SYRTO

SYstemic Risk TOmography
Signals, Measurements, Transmission Channels, and Policy Interventions

A Bayesian Beta Markov Random Field Calibration of the Term structure of implied risk neutral densities

Roberto Casarin, Fabrizio Leisen, German Molina, Enrique Ter Horst

SYRTO WORKING PAPER SERIES
Working paper n. 9 | 2015

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement n° 320270.

A Bayesian Beta Markov Random Field calibration of the term structure of implied risk neutral densities

Roberto Casarin[†] Fabrizio Leisen[‡]
German Molina[§] Enrique ter Horst[¶ ‖]

[†]University Ca' Foscari, Venice
[‡]University of Kent
[§]Idalion Capital US LP
[¶]CESA & IESA

Abstract

We build on Fackler and King (1990) and propose a 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. The model is a Bayesian dynamic beta Markov random field which allows for possible time dependence between densities with the same maturity and for dependence across maturities at the same point in time. The assumptions on the prior distribution allow us to compound the needs of model flexibility, parameter parsimony and information pooling across densities.

Keywords: Bayesian inference, Beta random fields, Exchange Metropolis Hastings, Markov chain Monte Carlo, Risk neutral measure.

1 Introduction

In financial mathematics, it is common to model stock prices as a geometric Brownian motion with mean drift equal to μ under the physical probability measure $\mathbb{P}$, and afterwards want to price options on such an asset. In order to

[‖]Corresponding author: enriqueterhorst@gmail.com. Author names are in alphabetical order.

do so one has to perform a change of measure to the asset process in order to make it risk neutral, meaning that it makes all investors neutral with respect to risk preferences. Such a probability measure is denoted as $\mathbb{Q}$ (Delbaen and Schachermayer, 2011). In general parametric stochastic process models, the mathematical problem of performing a change of measure from $\mathbb{P}$ to $\mathbb{Q}$ poses technical problems mainly due to the non-existeness of $\mathbb{Q}$ or its non-uniqueness (Delbaen and Schachermayer, 2011; Boyarchenko and Levendorskii, 2002). When departing from the regular geometric Brownian motion to the jump diffusion or geometric Lévy processes setup (Tankov and Cont, 2003), uniqueness of $\mathbb{Q}$ is not guaranteed and several methods such as the Esscher transform, among several, are used to circumvent these limitations (Esscher, 1932; Gerber and Shiu, 1994).

The economic literature has shown an increasing interest in nonparametric implied risk neutral densities (Fackler and King (1990); Lai (2011)) since not only do they allows us to gauge what are the economic agents thinking about the future and what are their economic expectation (Bliss and Panigirtzoglou, 2004; Rodriguez and ter Horst, 2008), but also the statistical superiority of nonparametric estimation of such risk-neutral densities (Lai, 2011). The Fackler and King (1990) calibration procedure of risk neutral densities, extracted from derivative prices, on the basis of observations on a variable of interest such as the underlying, allows us to obtain a density forecast for such a variable. Density forecast is now widely used in many applied economic contexts. The nonparametric calibration of implied risk neutral densities is now used in macroeconomics to generate prediction on inflation and interest rates (see Bhar and Chiarella (2000),Carlson et al. (2005), Vincent-Humphreys and Noss (2012), Vergote and Gutiérrez (2012); Vesela and Gutiérrez (2013), Sihvonen and Vähämaa (2014)).

Our contribution provides a dynamic estimation of the Radon-Nikodym derivative that allows us to move from a nonparametric estimation of $\mathbb{Q}$ to a nonparametric estimation of $\mathbb{P}$. This last result therefore 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 *Probability Integral Transforms*[1] (PIT) at different maturities and different dates for a given maturity. Since the PIT belong to the unit interval, the calibration function of our model makes use of beta densities as suggested by Fackler and King (1990). In order to

[1] A Probability Integral Transform is defined by a given realization of a random variable x_t and as $\text{PIT}_t = \int_{-\infty}^{x_t} f(y)dy$.

account for time and cross-maturity dependence, we propose a random field model with beta densities. 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 proper Bayesian inference framework, that allow us to include parameter 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.

Our paper is also related to the literature on density forecast. The use of densities for predicting quantities of interest is now common in economics and finance (see for a review) and many recent papers focused on the combination and the calibration of the predictive densities. Optimal linear pool of densities is considered e.g. in Hall and Mitchell (2007), Geweke and Amisano (2011), while more general approaches to density combination 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 which is solved in these papers. The issue of calibrating densities is considered instead in Gneiting et al. (2005) and Gneiting and Ranjan (2013), which also propose the use of beta densities to achieve a continuous deformation of the predictive density and to obtain well calibrated PITs. Our paper also contributes to this stream of literature, since it provides a general approach to the joint calibration of densities allowing for the pooling of information across different predictive densities (the risk neutral densities at different maturities). In despite of presence in the forecasting and financial literature of similar issues, such as the density calibration and combination, we shall notice that the implied risk neutral calibration literature differs substantially from the forecast calibration literature, in that the first one assumes the calibration model is generating the change of measure needed to obtain the physical measure from the risk neutral.

Finally, as an aside note, we would like think that this paper also contributes to the literature on modelling data on bounded domains. Our Bayesian beta Markov random field model, and the inference procedure, are original extensions to the multivariate context of the Bayesian beta models and inferences recently proposed in the statistic literature. See Branscum et al. (2007) for Bayesian beta regression and Casarin et al. (2012) for model selection in Bayesian beta autoregressive models and the references therein.

The paper is organized as follow. Section 2 introduces density calibration problem and our Bayesian beta random field model for the joint calibration. In Section 3, we discusses 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 some simulation experiments. In Section 5 we provide an application to the Euro currency while Section 6 concludes and discusses possible extensions.

