Bayesian Non-Parametric Portfolio Decisions with Financial Time Series*

Audrone Virbickaite[†], M. Concepción Ausín[‡] and Pedro Galeano[§]

Abstract

A Bayesian non-parametric approach for efficient risk management is proposed. A dynamic model is considered where optimal portfolio weights and hedging ratios are adjusted at each period. The covariance matrix of the returns is described using an asymmetric MGARCH model. Restrictive parametric assumptions for the errors are avoided by relying on Bayesian non-parametric methods, which allow for a better evaluation of the uncertainty in financial decisions. Illustrative risk management problems using real data are solved. Significant differences in posterior distributions of the optimal weights and ratios are obtained arising from different assumptions for the errors in the time series model.

Keywords: Asymmetric Multivariate GARCH; Bayesian Non-parametrics; Dirichlet Process Mixtures; Hedging; Portfolio Allocation;

JEL Classification: C11, C32, C53, C58, G11

*The first and second authors are grateful for the financial support from MEC grant ECO2011-25706. The third author acknowledges financial support from MCI grant MTM2008-03010 and MEC grant ECO2012-38442.

[†]Universidad Carlos III de Madrid, c/ Madrid 126, Getafe (Madrid) 28903, Dpto. Estadística (UC3M); tel.: 916249674; email: audrone.virbickaite@uc3m.es.

[‡]UC3M; tel.: 916245852; fax: 916249849; email: concepcion.ausin@uc3m.es.

[§]UC3M; tel.: 916248901; fax: 916249849; email: pedro.galeano@uc3m.es.

1 Introduction

The first cases of managing risk can be traced as far as to the ancient world, see McNeil *et al.* (2005). Ever since Markowitz (1952) introduced the mean-variance approach, the area of financial risk management has advanced immensely. Portfolio optimization and hedging are just few out of many risk management problems, however, in this volatile world, they are as relevant as ever for today's investor. While some individuals seek personal gain, others try to ensure stability and reduce risk. Both of these goals can be achieved via efficient allocation and protection of their assets. For more on modern portfolio theory, hedging and risk management in general see Korn and Korn (2001), Elton *et al.* (2003), McNeil *et al.* (2005), Fries (2007), Kwok (2008) and Hull (2012), among others. These papers show that in order to determine optimal portfolio weights or hedging ratios, it is very important to find appropriate financial models which describe adequately the individual variability of the assets and their correlations or mutual dependence.

Traditionally, risk management problems have assumed a time-invariant relationship structure. However, ever since Engle (1982) showed the existence of time-varying variances, the standard approach of using constant unconditional correlations and covariances is being debated in the financial literature. The overwhelming empirical evidence shows the advantages of employing a time-varying approach, see Rossi and Zucca (2002), Yang and Allen (2004), Giamouridis and Vrontos (2007), Lien (2009), Liu *et al.* (2010), Lee and Lee (2012) and Basak and Chabakauri (2012), among others. ARCH-family models, first introduced by Engle (1982) and then generalized by Bollerslev (1986), without doubt, are the most researched and used in practice to explain time-varying volatilities, see also Bollerslev *et al.* (1992), Bollerslev *et al.* (1994), Engle (2002b), Teräsvirta (2009) and Tsay (2010). When dealing with multivariate time series, one must also take into consideration the mutual dependence between returns. Correlations can also exhibit some stylized features, such as persistence and asymmetry. For multivariate GARCH (MGARCH), see Bauwens *et al.* (2006), Silvennoinen and Teräsvirta (2009) and Tsay (2010). In this paper we will consider a very general multivariate GARCH model which accounts for both types of asymmetries: in individual volatilities and in conditional correlations.

Whichever GARCH-type model is chosen, the distribution of the returns depends on the distributional assumptions for the error term. It is well known, that every prediction, in order to be useful, has to come with a certain precision measurement. In this way the agent can know the uncertainty of the risk she is facing. Distributional assumptions permit to quantify this uncertainty about the future. However, the traditional premises of Normal or Student-t distributions may be rather restrictive. Alternatively, in this paper, we propose a Bayesian non-parametric approach avoiding the specification of a particular parametric distribution for the return innovations. More specifically, we consider a Dirichlet Process Mixture Model (DPM), firstly introduced by Antoniak (1974), with a Gaussian base distribution. This is a very flexible model that can be viewed as an infinite mixture of Gaussian distributions which includes, among others, the Gaussian, Student-t, logistic, double exponential, Cauchy and generalized hyperbolic distributions, among others.

The Bayesian approach also helps to deal with parameter uncertainty in portfolio decision problems, see e.g. Jorion (1986), Greyserman *et al.* (2006), Avramov and Zhou (2010) and Kang (2011), among others. This is in contrast with the usual maximum likelihood estimation approach, which assumes a "certainty equivalence" viewpoint, where the sample estimates are treated as the true values, which is not always correct and has been criticized in a number of papers. As noted by Jorion (1986), this estimation error can gravely distort optimal portfolio selection. In this paper, we propose a Bayesian method which provides the posterior distributions of the one-step-ahead optimal portfolio weights and hedging ratios, which are more informative than simple point estimates. In particular, using the proposed approach, it is possible to obtain Bayesian credible intervals for the optimal portfolio weights and hedging ratios. Also, as seen in Ardia and Hoogerheide (2010), the Bayesian inference provides some other advantages over the classical maximum likelihood techniques. For example, it is easy to incorporate via priors complicated positivity constraints on the parameters to ensure positive variance and covariance stationarity. Additionally, it is possible to approximate the posterior distribution of any other non-linear function of the parameters, as will be done for the optimal portfolio weights and hedging ratios. Moreover, the results are reliable even for finite samples. And finally, the models we wish to compare do not necessarily have to be nested.

Therefore, the main contribution of this work is the application of Bayesian non-parametric techniques in portfolio decision problems and exploration of the differences in uncertainty between the proposed approach and conventional restrictive distributional assumptions. Our objective is to provide a more realistic evaluation of risk of financial decisions. More specifically, in this paper we solve time-varying portfolio allocation and hedging problems using the multivariate GARCH specification of Cappiello *et al.* (2006), combined with ideas of Hafner and Franses (2009), and univariate GJR-GARCH model of Glosten *et al.* (1993) for the individual volatilities. For the errors, we assume a Bayesian non-parametric model based on the class of multivariate scale Gaussian mixtures where the scale mixing distribution follows a Dirichlet Process (DP) prior (Ferguson, 1973), leading to a DPM model.

The outline of the paper is as follows: Section 2 introduces the static and time-varying portfolio optimization and hedging approaches. Section 3 describes the model, inference and prediction from a Bayesian perspective. Section 4 presents a short simulation study. Section 5 illustrates the proposed approach using two real data examples. Finally, Section 6 concludes.

2 Portfolio Decisions

In this section, we first introduce Global Minimum Variance (GMV) portfolio and solve portfolio allocation problem by maximizing agent's utility. Then, in order to protect the portfolio from market risk, we solve a hedging with futures problem. The financial applications are presented using static and time-varying approaches and comparing different distributional assumptions in a Bayesian context.

2.1 Portfolio Allocation and Hedging: Static Approach

The main objective of diversification is to increase investor's utility and reduce her exposure to risk. See Markowitz (1952) and Merton (1972) for some classical portfolio optimization references. In this paper we consider the cases where the investor maximizes her expected utility and minimizes the portfolio variance. The GMV portfolio can be found at the very peak of the efficient frontier. For the utility, assume quadratic preferences. The following are both optimization problems:

$$\begin{aligned} p_U^* &= \arg\max_p \mathrm{E}\left[U(\gamma, r_t^P)\right] = \arg\max_p \mathrm{E}[r_t^P] - \frac{\gamma}{2}\mathrm{Var}[r_t^P] : \ p'1_K = 1, \\ p_{GMV}^* &= \arg\min_p \mathrm{Var}[r_t^P] : \ p'1_K = 1, \end{aligned}$$

where p is the weight vector, γ is the risk-aversion coefficient, representing a trade-off between the expected return and risk, 1_K is a K-vector of ones and r_t^P is a vector of portfolio returns. Portfolio is composed of assets, where $r_t^P = p'r_t$ and r_t is a $K \times 1$ vector of asset returns, such that $\mathrm{E}[r_t] = \mu$ and $\mathrm{Cov}[r_t] = \Sigma$. The following are the closed-form solutions for both portfolios:

$$\begin{aligned} p_U^* &= \frac{1}{\gamma}\left(\Sigma^{-1}\mu - \frac{\mu'\Sigma^{-1}1_K - \gamma}{1_K'\Sigma^{-1}1_K} \cdot \Sigma^{-1}1_K\right), \\ p_{GMV}^* &= \frac{\Sigma^{-1}1_K}{1_K'\Sigma^{-1}1_K}. \end{aligned}$$

However, if we choose to impose the short sale constraint, i.e., $p_i \geq 0, \forall i = 1, \ldots, K$, the problem cannot be solved analytically anymore and it requires numerical optimization techniques.

