine Jump-Diffusions.” *Econometrica*, 68(6): 1343–1376. MR1793362. doi: <http://dx.doi.org/10.1111/1468-0262.00164>. 793
- Durham, G. (2013). “Risk-neutral Modeling with Affine and Nonaffine Models.” *Journal of Financial Econometrics*, 11(4): 650–681. doi: <http://dx.doi.org/10.1093/jjfinec/nbt009>. 792, 793
- Engle, R. and Rosenberg, J. (2000). “Testing the volatility term structure using option hedging criteria.” *Journal of Derivatives*, 8: 10–29. doi: <http://dx.doi.org/10.3905/jod.2000.319113>. 811
- Fackler, P. L. and King, R. P. (1990). “Calibration of Option-Based Probability Assessments in Agricultural Commodity Markets.” *American Journal of Agricultural Economics*, 72(9): 73–83. doi: <http://dx.doi.org/10.2307/1243146>. 791, 793, 794, 795, 797, 811
- Fawcett, N., Kapetanios, G., Mitchell, J., and Price, S. (2013). “Generalised Density Forecast Combinations.” Technical report, Warwick Business School. 792, 793, 794
- Ferrari, S. and Cribari-Neto, F. (2004). “Beta regression for modelling rates and proportions.” *Journal of Applied Statistics*, 31(7): 799–815. MR2095753. doi: <http://dx.doi.org/10.1080/0266476042000214501>. 794
- Filipovic, D. (2009). *Term-Structure Models: A Graduate Course*. Springer Finance, 1st edition. MR2553163. doi: <http://dx.doi.org/10.1007/978-3-540-68015-4>. 792, 793
- Fusai, G. and Roncoroni, A. (2000). *Implementing Models in Quantitative Finance: Methods and Cases*. Springer. MR2386793. 806
- Geweke, J. (1992). “Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments.” In: *Bayesian Statistics*, 169–193. University Press. MR1380276. 803
- Geweke, J. and Amisano, G. (2011). “Optimal Prediction Pools.” *Journal of Econometrics*, 164: 130–141. MR2821798. doi: <http://dx.doi.org/10.1016/j.jeconom.2011.02.017>. 793
- Gneiting, T., Raftery, A., Westveld, A. H., and Goldman, T. (2005). “Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation.” *Monthly Weather Review*, 133: 1098–1118. doi: <http://dx.doi.org/10.1175/MWR2904.1>. 794

- Gneiting, T. and Ranjan, R. (2013). “Combining Predictive Distributions.” *Electronic Journal of Statistics*, 7: 1747–1782. MR3080409. doi: <http://dx.doi.org/10.1214/13-EJS823>. 792, 793, 794
- Goldstein, R. (2000). “The term structure of interest rates as a random field.” *Review of Financial Studies*, 13: 365–384. doi: <http://dx.doi.org/10.1093/rfs/13.2.365>. 794
- Grunwald, G., Guttorp, P., and Raftery, A. (1993). “Prediction rules for exponential family state-space models.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 55: 937–943. 794
- Gzyl, H. and Mayoral, S. (2008). “Determination of risk pricing measures from market prices of risk.” *Insurance: Mathematics and Economics*, 43(3): 437–443. MR2479591. doi: <http://dx.doi.org/10.1016/j.insmatheco.2008.08.008>. 793
- Hall, S. G. and Mitchell, J. (2007). “Combining density forecasts.” *International Journal of Forecasting*, 23: 1–13. doi: <http://dx.doi.org/10.1016/j.ijforecast.2006.08.001>. 792, 793
- Hull, J. C. (2014). *Options, Futures, and Other Derivatives*. Prentice Hall, 9th edition. 791, 795
- Hutchinson, J., Lo, A., and Poggio, T. (1994). “A nonparametric approach to pricing and hedging derivative securities via learning networks.” *Journal of Finance*, 49(3): 851–889. doi: <http://dx.doi.org/10.1111/j.1540-6261.1994.tb00081.x>. 811
- Jaworski, P., Durante, F., and Haerdle, W. (2010). *Copula Theory and Its Applications*. Springer. MR3075361. doi: <http://dx.doi.org/10.1007/978-3-642-12465-5>. 792, 793
- Kennedy, D. (1994). “The term structure of interest rates as a Gaussian random field.” *Mathematical Finance*, 4(3): 247–258. MR1286709. doi: <http://dx.doi.org/10.1111/j.1467-9965.1994.tb00094.x>. 794
- Kimmel, R. (2004). “Modeling the term structure of interest rates: a new approach.” *Journal of Financial Economics*, 72(1): 143–183. doi: [http://dx.doi.org/10.1016/S0304-405X\(03\)00204-6](http://dx.doi.org/10.1016/S0304-405X(03)00204-6). 794
- Lai, W.-N. (2011). “Comparison of methods to estimate option implied risk-neutral densities.” *Quantitative Finance*, 1–17. 791
- Liang, F. (2010). “A double Metropolis Hastings sampler for spatial models with intractable normalizing constants.” *Journal of Statistical Computation and Simulation*, 80(9): 1007–1022. MR2742519. doi: <http://dx.doi.org/10.1080/00949650902882162>. 799
- Litterman, R. (1986). “Forecasting With Bayesian Vector Autoregressions – Five Years of Experience.” *Journal of Business and Economic Statistics*, 4: 25–38. 796
- Malz, A. (1997). “Estimating the probability distribution of the future exchange rate from option prices.” *Journal of Derivatives*, 5(2): 18–36. doi: <http://dx.doi.org/10.3905/jod.1997.407988>. 792, 811