2 A dynamic calibration model

Let x_{t,τ_i}, $i = 1, \ldots, M$, $t = 1, \ldots, T$, be a set of underlying realized forward levels, available at time t for the different maturities $\tau_1, \ldots, \tau_M$. Let $F^Q_{t,\tau_i}(x)$ and $F^P_{t,\tau_i}(x)$ denote the risk neutral and the physical cumulative density functions (cdf), respectively and $f^Q_{t,\tau_i}(x)$ and $f^P_{t,\tau_i}(x)$ their probability density functions (pdf).

We assume the following joint deformation model

$$F^P_t(x_{t,\tau_1}, \ldots, x_{t,\tau_M}) = C_t(F^Q_{t,\tau_1}(x_{t,\tau_1}), \ldots, F^Q_{t,\tau_M}(x_{t,\tau_M})) \tag{1}$$

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

$$f^P_{t,\tau}(x_{t,\tau_1}, \ldots, x_{t,\tau_M}) = c_t(F^Q_{t,\tau_1}(x_{t,\tau_1}), \ldots, F^Q_{t,\tau_M}(x_{t,\tau_M})) \prod_{j=1}^{M} f^Q_{t,\tau_j}(x_{t,\tau_j}) \tag{2}$$

where c_t is the mixed partial derivative of C_t with respect all the arguments. Let $y_{jt} = F^Q_{t,\tau_j}(x_{t,\tau_j})$, $j = 1, \ldots, M$, 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 $S = \{1, \ldots, M\}$ 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:

$$c_t(y_{1t}, \ldots, y_{Mt}) = \frac{1}{Z_t} \prod_{j=1}^{M} c_{jt}(y_{jt} | y_{N(j)}) \tag{3}$$

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 not know 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 lead to over-parametrized models and consequently to over-fitting problems. Thus, in

this paper we consider parsimonious beta MRF models. That is, 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 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 to depend on the past values of the PITs for the same maturity. We use the re-parametrization of beta pdf 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))

$$c_{jt}(y_{jt}|y_{N(j)}) = B_{jt} y_{jt}^{\mu_{jt}\gamma_{jt}-1}(1-y_{jt})^{(1-\mu_{jt})\gamma_{jt}-1} \tag{4}$$

with

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

and

$$\mu_{jt} = \varphi\left(\alpha_{0j} + \sum_{k=1}^{p} \alpha_{kj} y_{t-k,j} + \sum_{k \in N(j)} \beta_{kj} y_{t,k}\right) \tag{5}$$

$$\gamma_{jt} = \gamma_j \tag{6}$$

with $\varphi : \mathbb{R} \mapsto [0,1]$ a twice differentiable strictly monotonic link function. We assume a logistic function.

3 Bayesian inference

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

$$L(\mathbf{x}_{p+1:T}|\boldsymbol{\theta}) = \prod_{t=p+1}^{T} f^P_{t,\tau}(x_{t,\tau_1}, \ldots, x_{t,\tau_M}) \tag{7}$$

$$= \prod_{t=p+1}^{T} \frac{1}{Z_t} \prod_{j=1}^{M} B_{jt}(\mu_{jt}\gamma_j, (1-\mu_{jt})\gamma_j) \left(F^Q_{t,\tau_j}(x_{t,\tau_j})\right)^{\mu_{jt}\gamma_j - 1}$$

$$\left(1 - F^Q_{t,\tau_j}(x_{t,\tau_j})\right)^{(1-\mu_{jt})\gamma_j - 1} f^Q_{t,\tau_j}(x_{t,\tau_j})$$

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 Bayesian random field model we assume the following hierarchical specification of the prior distribution. For a given j, with $j = 1, \ldots, M$, we assume

$$\alpha_{kj} \overset{i.i.d.}{\sim} \mathcal{N}(\alpha_j, s_j^2)\, k = 0, \ldots, p \tag{8}$$

$$\beta_{kj} \overset{i.i.d.}{\sim} \mathcal{N}(\beta_j, g_j^2),\ k = 1, \ldots, m_j, \tag{9}$$

For the second level of the hierarchy we assume

$$\gamma_j \overset{i.i.d.}{\sim} \mathcal{G}a(\xi_1, \xi_2),\ j = 1, \ldots, M \tag{10}$$

$$\alpha_j \overset{i.i.d.}{\sim} \mathcal{N}(\alpha, s^2),\ j = 1, \ldots, M \tag{11}$$

$$\beta_j \overset{i.i.d.}{\sim} \mathcal{N}(\beta, g^2),\ j = 1, \ldots, M \tag{12}$$

where $m_j = \mathrm{Card}(N(j))$ is the number of elements of $N(j)$, $\mathcal{G}a(c, d)$ 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 consider the following re-parametrization $\sigma_j = \log(\gamma_j)$, $j = 1, \ldots, M$. We define the parameter vector $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_M, \alpha, \beta)$ where $\boldsymbol{\theta}_j = (\boldsymbol{\alpha}_j, \boldsymbol{\beta}_j, \sigma_j, \alpha_j, \beta_j)$, $\boldsymbol{\alpha}_j = (\alpha_{0j}, \alpha_{1j}, \ldots, \alpha_{pj})$ and $\boldsymbol{\beta}_j = (\beta_{1j}, \ldots, \beta_{m_j j})$. Then

the joint probability density function of the prior distribution is

$$f(\boldsymbol{\theta}) \propto \exp\left\{ -\frac{1}{2s_0^2}(\alpha - a)^2 - \frac{1}{2g_0^2}(\beta - b)^2 - \sum_{j=1}^{M}\left(\frac{1}{2s_j^2}(\alpha_j - \alpha)^2 \right.\right. \quad (13)$$
$$+\frac{1}{2g_j^2}(\beta_j - \beta)^2 + \frac{1}{2}(\boldsymbol{\alpha}_j - \boldsymbol{\mu}_j)' S_j^{-1}(\boldsymbol{\alpha}_j - \boldsymbol{\mu}_j)$$
$$\left.\left. +\frac{1}{2}(\boldsymbol{\beta}_j - \boldsymbol{\nu}_j)' G_j^{-1}(\boldsymbol{\beta}_j - \boldsymbol{\nu}_j) \right) \right\} \prod_{j=1}^{M} \exp\{-\xi/2\exp(\sigma_j)\} \exp(\xi/2\sigma_j)$$