The optimal portfolio with (p^*, r_t^P) minimizes the risk arising from fluctuations in the level of returns of *individual assets*, i.e. specific risk, generated by spot price fluctuations. However, it is still susceptible to the changes in the level of the *stock market*. Therefore, the investor can protect herself by hedging her portfolio. Normally a hedge consists of a risky asset and an offset of another related security, mostly futures and options. A portfolio can be hedged using futures of a certain market index, that portfolio is mirroring (see application using real data in Section 5).

As described in Hull (2012), "(...) a futures contract is an agreement between two parties to buy or sell an asset at a certain time in the future for a certain price". The asset can be either a commodity, or a financial asset, such as currencies of stock indices for example. Hedgers constitute a big part of the participants in the futures markets, because futures is an effective tool to reduce risk. An agent, who makes a deal to sell an asset has a short futures position, therefore it is called short hedge, and vice-versa. See Hull (2012) for more on hedging strategies using futures.

The optimal proportion of the futures contract that counterbalances the spot position is called the optimal hedge ratio, D^*. Say that r_t^{HP} is the total return of the hedged portfolio at time t, r_t^F is the return of the future contract and r_t^P is the return of the portfolio or asset of interest we want to hedge. The total return r_t^{HP} is the difference between the portfolio return and the futures return, scaled by the hedging strategy D:

$$r_t^{HP} = r_t^P - D \times r_t^F.$$

In some cases, commodities for example, the conventional hedging assumes unit correlation between the underlying asset and the financial derivative (hedging instrument), where the optimal strategy results in the hedging ratio equal to one. Nonetheless, in practice this approach has certain limitations, especially in other than commodity markets, see DeCovny and Tacchi (1991). Therefore, the investors should employ such a hedging strategy that minimizes the total variance of the hedged portfolio:

$$D^*_{GMV} = \arg\min_D \mathrm{Var}\left[r_t^{HP}\right],$$

with solution:

$$D^*_{GMV} = \mathrm{Cov}\left[r_t^F, r_t^P\right] / \mathrm{Var}\left[r_t^F\right] = \rho_{F,P} \cdot \sigma_P / \sigma_F.$$

However, this approach relies on unconditional variance, and is criticized in a number of papers, because the dependence between the futures and the underlying asset is likely to be time-varying. Park and Bera (1987) notice that simple OLS approach ignores heteroscedasticity.

The minimum total variance hedging quality criterion is not the only one that can be considered in searching for the optimal hedge ratio. Another is a utility-based approach, which maximizes hedger's preferences, as seen in Delbaen *et al.* (2002), Becherer (2004), Rossi and Zucca (2002), among others. As before, assume a risk-averse investor, who has a quadratic utility function $\mathrm{E}\left[U(r_t^{HP}, \gamma)\right] = \mathrm{E}\left[r_t^{HP}\right] - \gamma/2 \cdot \mathrm{Var}\left[r_t^{HP}\right]$. Therefore, solve for the optimal hedge ratio:

$$D^*_U = \arg\max_D \mathrm{E}\left[U(r_t^{HP}, \gamma)\right] = \arg\max_D \mathrm{E}\left[r_{t+1}^{HP}\right] - \frac{\gamma}{2}\mathrm{Var}\left[r_{t+1}^{HP}\right],$$

with solution:

$$D^*_U = \frac{\gamma \cdot \mathrm{Cov}(r_t^P, r_t^F) - \mu^F}{\gamma \cdot \mathrm{Var}[r_t^F]}.$$

2.2 Portfolio Allocation and Hedging: Dynamic Approach

The use of the time-varying covariance matrix to determine portfolio weights and hedging ratios leads to better performing portfolios, as shown by Yilmaz (2011), and reduces the hedged portfolio risk, as seen in Choudhry (2004). Giamouridis and Vrontos (2007) find that portfolios, constructed under dynamic approach, have lower average risk and higher out-of-sample risk-adjusted realized return.

To solve the portfolio allocation problem in our case, instead of $\Sigma = \mathrm{Cov}[r_t]$ we use estimated one-step-ahead conditional covariance matrix for the assets returns $\mathrm{Cov}\left[r_{t+1} | \mathcal{I}_t\right] = H_{t+1}$, which is adjusted continuously on the basis of available information up to time t: $\mathcal{I}_t$. Therefore, we are able to obtain optimal portfolio weights for each period:

$$p^*_{U,t+1}|\mathcal{I}_t = \frac{1}{\gamma}\left(H^{-1}_{t+1}\mu - \frac{\mu' H^{-1}_{t+1}1_K - \gamma}{1'_K H^{-1}_{t+1}1_K}\cdot H^{-1}_{t+1}1_K\right), \tag{1}$$

$$p^*_{GMV,t+1}|\mathcal{I}_t = \frac{H^{-1}_{t+1}1_K}{1'_K H^{-1}_{t+1}1_K}, \tag{2}$$

where μ is a mean vector of the returns, which is assumed to be constant. Same goes for subsequent portfolio hedging problems. We can extend our analysis into estimating time-varying hedge ratio, that is adjusted at every period:

$$D^*_{U,t+1}|\mathcal{I}_t = \frac{\gamma\cdot\mathrm{Cov}[r^P_t, r^F_t|\mathcal{I}_t] - \mu^F|\mathcal{I}_t}{\gamma\cdot\mathrm{Var}[r^F_t|\mathcal{I}_t]} = \frac{\gamma\cdot H^{(1,2)}_{t+1} - \mu^F}{\gamma\cdot H^{(2,2)}_{t+1}}, \tag{3}$$

$$D^*_{GMV,t+1}|\mathcal{I}_t = \frac{\mathrm{Cov}\left[r^P_{t+1}, r^F_{t+1}|\mathcal{I}_t\right]}{\mathrm{Var}\left[r^F_{t+1}|\mathcal{I}_t\right]} = \frac{H^{(1,2)}_{t+1}}{H^{(2,2)}_{t+1}}, \tag{4}$$

where here H_{t+1} is the one-step-ahead covariance matrix between the portfolio we want to hedge and the financial instrument, futures in particular.

In all portfolio and hedging solutions, the H_{t+1} is one-step-ahead conditional covariance matrix of the returns, that is estimated using some multivariate volatility model. One possibility is the use of multivariate GARCH models, since they are easy to implement and can capture the stylized facts, that are characteristic to financial returns. The use of MGARCH models in optimal allocation context was first suggested by Cecchetti *et al.* (1988). Since then, there has been a number of papers investigating the differences in estimated hedging ratios and evaluating their performance using various approaches, from simple OLS, to bivariate vector autorregression (VAR), to GARCH. They show that the use of GARCH-type models leads to the overall portfolio risk reduction, see Rossi and Zucca (2002), Kroner and Sultan (1993) and Yang and Allen (2004), among others.

Therefore, given a multivariate time series vector of returns r_t:

$$r_t = \mu + a_t = \mu + H^{1/2}_t\epsilon_t,$$
$$H_t = \mathrm{Cov}[r_t|\mathcal{I}_{t-1}] = H^{1/2}_t\mathrm{Cov}[\epsilon_t](H^{1/2}_t)',$$

where a_t is mean-corrected returns, ϵ_t is a random vector, such that $\mathrm{E}[\epsilon_t] = 0$ and $\mathrm{Cov}[\epsilon_t] = I_K$. There is a wide range of MGARCH models, where most of them differ in specifying H_t. In the next section we describe the MGARCH model used in this paper.

3 Model, Inference and Prediction

This section describes the asymmetric multivariate GARCH model used for modeling volatilities, the implementation of Bayesian non-parametric inference and the methodology of obtaining predictive densities of the returns and volatilities.

3.1 Asymmetric Generalized DCC Model

Financial returns exhibit two types of asymmetries: in individual volatilities and in conditional correlations. In order to incorporate asymmetric volatility effect, for individual time series we choose the GJR-GARCH model by Glosten *et al.* (1993). To model joint volatilities we use Asymmetric Generalized DCC (AGDCC) model, proposed by Cappiello *et al.* (2006) (based on the previous work by Engle (2002a)). We also incorporate the ideas of Hafner and Franses (2009), where the parameters in the correlation equation are vectors, not scalars, thus allows for asset-specific dynamics. This leads to the following final model:

$$r_t = \mu + H_t^{1/2}\epsilon_t, \text{ where } r_t - \mu = a_t \text{ and } \epsilon_t \sim \mathcal{F}_K, \tag{5}$$
$$H_t = D_t R_t D_t, \tag{6}$$
$$D_t^2 = \text{diag}(\omega_i) + \left[\text{diag}(\alpha_i) + \text{diag}(\text{diag}(\phi_i)I'_{t-1})\right] \odot a'_{t-1}a_{t-1} + \text{diag}(\beta_i) \odot D_{t-1}^2, \tag{7}$$
$$\varepsilon_t = D_t^{-1}a_t, \;\; \eta_t = \varepsilon_t \odot I(\varepsilon_t < 0), \tag{8}$$
$$Q_t = S(1 - \bar{\kappa}^2 - \bar{\lambda}^2 - \bar{\delta}^2/2) + \kappa\kappa' \odot \varepsilon'_{t-1}\varepsilon_{t-1} + \lambda\lambda' \odot Q_{t-1} + \delta\delta' \odot \eta'_{t-1}\eta_{t-1}, \tag{9}$$
$$R_t = (\text{diag}(Q_t))^{-1/2} Q_t (\text{diag}(Q_t))^{-1/2}, \tag{10}$$