- McKenzie, E. (1985). “An autoregressive process for Beta random variables.” *Management Science*, 31: 988–997. doi: <http://dx.doi.org/10.1287/mnsc.31.8.988>. 794
- Merton, R. (1973). “Theory of rational option pricing.” *Bell Journal of Economics and Management Science*, 4: 141–183. MR0496534. doi: <http://dx.doi.org/10.2307/3003143>. 802
- Møller, J., Pettitt, A. N., Reeves, R., and Berthelsen, K. K. (2006). “An efficient Markov chain Monte Carlo method for distributions with intractable normalizing constants.” *Biometrika*, 93: 451–458. MR2278096. doi: <http://dx.doi.org/10.1093/biomet/93.2.451>. 799
- Murray, I., Ghahramani, Z., and MacKay, D. J. C. (2006). “MCMC for doubly intractable distributions.” In: *Proceedings of the 22nd annual conference on uncertainty in Artificial Intelligence*. 799
- Panigirtzoglou, N. and Skiadopoulos, G. (2004). “A New Approach to Modeling the Dynamics of Implied Distributions: Theory and Evidence from the SP500 Options.” *Journal of Banking and Finance*, 28(7): 1499–1520. doi: [http://dx.doi.org/10.1016/S0378-4266\(03\)00127-4](http://dx.doi.org/10.1016/S0378-4266(03)00127-4). 791, 805
- Robert, C. P. and Rousseau, J. (2002). “A Mixture Approach To Bayesian Goodness of Fit.” Technical Report 02009, CEREMADE, University Paris Dauphine. 797
- Rodriguez, A. and ter Horst, E. (2008). “Bayesian Dynamic Density Estimation.” *Bayesian Analysis*, 3(2): 339–366. MR2407430. doi: <http://dx.doi.org/10.1214/08-BA313>. 791
- Rosenthal, J. S. (2011). “Optimal Proposal Distributions and Adaptive MCMC.” In: S. P. Brooks, G. J., A. Gelman and Meng, X.-L. (eds.), *Handbook of Markov Chain Monte Carlo: Methods and Applications*, 93–112. Chapman & Hall. MR2858446. 803
- Schonbucher, P. (2003). *Credit Derivatives Pricing Models: Models, Pricing and Implementation*. Wiley, 1st edition. 792, 793
- Shimko, D. (1993). “Bounds of probability.” *Risk*, 6(4): 33–37. 792, 811
- Sihvonen, J. and Vähämaa, S. (2014). “Forward-looking monetary policy rules and option implied interest rate expectations.” *The Journal of Futures Markets*, 34(4): 346–373. doi: <http://dx.doi.org/10.1002/fut.21596>. 792, 800
- Sims, C. A. and Zha, T. (1998). “Bayesian Methods for Dynamic Multivariate Models.” *International Economic Review*, 39(4): 949–968. doi: <http://dx.doi.org/10.2307/2527347>. 796
- Skiadas, C. (2009). *Asset Pricing Theory*. Princeton Series in Finance. Princeton University Press. 795
- Soderlind, P. and Svensson, L. (1997). “New techniques to extract market expectations from financial instruments.” *Journal of Monetary Economics*, 40(2): 383–429. doi: [http://dx.doi.org/10.1016/S0304-3932\(97\)00047-0](http://dx.doi.org/10.1016/S0304-3932(97)00047-0). 792