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

The joint posterior distribution can be written as

$$\pi(\boldsymbol{\theta}|\mathbf{x}_{p+1:T}) \propto \exp\left(-\sum_{t=p+1}^{T} \log Z_t - \sum_{t=p+1}^{T}\sum_{j=1}^{M} \log B_{jt} \right. \quad (14)$$
$$\left. + \sum_{t=p+1}^{T}\sum_{j=1}^{M} \left(A_{jt}\mu_{jt} + (1 - F^Q_{t,\tau_j}(x_{t,\tau_j})) \right) \exp(\sigma_j) \right) f(\boldsymbol{\theta})$$

where

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

and

$$A_{jt} = \log(F^Q_{t,\tau_j}(x_{t,\tau_j})/(1 - F^Q_{t,\tau_j}(x_{t,\tau_j})))$$

A major problem with this model is that the normalizing constants Z_t, $t = p+1, \ldots, T$, 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 computational expensive for many statistical models. An exact simulation algorithm for our beta MRF model is not available, thus in this paper we follow and alternative route and apply the double MH algorithm proposed by Liang (2010). 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 a MH algorithm with transition kernel $K_{\boldsymbol{\theta}'}(\mathbf{z}|\mathbf{x})$, then the n-step transition probability is

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

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

$$\rho(\boldsymbol{\theta}, \boldsymbol{\theta}', \mathbf{z}_{p+1:T}|\mathbf{x}_{p+1:T}) = \frac{f(\boldsymbol{\theta})q(\boldsymbol{\theta}|\boldsymbol{\theta}', \mathbf{x}_{p+1:T})}{f(\boldsymbol{\theta}')q(\boldsymbol{\theta}'|\boldsymbol{\theta}, \mathbf{x}_{p+1:T})} \frac{L(\mathbf{z}_{p+1:T}|\boldsymbol{\theta})}{L(\mathbf{x}_{p+1:T}|\boldsymbol{\theta})} \frac{P^n_{\boldsymbol{\theta}'}(\mathbf{z}_{p+1:T}|\mathbf{x}_{p+1:T})}{P^n_{\boldsymbol{\theta}'}(\mathbf{x}_{p+1:T}|\mathbf{z}_{p+1:T})} \tag{15}$$

If we chose $q(\boldsymbol{\theta}|\boldsymbol{\theta}', \mathbf{x}_{p+1:T})$ as a Metropolis transition kernel then the exchange is a MH step with transition $P^n_{\boldsymbol{\theta}'}(\mathbf{z}_{p+1:T}|\mathbf{x}_{p+1:T})$ and target distribution $L(\mathbf{z}_{p+1:T}|\boldsymbol{\theta})$, and the acceptance probability in Eq. 15 becomes

$$\rho(\boldsymbol{\theta}, \boldsymbol{\theta}', \mathbf{z}_{p+1:T}|\mathbf{x}_{p+1:T}) = \frac{L(\mathbf{z}_{p+1:T}|\boldsymbol{\theta})}{L(\mathbf{x}_{p+1:T}|\boldsymbol{\theta})} \frac{L(\mathbf{x}_{p+1:T}|\boldsymbol{\theta}')}{L(\mathbf{z}_{p+1:T}|\boldsymbol{\theta}')} \tag{16}$$

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 a MH algorithm starting with $\boldsymbol{\theta}$.

2. Generate the auxiliary variable $\mathbf{z}_{p+1:T} \sim P^n_{\boldsymbol{\theta}'}(\mathbf{z}_{p+1:T}|\mathbf{x}_{p+1:T})$ and accept it with probability $\min\{1, \rho(\boldsymbol{\theta}, \boldsymbol{\theta}', \mathbf{z}_{p+1:T}|\mathbf{x}_{p+1:T})\}$ given in Eq. 16

3. Set $\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}'$ if the auxiliary variable is accepted and $\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}$ otherwise.

As regards to 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 a 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 iteratively from the full conditional distributions of each site. By using the Markov property of our dynamic random field with respect the chosen neighbourhood system, the full conditional distribution of the j-th site conditionally on the remaining sites is a function of the sites in the neighbourhood of j, i.e.

$$\begin{aligned}
&\pi(x_{t,\tau_j}|x_{t,\tau_{j-1}}, x_{t,\tau_{j+1}}, \boldsymbol{\theta}) \propto \\
&B_{jt}\left(F^Q_{t,\tau_j}(x_{t,\tau_j})\right)^{\mu_{jt}\exp(\sigma_j)-1}\left(1-F^Q_{t,\tau_j}(x_{t,\tau_j})\right)^{(1-\mu_{jt})\exp(\sigma_j)-1} f^Q_{t,\tau_j}(x_{t,\tau_j}) \\
&\prod_{k=1}^{p} B_{j,t+k}\left(F^Q_{t+k,\tau_j}(x_{t+k,\tau_j})\right)^{\mu_{j,t+k}\exp(\sigma_j)-1}\left(1-F^Q_{t+k,\tau_j}(x_{t+k,\tau_j})\right)^{(1-\mu_{j,t+k})\exp(\sigma_j)-1} f^Q_{t+k,\tau_j}(x_{t+k,\tau_j}) \\
&\prod_{i\in N(j)} B_{it}\left(F^Q_{t,\tau_i}(x_{t,\tau_i})\right)^{\mu_{it}\exp(\sigma_i)-1}\left(1-F^Q_{t,\tau_i}(x_{t,\tau_i})\right)^{(1-\mu_{it})\exp(\sigma_i)-1} f^Q_{t,\tau_i}(x_{t,\tau_i})
\end{aligned} \tag{17}$$