where $\mathcal{F}_K$ is a K-dimensional distribution, specified later, "diag" stands for either taking just the diagonal elements from the matrix, or making a diagonal matrix from a vector, S is a sample correlation matrix of ε_t, $\odot$ denotes Hadamard matrix product operator. Parameters κ, λ and δ are $K \times 1$ vectors, $\bar{\kappa} = K^{-1}\sum_{i=1}^{K}\kappa_i$, $\bar{\lambda} = K^{-1}\sum_{i=1}^{K}\lambda_i$ and $\bar{\delta} = K^{-1}\sum_{i=1}^{K}\delta_i$. To ensure positivity and stationarity of Q_t, we impose $\kappa_i, \lambda_i, \delta_i > 0$ and $\kappa_i^2 + \lambda_i^2 + \delta_i^2/2 < 1$, $\forall i = 1, \ldots, K$. Individual volatilities are represented in the equation (7): $d_{iit}^2 = \omega_i + (\alpha_i + \phi_i I_{t-1})a_{it-1}^2 + \beta d_{iit-1}^2$, where I is an indicator function $(a_{it} < 0)$ and d_{iit}^2 are individual asset volatilities, following a GJR-GARCH model with parameters $\omega_i \geq 0$, $0 < \alpha_i, \phi_i, \beta_i < 1$, such that $\alpha_i + \beta_i + \phi_i/2 < 1$, $\forall i = 1, \ldots, K$.

The AGDCC by Cappiello *et al.* (2006) is just a special case where $\kappa_1 = \ldots = \kappa_K$, $\lambda_1 = \ldots = \lambda_K$ and $\delta_1 = \ldots = \delta_K$:

$$Q_t = S(1 - \kappa - \lambda - \delta/2) + \kappa \times \varepsilon'_{t-1}\varepsilon_{t-1} + \lambda \times Q_{t-1} + \delta \times \eta'_{t-1}\eta_{t-1}. \tag{11}$$

As for the distribution of $\epsilon_t \sim \mathcal{F}_K$, we model it as an infinite scale mixture of Normal distributions, where the density of ϵ_t is as follows:

$$f_\epsilon(\epsilon_t|G) = \int \mathcal{N}_K(\epsilon_t|0, \Lambda_t)dG(\Lambda_t^{-1}),$$

where $\mathcal{N}_K(\epsilon_t|0, \Lambda_t^{-1})$ denotes a K-dimensional conditional density function of multivariate Normal distribution with mean zero and scale matrix Λ_t, and G is the scale mixing distribution, which is unknown and modeled by a Dirichlet Process.

DP is a multi-parameter generalization of the Beta distribution and defines a distribution over distributions. DP leads to discrete probability measures, which is a disadvantage in practice. This problem can be overcome by using Dirichlet Process

Mixture model, as seen in Antoniak (1974):

$$\begin{aligned} \epsilon_t|\Lambda_t &\sim \mathcal{N}_K(0, \Lambda_t^{-1}), \\ \Lambda_t|G &\overset{\text{iid}}{\sim} G, \\ G|c, G_0 &\sim DP(c, G_0), \\ c &\sim \pi(c), \end{aligned} \tag{12}$$

where $c > 0$ is a scale parameter with prior density π, and G_0 is a prior probability measure. Observe that G is a random distribution drawn from the DP and because it is discrete, multiple Λ_i's can take the same value simultaneously, making it a mixture model. Hence, using the stick-breaking representation, the hierarchical model in (12) can be seen as an infinite mixture of distributions:

$$f(\epsilon_t|\Lambda, \rho) = \sum_{i=1}^{\infty} \rho_i \mathcal{N}_K(\epsilon_t|0, \Lambda_i^{-1}), \tag{13}$$

where the weights are obtained as follows: $\rho_1 = v_1$, $\rho_i = (1-v_1)\ldots(1-v_{i-1})v_i$, where v_i is Beta distributed: $v_i \sim \mathcal{B}\,(1, c)$ for $i = 2, 3, \ldots$. We assume a conjugate model, where G_0 is a Gamma distribution with parameters $(a/2, b/2)$ and also assume a Gamma hyper-prior on the concentration parameter $c \sim \mathcal{G}(a_0, b_0)$. Finally, we have assumed uniform prior distributions for the parameters of the GJR-AGDCC model.

3.2 Bayesian Inference

A number of papers in the field of GARCH-type models have explored different Bayesian procedures for inference and prediction and different approaches to modeling the fat-tailed errors and/or asymmetric volatility. The recent development of modern Bayesian computational methods, based on Monte Carlo approximations, such as importance sampling, and MCMC methods, such as the Metropolis-Hastings algorithms, have facilitated the usage of Bayesian techniques, see Robert and Casella (2004).

The following section describes the Bayesian non-parametric procedure used for the GJR-AGDCC model in (5)-(10). The algorithm is based on works by Walker (2007), Papaspiliopoulos and Roberts (2008), Papaspiliopoulos (2008) and Ausín *et al.* (2011).

Regarding the inference algorithms, there are two main types of approaches. On the one hand, the marginal methods, which rely on the Polya urn representation. All these algorithms are based on integrating out the infinite dimensional part of the model. One of the most recent papers, based on this method in MGARCH setting is by Jensen and Maheu (2012). Recently, another class of algorithms, called conditional methods, have been proposed. These approaches, based on the stick-breaking scheme, leave the infinite part in the model and sample a finite number of variables. These include the procedure by Walker (2007), who introduces slice sampling schemes to deal with the infiniteness in DPM, and the retrospective MCMC method of Papaspiliopoulos and Roberts (2008), that is later combined by Papaspiliopoulos (2008) with a slice sampling method by Walker (2007) to obtain a new composite algorithm, which is better, faster and easier to implement. Generally, the stick-breaking procedure, compared to the Polya urn, produces better mixing and simpler algorithms.

As seen in Walker (2007), in order not to sample an infinite number of values at each MCMC step, we introduce a latent variable u_t, such that the joint density of (ϵ, u) given (ρ, Λ) is given by

$$f(\epsilon_t, u_t|\rho, \Lambda) = \sum_{i=1}^{\infty} \mathbf{1}(u_t < \rho_i)\mathcal{N}_K(\epsilon_t|0, \Lambda_i^{-1}). \tag{14}$$

Let $A_\rho(u_t) = \{i : \rho_i > u_t\}$ be a set of size N_{u_t}, which is finite for all $u_t > 0$. Then the joint density of (ϵ_t, u_t) in (14) can be equivalently written as $f(\epsilon_t, u_t|\rho, \Lambda) = \sum_{i\in A_\rho(u_t)} \mathcal{N}_K(\epsilon_t|0, \Lambda_i^{-1})$. Integrating over u_t gives us the previous density of infinite mixture of distributions (13). Finally, given u_t, the number of mixture components is finite. In oder to simplify the likelihood, we also need to introduce further indicator latent variable z_t, which indicates the mixture component that ϵ_t comes from: $f(\epsilon_t, z_t = j, u_t|\rho, \Lambda) = \mathcal{N}_K(\epsilon_t|0, \Lambda^{-1})\mathbf{1}(j \in A_\rho(u_t))$. Define parameter sets $\Omega = (\rho, \Lambda)$ and $\Phi = (\mu, \omega, \alpha, \beta, \phi, \kappa, \lambda, \delta)$, where $\Theta = (\Omega, \Phi)$ is the set of all model parameters. Then, the log likelihood of Θ, given the latent variables u_t and z_t looks as follows:

$$l(\Theta|u_t, z_t) = -\frac{1}{2}\sum_{t=1}^{T}(k\log(2\pi) + \log|H_t^*| + a_t H_t^{*-1} a_t'), \tag{15}$$

where H_t^* is the new conditional covariance matrix, *adjusted* by the variance of the errors:

$$\mathrm{Cov}\left[r_t|\mathcal{I}_{t-1}, z_t\right] = H_t^{1/2}\Lambda_{z_t}^{-1}H_t^{1/2} = H_t^*.$$

Next, we describe the DPM model algorithm step by step.

Firstly, given z, the conditional posterior distribution of concentration parameter c is independent of the rest of the parameters, as in Escobar and West (1995). So, we first sample an auxiliary variable $\xi \sim \mathcal{B}(c+1, T)$ and then c from a Gamma mixture:

$$\pi_\xi \mathcal{G}(a_0 + z^*, b_0 - \log(\xi)) + (1-\pi_\xi)\mathcal{G}(a_0 + z^* - 1, b_0 - \log(\xi)),$$

where $z^* = \max(z_1, \ldots, z_T)$ and $\pi_\xi = (a_0 + z^* - 1)/(a_0 + z^* - 1 + T(b_0 - \log(\xi)))$.