- Vergote, O. and Puigvert, J. M. (2012). “Interest rate expectations and uncertainty during ECB governing council days: evidence from intraday implied densities of 3-month Euribor.” *Journal of Banking & Finance*, 36(10): 2804–2823. doi: <http://dx.doi.org/10.1016/j.jbankfin.2012.06.014>. 791, 792, 796, 800, 803, 805, 811
- Vesela, I. and Puigvert, J. (2014). “Interest rate forecasts, state price densities and risk premium from Euribor options.” *Journal of Banking & Finance*, 48: 210–223. doi: <http://dx.doi.org/10.1016/j.jbankfin.2014.03.028>. 791, 792, 796, 800, 811
- Vincent-Humphreys, R. D. and Noss, J. (2012). “Estimating probability distributions of future asset prices: empirical transformations from option-implied risk-neutral to real-world density functions.” Working Paper No. 455, Bank of England. 792
- Wallis, K. (1987). “Time series analysis of bounded economic variables.” *Journal of Time Series Analysis*, 8: 115–123. doi: <http://dx.doi.org/10.1111/j.1467-9892.1987.tb00425.x>. 794
- Wang, S. and Young, V. (1998). “Risk-Adjusted Credibility Premiums Using Distorted Probabilities.” *Scandinavian Actuarial Journal*, 2: 143–165. MR1659301. doi: <http://dx.doi.org/10.1080/03461238.1998.10413999>. 793
- Wang, S., Young, V., and Panjer, H. (1997). “Axiomatic Characterization of Insurance Prices.” *Insurance: Mathematics and Economics*, 21: 173–183. MR1604936. doi: [http://dx.doi.org/10.1016/S0167-6687\(97\)00031-0](http://dx.doi.org/10.1016/S0167-6687(97)00031-0). 793
- West, M. (1992). “Modelling agent forecast distributions.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 54: 553–567. <http://www.stat.duke.edu/~mw/MWextrapubs/West1992d.pdf> MR1160482. 792
- West, M. and Crosse, J. (1992). “Modelling of probabilistic agent opinion.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 54: 285–299. <http://www.stat.duke.edu/~mw/MWextrapubs/West1992c.pdf> MR1157726. 792
- West, M. and Harrison, P. J. (1997). *Bayesian Forecasting & Dynamic Models*. Springer Verlag, 2nd edition. MR1482232. 792

Acknowledgments

We thank seminar participants at the George Box Workshop 2014, and the Mathematics Department of Universidad Nacional de Medellín, Colombia. We also thank Laura Gunn for her contributions to improve this paper. Roberto Casarin’s research is supported by funding from the European Union, Seventh Framework Programme FP7/2007–2013 under grant agreement SYRTO-SSH-2012-320270, by the Institut Europlace de Finance, “Systemic Risk grant”, the Global Risk Institute in Financial Services, the Louis Bachelier Institute, “Systemic Risk Research Initiative”, and by the Italian Ministry of Education, University and Research (MIUR) PRIN 2010–11 grant MISURA. Fabrizio Leisen’s research is supported by the European Community’s Seventh Framework Programme FP7/2007–2013 under grant agreement number 630677.