for $t = p+1, \ldots, T$ and $j = 1, \ldots, M$. The full conditionals are not easy to simulate from, thus we apply a MH step with proposal distribution

$$q(x|x_{t,\tau_{j-1}}, x_{t,\tau_{j+1}}, \boldsymbol{\theta}) \propto \left(F^Q_{t,\tau_j}(x)\right)^{\mu_{jt}\gamma_{jt}-1}\left(1-F^Q_{t,\tau_j}(x)\right)^{(1-\mu_{jt})\gamma_{jt}-1} f^Q_{t,\tau_j}(x)$$

which can be simulated exactly as follows: $y^* \sim \mathcal{B}e(\mu_{jt}\exp(\sigma_j), (1-\mu_{jt})\exp(\sigma_j))$ and $x^* = F^{Q,-1}_{t,\tau_j}(y^*)$. This choice of the proposal distribution leads to a simplification of the logarithmic acceptance probability:

$$\log\tilde{\rho} = \sum_{i\in N(j)} (\log B_{it} - \log B^*_{it} + A_{it}(\mu^*_{it} - \mu_{it})\exp(\sigma_i))$$

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 $j \in N(i)$.

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 other

more exotic derivatives but for central bankers as well and policy makers (Aït-Sahalia and Duarte, 2003; Rouah and Vainberg, 2007). Recently a great deal of interest has grown in predicting the both the nonparametric risk-neutral and its physical counterpart simultaneously for the 3-month Euribor interest rate using the beta calibration function as provided by Vesela and Gutiérrez (2013) 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 Gutiérrez (2012). These constant maturity risk-neutral densities are interpolated in practice from fixed expiration densities as done in Vergote and Gutiérrez (2012).

In this section we run 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. This exercise consists of several layers according to the following sequence:

- First we produce the data under the physical measure, which will be common to all the simulation exercises. We simulate price paths under the physical measure for 3, 6 and 12 months for a time interval of $T = 2$ years, $\mu = 0.20$, $r = 0.05$, $\sigma = 0.15$, $\tau_1 = 0.25$ (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 the parameters of this risk neutral measure. For this purpose, we assume two potential scenarios that cover the two extremes:

    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 $\sigma = 0.10$.
    3. In both cases, we are using r different than μ.

- For each of the cases above, and for each of the maturities in the simulation exercise, we compare two curves (Figure (1)):

    1. NC Curve: This is the non-calibrated curve. 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 calibrated data using the β-MRF process using the risk neutral data, under the stated value of the volatility.

- As a reference, the 45 degree line represents the perfect scenario where the PITs are not autocorrelated, and they are uniformly distributed.

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

$$S_t = S_0 + \int_0^t S_u \mu du + \int_0^t S_u \sigma dW(u) \tag{18}$$

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:

$$f^Q_{t,\tau_j}(S_{t+\tau_j}) = \frac{1}{S_{t+\tau_j}\sqrt{2\pi\sigma^2\tau_j}} \exp\left[-\frac{[\log(S_{t+\tau_j}/S_t) - (r - 0.5\sigma^2)\tau_j]}{2\sigma^2\tau_j}\right] \tag{19}$$

$j = 1, 2, 3$. Once we observe 3 months later a price level of $S_{t+\tau_1}$ under the historical measure, then we proceed to compute the 3, 6 and 12 months PITs at time t as follows:

$$y_{t,\tau_j} = \int_{-\infty}^{S_{t+\tau_j}} f^Q_{t,\tau_j}(S_{t+\tau_j}) dS_{t+\tau_j} \tag{20}$$

$j = 1, 2, 3$. The next day at time $t_1 = t + 1$, we recompute the PITs in the same way as equation (20), obtaining a vector $\mathbf{x}_t = (\mathbf{x}_t, \mathbf{x}_{t+1}, ..., \mathbf{x}_{t+T})$ where again $\mathbf{x}_s = (F^Q_{s,\tau_1}(x_{s,\tau_1}), F^Q_{s,\tau_2}(x_{s,\tau_2}), F^Q_{s,\tau_3}(x_{s,\tau_3}))$, and where the components of $\mathbf{x}_s$ will be very likely correlated. 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. Each period t we obtain a new PIT which is the outcome of the physical process under a 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 today 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 that is embedded in any overlapping data. Classical approaches include a mere thinning of the data (which we do in our simulation exercises) to take only non-overlapping periods. 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 more general, since it takes into account in the modelling the different sources of correlation 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.

We apply our Bayesian β-MRF calibration model with the following hyper parameter 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 MCMC algorithm we consider 5,000 iterations after converge (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 $\boldsymbol{\theta}$ from q has been setted in a way to achieve a 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).

With regards to Table (1):

- $\boldsymbol{\alpha}_j$ are the autoregressive parts of the MBRF (time factor) representing the time-dependence.

- $\boldsymbol{\beta}_j$ are the parameters linking the different maturities (maturity factor) Representing the cross-maturity dependence.

It seems that the autocorrelation over time decreases as the maturity increases. This can be seen in the value of the alphas. Also, β_1 and β_2 represent the neighboring maturities correlation parameters before and after respectively. So β_{12} represents the correlation parameters between maturity 1 and maturity 2, while β_{12} represents the correlation parameter between maturity 2 and maturity 1. Furthermore, it seems that the panels c and d are pooling across maturities. This is interesting, because it assumes the same autoregressive structure over time for the PITs across their maturities. **However, note that the values of gamma are extremely different we need to make sure we interpret correctly the meaning of this parameter lets discuss this bit.**

The results of the calibration exercises are given in Tab. 1 and Fig. 1. It is important to note the following salient features:

- The autoregressive coefficient is significant at all maturities. The proximity parameter is significant only for the last maturity. The value of the precision parameter increases with the maturity. Fig. 1 shows the non-calibrated and calibrated PITs. Fig. 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 $j = 1, \ldots, M$. The results are given in Tab. 1.