In the **second step** sample the weights of the components v_j for $j = 1, \ldots, z^*$, where the prior for $v \sim \mathcal{B}(1, c)$ and, given the data and z:

$$v_j|z \sim \mathcal{B}(n_j + 1, T - \sum_{l=1}^{j} n_l + c),$$

where n_j is the number of observations in the j^{th} component and $\sum_{l=1}^{j} n_l$ gives the cumulative sum of the groups. Also, $\rho_1 = v_1$, $\rho_j = (1-v_1)\ldots(1-v_{j-1})v_j$, for

$j = 2, \ldots, z^*$.

At the **third step** update $u_t \sim \mathcal{U}(0, \rho_{z_t})$, for $t = 1, \ldots, T$.

In the **fourth step** sample all the values of ρ_j that are larger than u_t. As Walker (2007) showed we need to find the smallest j^* such that $\sum_{j=1}^{j^*} \rho_j > u^*$ and then update v_j and ρ_j for $j = z^* + 1, \ldots, j^*$, where $u^* = \min(u_1, \ldots, u_T)$.

In multivariate setting the conjugate prior for the inverse of covariance matrix is the Wishart distribution, which can be seen as a matrix generalization of the chi-square distribution (Eaton, 2007). Therefore, the prior for the scale matrix Λ is chosen to be Wishart (V, df, K), where V is positive semi-definite symmetric scale matrix of dimensions $K \times K$, df is degrees of freedom and K is the dimension parameter. The real degrees of freedom can be obtained as $df + K - 1$. The mean vector of the variables distributed as Wishart is $\mathrm{E}[\Lambda] = (df + K - 1) \times V$. Sampling variability is large when the degrees of freedom is small. Thus, df has to be larger than $K-1$, and V such that the expectation of the Λ_t is identity: $V = \frac{1}{(df+K-1)} I_K$, as seen in Jensen and Maheu (2012). See more on the properties of Wishart distribution in Eaton (2007). Next, update the Λ, whose posterior distribution is independent of (ρ, u_t):

$$\Lambda_j | \ldots \sim \mathcal{W}\left(df + n_j, \left(V^{-1} + \mathrm{Cov}\left[\epsilon_{z_t=j}\right]\right)^{-1}\right).$$

In the **sixth step** update to which component the observations belong to by using the following (as seen in Walker (2007)):

$$\text{Probability}(z_t = j | \ldots) \propto \mathbf{1}\left(j \in A_\rho\left(u_t\right)\right) \mathcal{N}_K(\epsilon_t | 0, \Lambda_j^{-1}),$$

where $A_\rho(u_t) = \{j : \rho_j > u_t\}$, which is not empty.

The rest of the steps of the algorithm concern updating the parameters of the GJR-AGDCC model. We use the Random Walk Metropolis Hasting (RWMH), where for each parameter $\theta \in \Theta$ a candidate value $\tilde{\theta}$ is generated from a K-variate Normal distribution with mean equal to the previous value of the parameter and variance calibrated to achieve a desired acceptance probability. This procedure is repeated at each MCMC iteration. The probability of accepting a proposed value $\tilde{\theta}$, given the current value θ, is $\alpha(\theta, \tilde{\theta}) = \min\left\{1, \prod_{t=1}^{T} l(r_t|\theta) / \prod_{t=1}^{T} \tilde{l}(r_t|\tilde{\theta})\right\}$, where the likelihood used is as in (15), see e.g. Robert and Casella (2004). In this paper the acceptance probabilities are adjusted to be between 20% and 50%. For more detailed explanation of the algorithm in univariate setting see Ausín *et al.* (2011).

3.3 Prediction

In this section, we are mainly interested in estimating the one-step-ahead predictive density of the returns:

$$f(r_{t+1}|r_1, \ldots, r_t) = \int f(r_{t+1}|\Phi) f(\Theta|r_1, \ldots, r_t) d\Theta. \tag{16}$$

Although this integral is not analytically tractable, we can approximate it using the MCMC output. For this, we make use of the the procedure described in Walker

(2007), where it is explained how to obtain a sample from the predictive density of the errors $f(\epsilon_{t+1}|\epsilon_1,\ldots,\epsilon_t)$. Thus, having this sample and a posterior sample of the remaining model parameters Φ, it is easy to obtain a sample from the predictive (16) using equations (5)-(10). The following is the detailed explanation of the procedure.

At each MCMC iteration there are weights ρ_j and corresponding precision matrices Λ_j. We sample a random variable $r \sim \mathcal{U}(0,1)$ and take such Λ_j for which $\rho_{j-1} < r < \rho_j$. If we need more weights (i.e. $\sum \rho_j < 1$), we can sample additional ρ_j as before[1], and Λ_j from the prior. Once we have Λ_j selected, Walker (2007) proposes to draw one observation from each $\mathcal{N}\left(0,\Lambda_j^{-1}\right)$ in order to have a sample $\{\epsilon_{t+1,m}\}_{m=1}^{M}$ of the size of the MCMC chain from the predictive distribution of $f(\epsilon_{t+1}|\epsilon_1,\ldots,\epsilon_t)$. Alternatively, instead of sampling just one observation, we suggest to assemble the entire collection of precision matrices, $\left\{\Lambda_{z_{t+1},m}\right\}_{m=1}^{M}$. This approach has several advantages. Firstly, it allows to incorporate prior information about the variance of the errors by having the small probability of being sampled from the prior. Secondly, it allows to increase the sample size for the predictive density of $f(\epsilon_{t+1}|\epsilon_1,\ldots,\epsilon_t)$, because instead of sampling just one observation as in Walker (2007), we can sample as many as we choose at each MCMC iteration: $\epsilon_i \overset{\text{iid}}{\sim} \mathcal{N}(0,\Lambda^{-1})$, $i=1,\ldots$. And finally, it provides a sample of one-step-ahead volatilities $\left\{H^*_{t+1,m}\right\}_{m=1}^{M}$, that we will use in the portfolio allocation and hedging problems in the following section.

3.4 Bayesian Portfolio Decisions

As commented in the introduction, optimal allocation is greatly affected by the parameter uncertainty, which has been recognized in a number of papers, see Jorion (1986) and Greyserman *et al.* (2006), among others. They conclude that in frequentist setting the estimated parameter values are considered to be the true ones, therefore, the optimal portfolio weights tend to inherit this estimation error. Instead of solving the optimization problem on the basis of the choice of unique parameter values, the investor can choose the Bayesian approach, because it accounts for parameter uncertainty, as seen in Kang (2011) and Jacquier and Polson (2012), for example.

Portfolio decision problems in Bayesian setting are usually solved by choosing such portfolio weights, that maximize the expected utility of the portfolio **with respect to the predictive density of the one-step-ahead returns**:

$$\hat{p}^*_{t+1} = \arg\max_p \int U(p' r_{t+1}) f(r_{t+1}|r_1,\ldots,r_t) dr_{t+1}.$$

Therefore, the investor would obtain point optimal portfolio weights, where the parameter uncertainty has been accounted for. In our case, we do not have the analytically tractable posterior distribution of the returns, just a sample of size M.

[1] $\rho_1 = v_1$, $\rho_j = (1-v_1)\ldots(1-v_{j-1})v_j$, $v_j \sim \mathcal{B}(1,c)$

We can approximate the solution:

$$\hat{p}^*_{t+1} \approx \arg\max_p \frac{1}{M} \sum_{m=1}^{M} U(p' r_{t+1,m}),$$

where $\{r_{t+1,m}\}_{m=1}^M$ is a predictive sample of one-step-ahead returns, obtained as explained in the previous section. However, this approach provides only with a point estimation of the optimal portfolio weights. Since the analytically tractable posterior distribution is not available, it is not straightforward to obtain measures of uncertainty as credible intervals in order to asses the quality of this estimation, see Brandt (2009), for example. Moreover, this approach does not provide the measure of uncertainty for subsequent portfolio characteristics, such as the portfolio expected return, variance or expected utility.

Alternatively, we propose to obtain samples from the entire **posterior distribution of optimal portfolio weights** $f(p^*_{t+1}|r_1, \ldots, r_t)$. This approach relies on solving the allocation problem at every MCMC iteration and approximate for example the posterior mean of the optimal portfolio weights by:

$$\mathrm{E}[p^*_{t+1}|r_1, \ldots, r_t] = \int p^*_{t+1} f(\Theta|r_1, \ldots, r_t) d\Theta \approx \frac{1}{M} \sum_{m=1}^{M} p^*_{t+1,m},$$

where $\{p^*_{t+1,m}\}_{m=1}^M$ is a posterior sample of optimal portfolio weights obtained for each value of the model parameters in the MCMC sample. In other words, since we have assembled M one-step-ahead volatility matrices and mean vectors, we can solve the portfolio allocation problem M times. Similarly, we can approximate the posterior median of p^*_{t+1} and credible intervals by using the quantiles of the sample of optimal portfolio weights. In this manner, we are able to obtain a sample from the posterior distribution of portfolio expected returns $\{\mathrm{E}[r^P_{t+1}]_m\}_{m=1}^M = \{p^{*'}_{t+1,m}\mu_m\}_{m=1}^M$, variance $\{\mathrm{Var}[r^P_{t+1}]_m\}_{m=1}^M = \{p^{*'}_{t+1,m} H^{*'}_{t+1,m} p^*_{t+1,m}\}_{m=1}^M$ and expected utility $\{\mathrm{E}[U_{t+1}]_m\}_{m=1}^M$.