5 An application to the Euro currency

We apply our methodology to Over-the-Counter annualized implied volatilities on the Euro currency for different tenors (one month, two months, and six months), spanning from the 01/01/2010 until 01/04/2013. For the computation of the risk neutral densities for the Euro, 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 Gutiérrez (2012), in order to transform back to the option price space and take the second derivative to yield the risk neutral density[2]. For an

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

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

	τ_j, $j=1$		τ_j, $j=2$		τ_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) ($\sigma = 0.2$)

	τ_j, $j=1$		τ_j, $j=2$		τ_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)	14.03.88	(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% credibility 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 vale of the scale parameter is $\sigma = 0.15$.

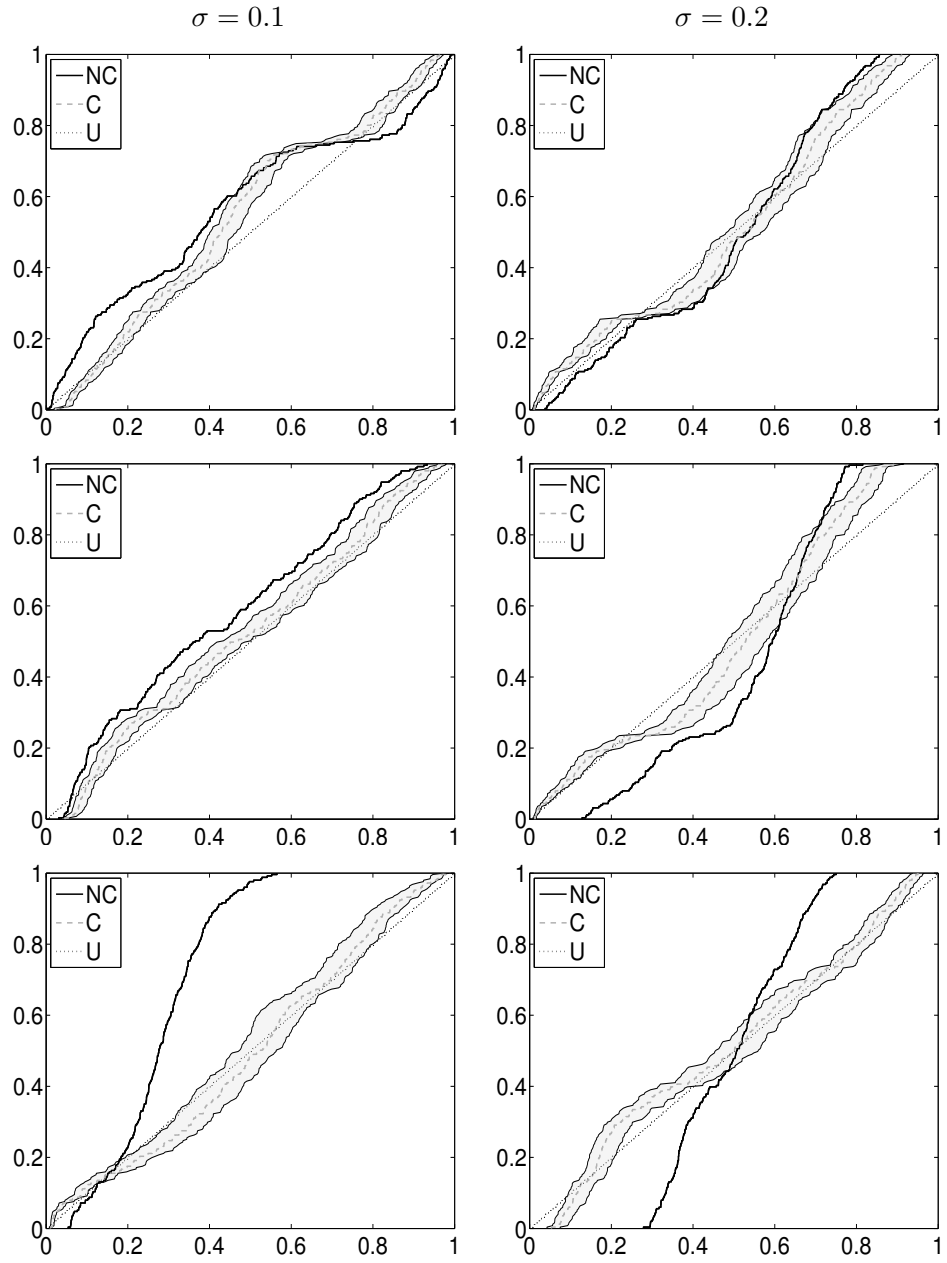


Figure 1: Non calibrated and calibrated risk neutral distribution for different maturities (rows) and volatility levels (columns).