In hedging exercise, same as in the portfolio allocation problem above, it is possible to obtain a sample from the posterior distribution of the optimal hedge ratios D^*_{t+1}. As for the point estimate, the investor can choose either the posterior sample mean or median, which can be obtained from the collection of optimal hedge ratios $\{D^*_{t+1,m}\}_{m=1}^M$.

4 Simulation Study

The goals of this simulation study are to show the flexibility and adaptability of the DPM model and to explain some bimodal posterior densities that we later observe in real data applications. Here we use the basic GJR-AGDCC model as in (11) by Cappiello *et al.* (2006). We have generated three bivariate time series of 3000 observations with the following errors: Gaussian $\mathcal{N}(0, I_2)$, Student-t $\mathcal{T}(I_2, \nu = 8)$ and a mixture of two bivariate Normals $0.9\mathcal{N}(0, \sigma_1 = 0.8, \sigma_{12} = 0.0849, \sigma_2 = 0.9) + 0.1\mathcal{N}(0, \sigma_1 = 2.8, \sigma_{12} = -0.7637, \sigma_2 = 1.9)$. In the mixture data we have chosen a bigger variance of the second component for the first series, to make it more volatile than the second. Then, estimate all three data sets using a DPM model. The point estimates are not reported in the paper because of the limited space. All parameters were estimated well, with true parameters always inside the 95% credible intervals. The length of MCMC chain is 10,000 burn-in plus 20,000 iterations. The contour plots in Figure 1 compare true predictive densities of returns $f(r_{t+1}|\mathcal{I}_t)$ with the estimated ones, which were obtained by sampling 5 observations at each MCMC step, resulting in a sample size of 100,000, as explained in the previous section. As we can see, the estimation results are very precise compared to the true contour of one-step-ahead returns. The contours can be seen as a summary of the estimation results for all 13 parameters of the model $\Phi = (\mu, \omega, \alpha, \beta, \phi, \kappa, \lambda, \delta)$ and the distribution for the error term. Therefore, the infinite mixture model is a flexible tool that is able to adjust to whatever distribution the data comes from.

Figure 1 goes here

Next, the top part of the Figure 2 presents the marginal kernel smoothing densities for one-step-ahead errors, $f(\epsilon_{t+1}|r_1, \ldots, r_r)$. The densities are symmetrical, thus we present only the left and right tails for each marginal series. The Student-t data predicts fatter right and left tails, consequently, allowing for more extreme observations, which in turn increase the volatility. The mixture model predicts fat tails just for the first marginal series, but not for the second. This is due to the nature of the simulated data: the second mixture component is more volatile for the first series than for the second ($\sigma_1 = 2.8$, $\sigma_2 = 1.9$). The effect of how fat tails increase volatility can be seen in the middle row of Figure 2, where the kernel smoothing densities of the elements of the error covariance matrices $\{\Lambda^{-1}_{t+1,m}\}_{m=1}^{M}$ are presented. The Gaussian error data does not allow for extreme observations, therefore does not allow for high volatilities. However, in a Student-t and mixture data the predictive density of the variance of the errors has a very fat right tail, which means high volatilities. Because of the more extreme returns, the density of the variance of the errors is not symmetrical. And finally, the bottom row of Figure 2 presents the predictive densities of the volatilities of the returns $\{H^*_{t+1,m}\}_{m=1}^{M}$. Here the effect of the fat tails is even more pronounced, since we also take into account the asymmetric volatility effect. The middle graph, which in the middle row was symmetrical, here is also bimodal, because it models the asymmetric correlation effect. This simulation study helps to understand and explain the bimodal posterior distributions that appear in the real data application in the next Section.

Figure 2 goes here

5 Data and Results

In the illustration using real data, same as in the simulation study, for the sake of simplicity we use the basic GJR-AGDCC model, where κ, λ and δ are scalars as in (11) by Cappiello *et al.* (2006).

In this section we illustrate the financial applications described in Section 2. At first we solve portfolio allocation problem for bivariate time series for utility-based and GMV approaches. Then, once we have the portfolio, we find the optimal hedge ratios under minimum variance and maximum utility criteria using futures of a certain index. All data was obtained from IHS Global Insight database and Yahoo Finance.

5.1 Portfolio Allocation

For the first illustration we use the daily price data of Apple Inc. company (P_t^A) and NASDAQ Industrial index (P_t^N) from January 1, 2000 till May 7, 2012. Then, daily prices are transformed into daily logarithmic returns (in %), resulting in 3098 observations. Table 1 provides the basic descriptive statistics, and Figure 3 illustrates the dynamics of returns.

Table 1. Descriptive Statistics of the Apple Inc. and NASDAQ Ind. Return Series

	$100 \times \ln\left(\frac{P_t^A}{P_{t-1}^A}\right)$	$100 \times \ln\left(\frac{P_t^N}{P_{t-1}^N}\right)$
Mean	0.0973	0.0020
Median	0.1007	0.0766
Variance	9.7482	3.1537
Skewness	−4.2492	−0.1487
Kurtosis	102.0411	7.1513
Correlation	0.5376	

Figure 3 goes here

As expected, the Apple Inc. has higher overall variance because of the higher mean return. Both returns do not exhibit any evidence of auto-regressive behavior. Apple Inc. returns contain one atypical data point, corresponding to September 29, 2000. The very low return is due to an announcement the day before about lower than expected sales. The estimation results of the AGDCC model are reported in the Table 2. The ML and RWMH estimation approaches, assuming Gaussian innovations, as expected, provide very similar estimates. DPM model exhibits differences in point estimations for the ϕ parameter, which measures the asymmetric volatility effect for marginal return series. Also, ω parameter for both Gaussian models is larger than in DPM, probably because of the atypical data in the return series. Gaussian error models have to inflate the constant volatility parameter in order to accommodate this atypical data point, whereas in the DPM model is has been accounted for in one of the mixture components of the distribution of the errors.

Table 2. Estimation Results for Apple Inc. (1) and NASDAQ Ind. (2) Returns, 20,000 iterations plus 10,000 burn-in

	ML-Gaussian Errors Parameter (st.dev.)	Bayesian Gaussian Errors Post. mean (Post.st.dev.)	Bayesian DPM Post. mean (Post.st.dev.)
μ_1, μ_2	0.0973, 0.0020	0.1628,0.0231	0.1404, 0.0417
	(0.0296),(0.0149)	(0.0408),(0.0216)	(0.0362),(0.0215)
ω_1, ω_2	0.1751, 0.0232	0.2676, 0.0277	0.1324, 0.0206
	(0.0142),(0.0024)	(0.0599),(0.0053)	(0.0450),(0.0055)
α_1, α_2	0.0672, 0.0059	0.0923, 0.0146	0.0692, 0.0125
	(0.0050),(0.0047)	(0.0135),(0.0073)	(0.0149),(0.0066)
β_1, β_2	0.8725, 0.9250	0.8396, 0.9226	0.8893, 0.9250
	(0.0042),(0.0053)	(0.0166),(0.0081)	(0.0173),(0.0078)
ϕ_1, ϕ_2	0.1095, 0.1189	0.1090, 0.1016	0.0506, 0.0804
	(0.0094),(0.0075)	(0.0252),(0.0124)	(0.0234),(0.0168)
κ	0.0231	0.0113	0.0140
	(0.0027)	(0.0090)	(0.0078)
λ	0.9628	0.9794	0.9516
	(0.0047)	(0.0195)	(0.0314)
δ	0.0061	0.0051	0.0182
	(0.0049)	(0.0042)	(0.0117)

After the estimation was carried out using the Gaussian errors and the infinite mixture of Gaussian distributions, we are able to approximate the predictive one-step-ahead return distributions $f(r_{t+1}|\mathcal{I}_t)$. Figure 4 shows the tails of the kernel smoothing marginal densities of the one-step-ahead returns. We can observe the differences in tails arising from different specification of the errors. The DPM model permits for a more flexible distribution, therefore, for more extreme returns, i.e. fatter tails. The kernel smoothing densities were obtained using the same procedure as in the simulation study: by sampling 5 observations at each MCMC step, resulting in a sample size of 100,000.

Figure 4 goes here

Table 3 presents the estimated mean, median and 95% credible intervals of one-step-ahead volatility matrices in Bayesian context, obtained from the collection of M matrices $\{H^*_{t+1,m}\}_{m=1}^{M}$. Both Gaussian models provide very similar volatility estimates. In the constant unconditional approach, the individual volatilities and covariances are much greater. Unconditional covariance matrix gives estimated volatility of 9.75 for Apple Inc. and 3.15 for NASDAQ Ind. with the correlation of 0.5375, whilst the dynamic Gaussian and DPM models for $t+1$ estimate 7.42 and 7.50 for Apple Inc. and twice as small for NASDAQ Ind. 1.59 and 1.56 correspondingly, with correlations of 0.5329 and 0.4972 respectively. In Gaussian model the means of the marginal volatilities are very similar to the medians, meaning, that the distributions are symmetric. In DPM model the means are greater than the medians, which means that the posterior distributions are skewed to the right, i.e. they have longer right tails. This happens because DPM model permits some returns to come from a very volatile component, as seen in Figure 4 and explained in the simulation study in Section 4. Figure 5 illustrates the right tail of the posterior distribution of one-step-ahead volatilities of Apple Inc. returns, where, $P\left(H_{t+1}^{*(1,1)} > 13.45|\mathcal{I}_t\right) = 0.050$, $P\left(H_{t+1}^{*(1,1)} > 60.51|\mathcal{I}_t\right) = 0.010$. This reads, that there is 1% chance to observe a data point, that is more volatile than 60.51 and 5% chance to observe volatility greater than 13.45.

Figure 5 goes here

Table 3. Estimated Means, Medians and 95% Credible Intervals of One-Step-Ahead Volatilities of the Apple Inc. and NASDAQ Ind. Return Series

Matrix element		(1,1)	(1,2), (2,1)	(2,2)
Constant	Σ	9.7482	2.9805	3.1537
ML Gaussian	H_{t+1}	7.4631	1.6646	1.5729
Bayesian Gaussian	mean	7.4185	1.8330	1.5947
	median	7.4272	1.8225	1.5930
	95%	6.9514, 7.8445	1.6572, 2.0515	1.4759, 1.7237
Bayesian DPM	mean	7.5016	1.7010	1.5605
	median	4.7396	1.5844	1.4810
	95%	3.2745, 44.9882	0.8202, 3.4372	0.9724, 3.1264

Utility and GMV portfolios. Here we solve for the utility-based and GMV portfolios without the short-sale constraint, as in (1)-(2). The risk aversion coefficient $\gamma = 0.03$, which means that the penalty for increased variance is $\gamma/2$. The choice of the risk-aversion coefficient is arbitrary. Table 4 shows the estimated means and credible intervals of optimal portfolio weights p_{t+1}^{*}, portfolio return r_{t+1}^{P}, portfolio variance σ_{t+1}^{2P} and investor's utility from the portfolio U_{t+1}^{P}. In the utility-based case, the optimal portfolio weights for constant model are very different for the rest of the models: it suggests investing 48% of the wealth in Apple Inc. shares, and the rest in the index. Gaussian and DPM models estimate more similar portfolio weights, 0.83 and 0.94, where both mean estimates enter in the credible intervals of the others. Gaussian 95% credible interval is around two times smaller than the DPM, which is expected, since DPM model permits for fatter tails in returns and volatilities. The rest of the columns present the estimated expected returns, variances and utilities for the utility-based portfolio. Gaussian and DPM models result in very similar point estimates, whilst the DPM model permits for fatter tails in all cases. The 95% credible intervals for portfolio returns, variances and utility is from 25 to 30% wider in the DPM setting. This illustrates, that if the investor chooses to impose a restrictive distribution of the errors, she would not be able to measure the uncertainty of her financial decisions correctly. In Gaussian setting, she would be overconfident. Similar conclusions can be drawn from the GMV-based portfolio estimation results. The constant covariance model proposes investing 2.5% of the funds in Apple Inc. shares, however, the total portfolio risk is as twice as big as in Gaussian and DPM models. Both models propose short-selling the risky asset and investing all the funds plus the income from short-selling into a less risky asset - the index. In GMV-based portfolio the differences in uncertainty are even more obvious: the credible interval for $p^*_{GMV,t+1}$ for DPM model is more than three times wider. The general differences in tails can be observed in the Figures 6 and 7.

Table 4. Estimated Utility-Based and GMV Portfolio Weights of the Apple Inc. and NASDAQ Ind. Return Series

	$p^*_{U,t+1}$			$p^*_{GMV,t+1}$		
	Constant	Bayesian Gauss	Bayesian DPM	Constant	Bayesian Gauss	Bayesian DPM
p^*_{t+1}	0.4825	0.8253	0.9359	0.0249	-0.0453	-0.0312
(95%)		(0.4070, 1.2277)	(0.0682, 1.9681)		(-0.0818, -0.0132)	(-0.1833, 0.1215)
r^P_{t+1}	0.0479	0.1456	0.1425	0.0044	0.0168	0.0386
(95%)		(0.0406, 0.2807)	(0.0324, 0.3332)		(-0.0258, 0.0604)	(-0.0055, 0.0828)
σ^{2P}_{t+1}	4.6022	5.8757	4.9925	3.1493	1.5822	1.5301
(95%)		(2.6560, 10.2364)	(1.7159, 11.1193)		(1.4600, 1.7131)	(0.9398, 3.0042)
U^P_{t+1}	-0.0211	0.0575	0.0676			
(95%)		(-0.0074, 0.1302)	(-0.0095, 0.1704)			

Figures 6 and 7 present the kernel smoothing densities of portfolio returns, variances and utilities for Gaussian and DPM models. In all cases DPM model exhibits fatter right tail and wider credible intervals, which arises as a consequence of the fat-tailed return distribution.

Figure 6 goes here

Figure 7 goes here

To sum up, these portfolio allocation exercises helped to illustrate the direct consequences of return distribution to the uncertainty of financial decisions. DPM model permits the investor to perform inference and prediction about the returns and their volatilities without imposing arbitrary restrictions on the data generating process.

5.2 Hedging

Once the optimal portfolio has been selected, the investor might want to eliminate risk associated from the movements in the market. Say, that the S&P 500 Index represents our well-diversified portfolio, therefore, S&P 500 Index futures can be used to hedge it. As Hull (2012) notices, hedging using index futures eliminates the market risk and the performance of portfolio depends only on the performance relative to the market.

For this illustration we use the daily price data of S&P 500 index (P^s_t) and E-mini S&P 500[2] (from now on, just E-mini) (P^f_t) from January 1, 2000 till May 11,

[2]E-mini S&P 500 is a stock market index futures contract, traded on the Chicago Mercantile Exchange's Globex electronic trading platform. The value of the future contract is 50 US dollars times the E-mini S&P 500 futures price. It is a smaller version ($1/5^{th}$) of regular S&P 500 futures contract, made to be available to smaller investors.

2012. The E-mini prices are quoted for the second expiration future nearby. Daily prices are transformed into daily logarithmic returns (in %), resulting in 3109 observations. Table 5 provides the basic descriptive statistics, and Figure 8 illustrates the dynamics of returns.

Table 5. Descriptive Statistics of the S&P 500 Index and E-mini Futures Return Series

	$100 \times \ln\left(\frac{P_t^s}{P_{t-1}^s}\right)$	$100 \times \ln\left(\frac{P_t^f}{P_{t-1}^f}\right)$
Mean	-0.0063	-0.0023
Median	0.0592	0.0551
Variance	1.9417	1.8805
Skewness	0.0047	-0.1576
Kurtosis	14.5295	10.1901
Correlation	0.9792	

Figure 8 goes here

The estimation was carried out using Gaussian distribution and the infinite mixture of Gaussian distributions for the errors. Estimation results are presented in Table 6. Both Bayesian approaches provide similar estimates, whereas ML has problems in estimating α and δ. Next, Figure 9 presents the predictive kernel smoothing densities for marginal returns, obtained by sampling 5 returns at each MCMC step, resulting in a sample of 100,000. Observe, that the DPM model allows for fatter tails - more extreme returns.