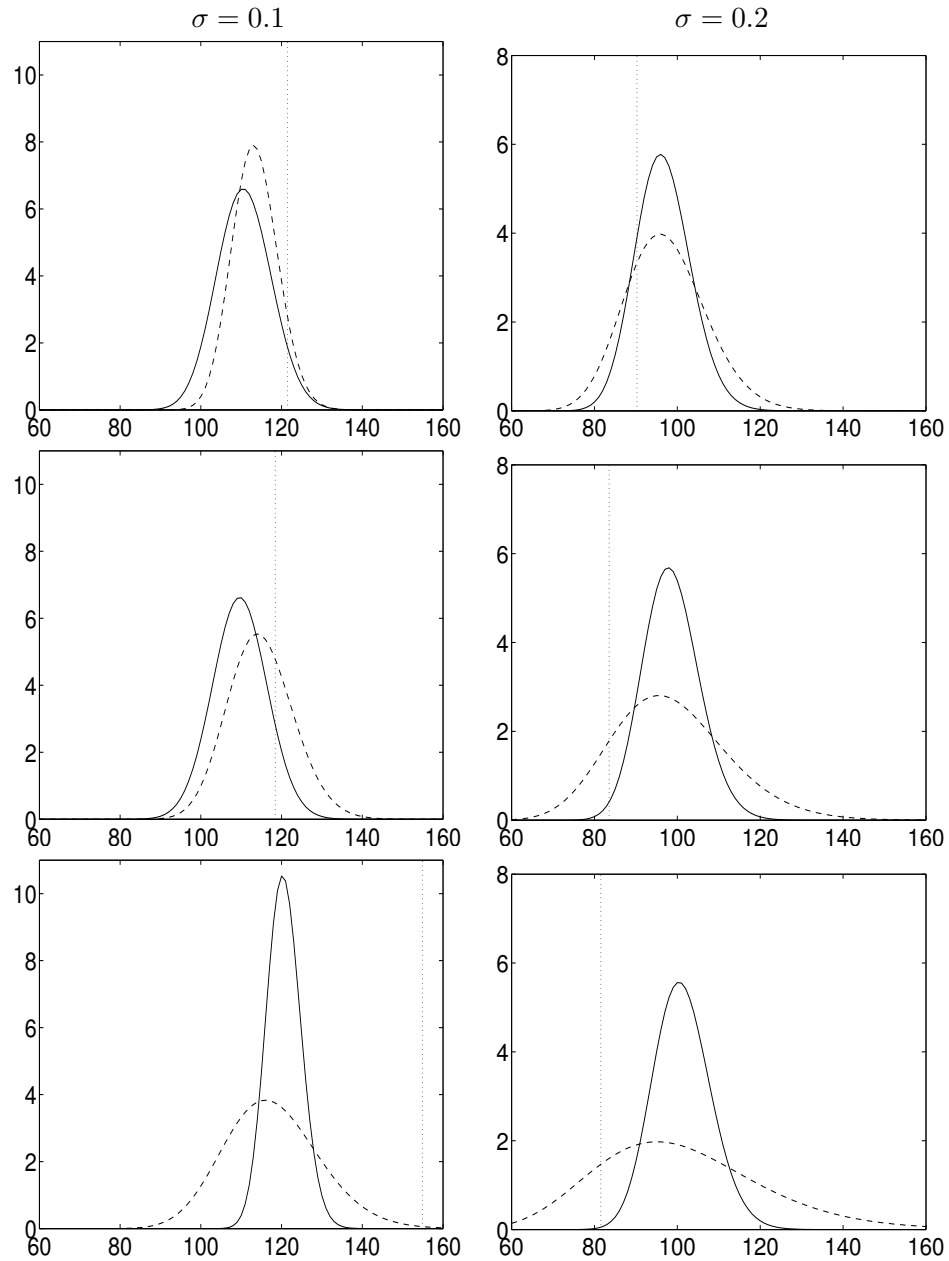


Figure 2: Non calibrated (dashed line) and calibrated (solid line) risk neutral distribution and price level (vertical dotted line) at last point of the sample, i.e. $t = 504$, for different maturities (rows) and different volatility levels (columns).

extensive review on how to extract risk-neutral densities from option prices with Matlab code included, please see Fusai and Roncoroni (2000). We can apply this methodology both for the case where we assume that each tenor has its own calibration function and also for the case where we assume that there is a single calibration function that works across several tenors by setting $\beta_{kj} = 0$ in the specification of μ_{jt}.

Figure 3 show the time series (left column) and the histograms (right column) of the different PIT series. Even though most histograms of the time series of the PITs look very uniform, an interesting feature is that the longer the tenor, the stronger is the autoregressive component.

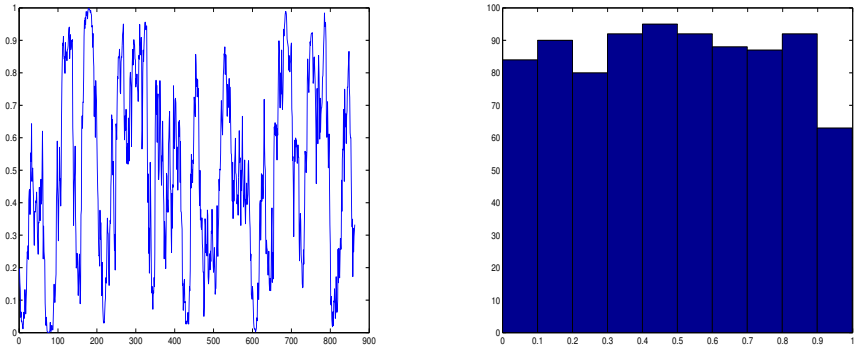
We further display below the risk neutral densities estimated on the 01/04/2013 for the different maturities as well as their physical densities computed by applying the calibration function to each of the risk-neutral densities. We apply our beta MRF calibration model with the prior and MCMC setting used in the simulated experiments (see previous section). **The results are given in Fig. 4. As it results from panel (a) in Tab. 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 one 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.**

6 Conclusion

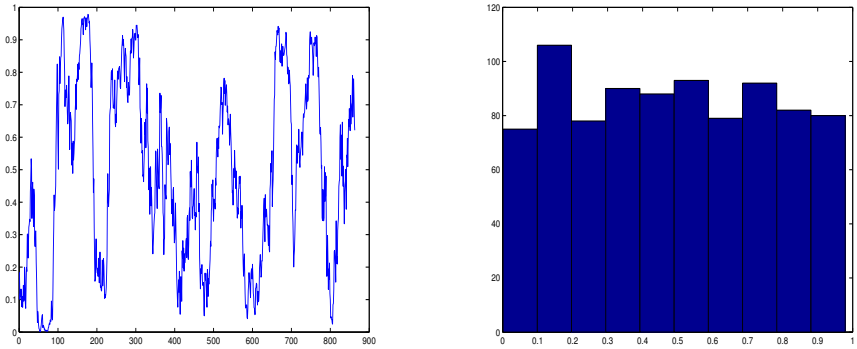
This paper provides an extension in the context of random fields of beta models existing in the literature and builds upon the Bayesian autoregressive model by Casarin et al. (2012). We have provided a new modelling framework using both the derivative and spot markets for the term structure of the implied probability, which accounts for the possible dependence between pits at different maturities and different dates for a given maturity, therefore allowing borrowing of information between the different tenors for both the risk-neutral and the physical measures. We also provide a proper inferential Bayesian framework that allows us to include parameter uncertainty in the density calibration functions, and therefore in the physical densities.