Table 6. Estimation Results for S&P 500 (1) and E-mini (2) Returns, 20,000 iterations plus 10,000 burn-in

	ML-Gaussian Errors Parameter (st.dev.)	Bayesian Gaussian Errors Post. mean (Post.st.dev.)	Bayesian DPM Post. mean (Post.st.dev.)
μ_1, μ_2	-0.0063, -0.0023	-0.0159, -0.0142	0.0222, 0.0219
	(0.0104),(0.0105)	(0.0150),(0.0153)	(0.0152),(0.0157)
ω_1, ω_2	0.0188, 0.0166	0.0199, 0.0206	0.0161, 0.0174
	(0.0009),(0.0010)	(0.0023),(0.0026)	(0.0032),(0.0036)
α_1, α_2	3.9×10^{-9}, 5.9×10^{-9}	0.0126, 0.0208	0.0165, 0.0224
	(4.6×10^{-6}), (1.7×10^{-6})	(0.0060),(0.0073)	(0.0069),(0.0082)
β_1, β_2	0.9059, 0.9142	0.9099, 0.9046	0.9122, 0.9084
	(0.0026),(0.0027)	(0.0068),(0.0082)	(0.0087),(0.0098)
ϕ_1, ϕ_2	0.1583, 0.1458	0.1162, 0.1123	0.1076, 0.1075
	(0.0055),(0.0055)	(0.0094),(0.0096)	(0.0180),(0.0177)
κ	0.1336	0.1402	0.1153
	(0.0061)	(0.0153)	(0.0184)
λ	0.5802	0.4114	0.5276
	(0.0328)	(0.1067)	(0.0866)
δ	4.6×10^{-8}	0.0083	0.0106
	(1.8×10^{-5})	(0.0077)	(0.0100)

Figure 9 goes here

Then, as in the portfolio optimization problem, we are able to obtain one-step-ahead volatilities, where the estimated variances and covariances are presented in the Table 7. The mean, median and 95% credible intervals of one-step-ahead volatility matrices in Bayesian context are obtained from the collection of M matrices $\{H^*_{t+1,m}\}_{m=1}^M$. Here, the constant unconditional approach estimates variances that are twice as big than in dynamic models, for both time series. The mean correlation is similar in all three cases: 0.9792 for unconditional, 0.9723 for Gaussian and 0.9757 for DPM model. The means and medians in Gaussian model are almost the same, therefore, the posterior one-step-ahead distribution of volatilities is symmetric. However, in DPM model the means are greater than the medians, meaning that the posterior distribution is positively skewed. In order to investigate in more detail the posterior of one-step-ahead volatility matrix in DPM model we draw the kernel smoothing densities, which are presented in Figure 10. Obviously, they are bimodal. As seen in simulation study in Section 4, this can be explained by the fact that there are some returns coming from a more than usual volatile component (-s). In addition, the asymmetric volatility pronounces this effect even more for both marginal variances and asymmetric correlation effect causes the right graph of Figure 10 to have bimodal-type form.

Figure 10 goes here

Table 7. Estimated Means, Medians and 95% Credible Intervals of One-Step-Ahead Volatilities of the S&P 500 and E-mini Return Series

Matrix element		(1,1)	(1,2), (2,1)	(2,2)
Constant	Σ	1.9417	1.8712	1.8805
ML Gaussian	H_{t+1}	0.8851	0.8128	0.7885
Bayesian Gaussian	mean	0.8197	0.7532	0.7315
	median	0.8196	0.7531	0.7311
	95%	0.7794, 0.8607	0.7170, 0.7908	0.6949, 0.7695
Bayesian DPM	mean	0.8501	0.7853	0.7620
	median	0.7034	0.6593	0.6534
	95%	0.3540, 2.0363	0.3378, 1.7537	0.3432, 1.6121

Utility and GMV Hedging Ratios. Finally, using the estimated one-step-ahead volatilities and means, we solve the hedging problem in two cases: using the minimum variance criteria and the maximization of expected utility, as in (3)-(4). Here we use $\gamma = 0.3$, which means that the penalization for increased variance is greater than in portfolio allocation problem, because our principle objective now is the reduction of variance rather than gain from the return. Table 8 presents the means and 95% credible intervals of optimal hedge ratios D^*_{t+1}, the total hedged portfolio return r^{HP}_{t+1}, the hedged portfolio variance σ^{2HP}_{t+1} and utility U^{HP}_{t+1}. The mean optimal utility-based and GMV hedge ratios for Gaussian model enters the DPM credible intervals, but not vice versa. The medians for utility-based portfolio are 1.09677 and 0.9194 for Gaussian and DPM models respectively, indicating, that the posterior distribution of optimal hedge ratios for DPM model is positively skewed, meanwhile the Gaussian model provides symmetrical posterior densities. The same is observed for the GMV portfolio, where the medians are 1.0297 and 1.0053 for Gaussian and DPM models respectively. Even though the hedging ratios are different for both approaches, they result into statistically equal hedged portfolio variances and utilities. The constant covariance matrix hedging approach, compared to both time-varying Bayesian models, provides worse hedged portfolio with high variance and small utility. In all the cases, credible intervals for DPM model are from 20% up to six times wider. This also can be observed from Figures 11 and 12.

Table 8. Estimated Utility-Based and GMV Optimal Hedging Ratios of the S&P 500 and E-mini Return Series

		$D^*_{U,t+1}$			$D^*_{GMV,t+1}$	
	Constant	Bayesian Gauss	Bayesian DPM	Constant	Bayesian Gauss	Bayesian DPM
D^*_{t+1}	0.9992	1.0947	0.8974	0.9950	1.0298	1.0164
(95%)		(0.9539, 1.2265)	(0.6089, 1.1006)		(1.0178, 1.0422)	(0.9479, 1.1188)
r^T_{t+1}	-0.0039	0.0007	0.0039	-0.0039	-0.0013	-3$\times10^{-5}$
(95%)		(-0.0071, 0.0110)	(-0.0040, 0.0180)		(-0.0077, 0.0050)	(-0.0065, 0.0062)
σ^{2T}_{t+1}	0.0799	0.0506	0.0522	0.0798	0.0440	0.0390
(95%)		(0.0423, 0.0723)	(0.0239, 0.1374)		(0.0413, 0.0469)	(0.0205, 0.1283)
U^T_{t+1}	-0.0159	-0.0069	-0.0039			
(95%)		(-0.0138, 0.0008)	(-0.0193, 0.0064)			

Next, Figures 11 and 12 present kernel smoothing densities of optimal hedge ratios, hedged portfolio expected return, variance and expected utility for both hedging approaches. As commented before, the DPM credible intervals are wider due to the fatter tails in the return distribution. In Figure 12 the second and third graphs demonstrate strong bimodality of the posterior densities, which is a direct consequence of the bimodality in predictive volatility densities.

Figure 11 goes here

Figure 12 goes here

Here we present a short numerical illustration for the portfolio hedging problem. As seen in Hull (2012), the optimal number of future contracts is given by $N^* = D^*(V_P/V_F)$, where D^* is the optimal hedge ratio, V_P is the value of the portfolio and V_F is the value of one futures contract. In time-varying setting, the value of the portfolio or futures contract depends on the prices at $t+1$, that can be obtained as $P_{t+1} = P_t \cdot \exp\{r_{t+1}/100\}$, where P_t is the current price. Therefore, the optimal number of contracts at each MCMC iteration is:

$$N^*_{m,t+1} = D^*_{m,t+1} \cdot \frac{\text{shares}}{50} \cdot \frac{P^P_t \exp\{r^P_{t+1}/100\}}{P^F_t \exp\{r^F_{t+1}/100\}}.$$

Say, an agent owns 1,000 shares, the price of one share of the portfolio is $1,412.00$ and the E-mini futures price for the second expiration contract nearby at the moment is $1,400.00$. The optimal number of contracts for constant covariance model is

20.0715, Gaussian - 20.7727, DPM - 20.5022. In Bayesian models the optimal number of contracts is a sample mean of MCMC chain, whose posterior kernel smoothing densities can be seen in Figure 13. So, if the investor employs a time-invariant approach, she would short $N^*_{t+1} = 20$ E-mini futures contracts, however, in order to minimize the portfolio variance at $t + 1$, the investor should short $N^*_{t+1} = 21$ in Gaussian and $N^*_{t+1} = 21$ in DPM. The 95% credible interval for Gaussian model is (20.5298, 21.0221) and for DPM - (19.1196, 22.5682), which is 7 times wider. Since Gaussian errors do not permit extreme returns and high volatilities, the Gaussian credible interval for N^*_{t+1} is smaller than in reality, making the investor overconfident about her decisions and ignore risk arising from the extreme returns.

Figure 13 goes here

In portfolio allocation or in hedging context, adjusting portfolio weights at each period might lead to high transaction costs, thus the investor will adjust her portfolio only if the expected utility after the adjustment minus the transaction costs is greater than the expected utility without the adjustment. Both illustrations have shown the differences in error specifications in using real data. We have illustrated how quantification of uncertainty reflects distributional assumptions of the errors.

6 Conclusion

In this paper we have considered the constant and dynamic portfolio allocation and hedging problems, where the time-varying covariance matrix was estimated using a GJR-AGDCC model, that captures asymmetric volatilities and correlations. For the error term we have proposed a flexible infinite mixture of Gaussian distributions, which was handled using Bayesian non-parametric approach. We have presented a short simulation study that illustrates the differences arising from different assumptions for the errors and shows the adaptability of the DPM model.

We have employed the proposed approach to solve the portfolio allocation and hedging problems using real data of asset returns. In both applications we have showed that even though the point estimates for optimal hedge ratios and optimal portfolio weights are very similar for Gaussian and infinite mixture models, the non-parametric credible intervals are wider. Therefore, the normality assumptions forces the investor to be overconfident about her estimates. Moreover, the non-parametric model allowed for some one-step-ahead volatilities come from a very volatile component, thus making the posterior distribution of covariance matrix asymmetric.

The explained methodology and obtained results are not limited to the two specific risk management problems and could be expanded into various other topics in applied finance and risk management.