The literature on density forecast and the construction of calibration function has seen a renewed interest since Fackler and King (1990). The use of densities for predicting quantities of interest is now widely used

One month



Two months



Six months

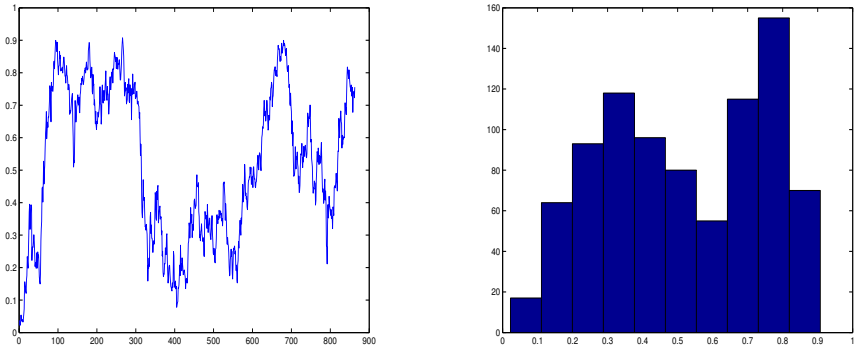


Figure 3: PIT time series (first column) and histogram (second column).

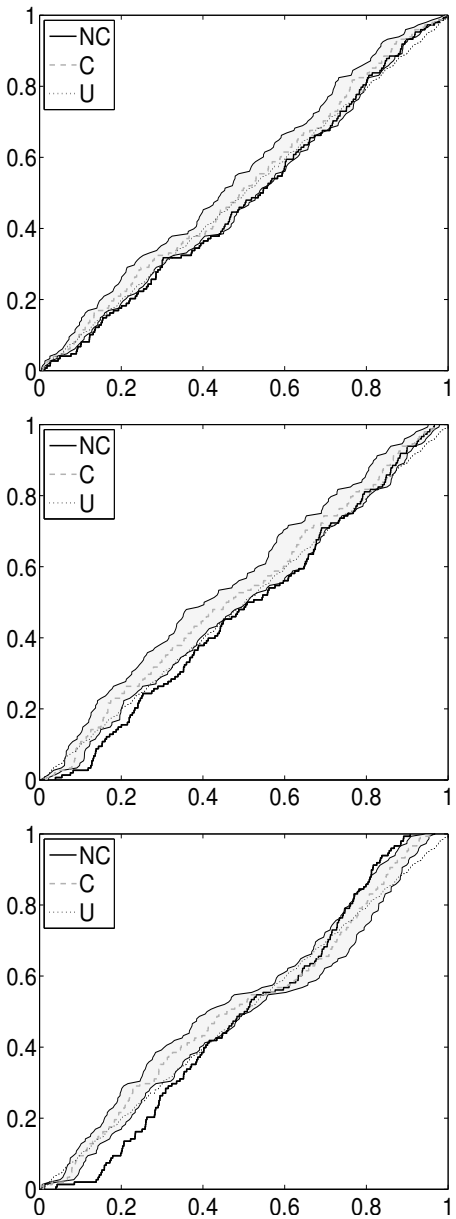


Figure 4: Non calibrated (solid line), perfectly calibrated (dotted line), and β-MRF calibrated (gray dashed line) risk neutral distributions for the three different maturities (rows): one, two and six months. In each plot, gray areas represent the 95% HPD region.

Panel (a) (Original data sample)

	τ_j, $j=1$		τ_j, $j=2$		τ_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)

	τ_j, $j=1$		τ_j, $j=2$		τ_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% credibility 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.

in many contexts (Vergote and Gutiérrez, 2012; Vesela and Gutiérrez, 2013) . Modelling the time evolution of the predictive densities and the relationship between densities from many sources is a challenging issue. For example, 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 Gutiérrez, 2012; Vesela and Gutiérrez, 2013). Using independent pits has the drawback of requiring a lot of data to have a reliable calibration function, allowing to use 4 data points per year for 3 month maturity densities, and 2 data points per year for 6 month maturity risk neutral densities. Our methodology can use data from every day in a rolling window fashion since it can adapt dependent pits.

Further research will include adapting our methodology to other indexes as the Euribor, stock, commodity and fixed income indexes by interpolating the risk neutral densities for different constant maturities from fixed expiry contracts as done in (Vergote and Gutiérrez, 2012).

Acknowledgments

We thank seminar participants at the George Box Workshop 2014, Mathematics Department of Universidad Nacional of Medellin, Colombia, 2014. 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, and by the Italian Ministry of Education, University and Research (MIUR) PRIN 2010-11 grant MISURA.

A Risk Neutral density estimation

Our real data consists of the daily implied annualized volatility on the Euro dollar for different expirations such as one week, one month, two months, three months and six months. The full data consists of the closing snapshots at the end of the London Business day for spot, forwards and implied volatilities for the period of the 1/1/2010 to the 23/4/2013. We follow Campa et al. (1997); Vergote and Gutiérrez (2012) and transform the option prices (y-axis) and strikes (x-axis) to the sigma (y-axis) and delta (x-axis) space in order to fit a cubic smoothing spline to the volatility smile. The reason for working in the sigma-delta space instead of the regular option price space is that undesired noise in the option data is introduced through high liquidity and transaction volumes which then makes difficult the interpolation of option prices. 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 option data by Shimko (1993); Hutchinson et al. (1994); Malz (1997); Ait-Sahalia and Lo (1998); Engle and Rosenberg (2000); Bliss and Panigirtzoglou (2002).

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

$$\min \lambda \sum_{i=1}^{n} \omega_i (\sigma_i - \sigma(\hat{\Theta})_i)^2 + (1-\lambda) \int_0^1 g''(\delta, \Theta) d\delta \tag{21}$$

where δ is the partial derivative of the Black and Scholes option call price with respect to the underlying, σ_i, $\sigma(\hat{\Theta})$, and $\omega_i = \frac{\nu_i}{\sum_i^n \nu_i}$ are the observed 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, $g()$ is the cubic spline function. The value for λ used is equal to 0.99. It is worthwhile noting that the BlackScholes formula is used solely to convert the option prices into/from their implied volatilities, in order to make the smoothing more effectively. This assumption does not assume that we are assuming the Black and Scholes pricing formula is the correct one, but only merely a way to make the smoothing more effective in the interpolation[3].

[3]The function csaps performs cubic smoothing spline interpolation in Matlab

References

Aït-Sahalia, Y. and Duarte, J. (2003). Nonparametric option pricing under shape restrictions. *Journal of Econometrics*, **116** 9–47.

Ait-Sahalia, Y. and Lo, A. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *Journal of Finance*, **53(2)** 449–547.

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.

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.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, **81** 637–654.

Bliss, R. and Panigirtzoglou, N. (2002). Testing the stability of implied probability density functions. *Journal of Banking and Finance*, **26(2)** 381–422.

Bliss, R. and Panigirtzoglou, N. (2004). Option-implied risk aversion estimates. *Journal of Finance*, **59** 407–446.

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** 215–225.

Boyarchenko, S. I. and Levendorskii, S. Z. (2002). *Non-Gaussian Merton-Black-Scholes theory*. Wold Scientific. Advanced Series on Statistical Science and Applied Probability-Vol 9.

Branscum, A. J., Johnson, W. and Thurmond, M. (2007). Bayesian beta regresssion: Applications to household expenditure data and genetic distance between foot-and-mouth disease viruses. *Australian & New Zealand Journal of Statistics*, **49** 287–301.

Bremaud, P. (1999). *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*. Springer.

Campa, J., Chang, P. and Reider, R. (1997). ERM bandwidths for EMU and after: evidence from foreign exchange options. *Economic Policy*, **24** 55–89.

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.

Casarin, R., Valle, L. D. and Leisen, F. (2012). Bayesian model selection for beta autoregressive processes. *Bayesian Analysis*, **7** 1–26.

Delbaen, F. and Schachermayer, W. (2011). *The Mathematics of Arbitrage*. Springer Finance.

Engle, R. and Rosenberg, J. (2000). Testing the volatility term structure using option hedging criteria. *Journal of Derivatives*, **8** 10–29.

Esscher, F. (1932). On the probability function in the collective theory of risk. *Skandinavisk Aktuarietidskrift*, **15** 175–195.

Fackler, P. L. and King, R. P. (1990). Calibration of option-based probability assesments in agricoltural commodity markets. *American Journal of Agricultural Economics*, **72** 73–83.

Fawcett, N., Kapetanios, G., Mitchell, J. and Price, S. (2013). Generalised density forecast combinations. Tech. rep., Warwick Business School.

Fusai, G. and Roncoroni, A. (2000). *Implementing Models in Quantitative Finance: Methods and Cases*. Springer.

Gerber, H. U. and Shiu, S. (1994). Option pricing by esscher transforms. *Transactions of the Society of Actuaries*, **46** 99–191.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In *In Bayesian Statistics*. University Press, 169–193.

Geweke, J. and Amisano, G. (2011). Optimal prediction pools. *Journal of Econometrics*, **164** 130–141.

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.

Gneiting, T. and Ranjan, R. (2013). Combining predictive distributions. *Electronic Journal of Statistics*, **7** 1747–1782.

Hall, S. G. and Mitchell, J. (2007). Combining density forecasts. *International Journal of Forecasting*, **23** 1–13.

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.

Lai, W.-N. (2011). Comparison of methods to estimate option implied risk-neutral densities. *Quantitative Finance* 1–17.

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.

Malz, A. (1997). Estimating the probability distribution of the future exchange rate from option prices. *Journal of Derivatives*, **5(2)** 18–36.

Merton, R. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science*, **4** 141–183.

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.

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*.

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*.

Robert, C. P. and Rousseau, J. (2002). A mixture approach to Bayesian goodness of fit. Technical Report 02009, CEREMADE, University Paris Dauphine.

Rodriguez, A. and ter Horst, E. (2008). Bayesian dynamic density estimation. *Bayesian Analysis*, **3** 339–366.

Rosenthal, J. S. (2011). Optimal Proposal Distributions and Adaptive MCMC. In *Handbook of Markov Chain Monte Carlo: Methods and*

Applications (G. J. S.P. Brooks, A. Gelman and X.-L. Meng, eds.). Chapman & Hall, 93–112.

Rouah, F. and Vainberg, G. (2007). Option pricing models and volatility using excel-vba. *John Wiley and Sons*.

Shimko, D. (1993). Bounds of probability. *Risk*, **6(4)** 33–37.

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.

Tankov, P. and Cont, R. (2003). *Financial Modelling with Jump Processes*. Chapman & Hall/CRC.

Vergote, O. and Gutiérrez, J. M. P. (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** 2804–2823.

Vesela, I. and Gutiérrez, J. P. (2013). Getting real forecasts, state price densities and risk premium from Euribor options. *Available at SSRN: http://ssrn.com/abstract=2178428*.

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.