References

Antoniak CE. 1974. Mixtures of Dirichlet Processes with Applications to Bayesian Nonparametric Problems. *The Annals of Statistics* **2**: 1152–1174.

Ardia D, Hoogerheide LF. 2010. Efficient Bayesian Estimation and Combination of GARCH-Type Models. In *Rethinking Risk Measurement and Reporting: Examples and Applications from Finance: Vol II*, Bocker K (ed), chap 1. RiskBooks: London.

Ausín MC, Galeano P, Ghosh P. 2011. A Semiparametric Bayesian Approach to GARCH-Type Models with Application to Value at Risk Estimation. *Working Paper* : 1–33

Avramov D, Zhou G. 2010. Bayesian Portfolio Analysis. *Annual Review of Financial Economics* **2**: 25–47. ISSN 1941-1367. DOI: 10.1146/annurev-financial-120209-133947

Basak S, Chabakauri G. 2012. Dynamic Hedging in Incomplete Markets: a Simple Solution. *Review of Financial Studies* **25**: 1845–1896. ISSN 0893-9454. DOI: 10.1093/rfs/hhs050

Bauwens L, Laurent S, Rombouts JVK. 2006. Multivariate GARCH Models: a Survey. *Journal of Applied Econometrics* **21**: 79–109. ISSN 0883-7252. DOI: 10.1002/jae.842

Becherer D. 2004. Utility-Indifference Hedging and Valuation via Reaction-Diffusion Systems. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **460**: 27–51. ISSN 1364-5021. DOI: 10.1098/rspa.2003.1234

Bollerslev T. 1986. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* **31**: 307 – 327.

Bollerslev T, Chou RY, Kroner KF. 1992. ARCH Modeling in Finance: a Review of the Theory and Empirical Evidence. *Journal of Econometrics* **52**: 5–59.

Bollerslev T, Engle RF, Nelson DB. 1994. ARCH Models. In *Handbook of Econometrics, Vol. 4*, Engle RF, McFadden D (eds), November. Elsevier: Amsterdam. ISBN 978-0-444-88766-5.

Brandt MW. 2009. Portfolio Choice Problems. In *Handbook of Financial Econometrics*, Ait-Sahalia Y, Hansen L (eds), chap 5. Elsevier: North Holland. ISBN 9780444535542.

Cappiello L, Engle RF, Sheppard K. 2006. Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns. *Journal of Financial Econometrics* **4**: 537–572.

Cecchetti SG, Cumby RE, Figlewski S. 1988. Estimation of the Optimal Futures Hedge. *The Review of Economics and Statistics* **70**: 623–630.

Choudhry T. 2004. The Hedging Effectiveness of Constant and Time-Varying Hedge Ratios Using Three Pacific Basin Stock Futures. *International Review of Economics & Finance* **13**: 371–385. ISSN 10590560. DOI: 10.1016/j.iref.2003.04.002

DeCovny S, Tacchi C. 1991. *Hedging Strategies*. Woodhead-Faulkner: New York.

Delbaen F, Grandits P, Rheinlander T, Samperi D, Schweizer M, Stricker C. 2002. Exponential Hedging and Entropic Penalties. *Mathematical Finance* **12**: 99–123. ISSN 0960-1627. DOI: 10.1111/1467-9965.02001

Eaton ML. 2007. The Wishart Distribution. In *Multivariate Statistics: A Vector Space Approach*, 4, chap 8. Institute of Mathematical Statistics: Beachwood, Ohio. ISBN 978094060069.

Elton EJ, Gruber MJ, Brown SJ, Goetzmann WN. 2003. *Modern Portfolio Theory and Investment Analysis*. John Wiley & Sons, Inc.: USA, 6th edn. ISBN 0-471-23854-6.

Engle RF. 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica* **50**: 987.

Engle RF. 2002a. Dynamic Conditional Correlation: a Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business and Economic Statistics* **20**: 339–350. ISSN 0735-0015. DOI: 10.1198/073500102288618487

Engle RF. 2002b. New Frontiers for ARCH Models. *Journal of Applied Econometrics* **17**: 425–446. ISSN 0883-7252. DOI: 10.1002/jae.683

Escobar MD, West M. 1995. Bayesian Density Estimation and Inference Using Mixtures. *Journal of the American Statistical Association* **90**: 577–588.

Ferguson TS. 1973. A Bayesian Analysis of Some Nonparametric Problems. *The Annals of Statistics* **1**: 209–230.

Fries C. 2007. *Mathematical Finance: Theory, Modeling, Implementation*. John Wiley & Sons, Inc.: Hoboken, New Jersey. ISBN 978-0-470-04722-4.

Giamouridis D, Vrontos ID. 2007. Hedge Fund Portfolio Construction: a Comparison of Static and Dynamic Approaches. *Journal of Banking & Finance* **31**: 199–217. ISSN 03784266. DOI: 10.1016/j.jbankfin.2006.01.002

Glosten LR, Jagannathan R, Runkle DE. 1993. On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *The Journal of Finance* **48**: 1779–1801.

Greyserman A, Jones DH, Strawderman WE. 2006. Portfolio Selection Using Hierarchical Bayesian Analysis and MCMC Methods. *Journal of Banking & Finance* **30**: 669–678. ISSN 03784266. DOI: 10.1016/j.jbankfin.2005.04.008

Hafner CM, Franses PH. 2009. A Generalized Dynamic Conditional Correlation Model: Simulation and Application to Many Assets. *Econometric Reviews* **28**: 612–631. ISSN 0747-4938. DOI: 10.1080/07474930903038834

Hull JC. 2012. *Options, Futures and Other Derivatives*. Pearson Education Limited: Essex, 8th edn. ISBN 9780273759072.

Jacquier E, Polson NG. 2012. Asset Allocation in Finance: a Bayesian Perspective. In *Forthcoming in Hierarchical models and MCMC: a Tribute to Adrian Smith*, Damien, Dellaportas, Polson, Stephen (eds). Oxford University Press.

Jensen MJ, Maheu JM. 2012. Bayesian Semiparametric Multivariate GARCH Modeling. *Working Paper* : 1–37

Jorion P. 1986. Bayes-Stein Estimation for Portfolio Analysis. *The Journal of Financial and Quantitative Analysis* **21**: 279–292.

Kang L. 2011. Asset Allocation in a Bayesian Copula-GARCH Framework: an Application to the "Passive Funds Versus Active Funds" Problem. *Journal of Asset Management* **12**: 45–66. ISSN 1470-8272. DOI: 10.1057/jam.2010.6

Korn R, Korn E. 2001. *Option Pricing and Portfolio Optimization: Modern Methods of Financial Mathematics*, vol 31. American Mathematical Society: USA. ISBN 0-8218-2123-7.

Kroner KF, Sultan J. 1993. Time-Varying Distributions and Dynamic Hedging with Foreign Currency Futures. *The Journal of Financial and Quantitative Analysis* **28**: 535–551.

Kwok YK. 2008. *Mathematical Models of Financial Derivatives*. Springer Berlin Heidelberg: Berlin, 2nd edn. ISBN 978-3-540-42288-4.

Lee CL, Lee ML. 2012. Hedging Effectiveness of REIT Futures. *Journal of Property Investment & Finance* **30**: 257–281. ISSN 1463-578X. DOI: 10.1108/14635781211223824

Lien D. 2009. A Note on the Hedging Effectiveness of GARCH Models. *International Review of Economics & Finance* **18**: 110–112. ISSN 10590560. DOI: 10.1016/j.iref.2007.07.004

Liu SD, Jian JB, Wang YY. 2010. Optimal Dynamic Hedging of Electricity Futures Based on Copula-GARCH Models. In *IEEE IEEEM*. ISBN 9781424485024, 2498–2502.

Markowitz HM. 1952. Portfolio Selection. *Journal of Finance* **7**: 77–91.

McNeil AJ, Frey R, Embrechts P. 2005. *Quantitative Risk Management*. Princeton University Press: Princeton. ISBN 0691122555.

Merton RC. 1972. An Analytic Derivation of the Efficient Portfolio Frontier. *The Journal of Financial and Quantitative Analysis* **7**: 1851–1872.

Papaspiliopoulos O. 2008. A Note on Posterior Sampling from Dirichlet Mixture Models. *Working Paper* : 1–8

Papaspiliopoulos O, Roberts GO. 2008. Retrospective Markov Chain Monte Carlo Methods for Dirichlet Process Hierarchical Models. *Biometrika* **95**: 169–186. ISSN 0006-3444. DOI: 10.1093/biomet/asm086

Park HY, Bera AK. 1987. Interest-Rate Volatility, Basis Risk and Heteroscedasticity in Hedging Mortgages. *Real Estate Economics* **15**: 79–97. ISSN 1080-8620. DOI: 10.1111/1540-6229.00420

Robert CP, Casella G. 2004. *Monte Carlo Statistical Methods*. Springer Texts in Statistics, 2nd edn. ISBN 978-0-387-21239-5.

Rossi E, Zucca C. 2002. Hedging Interest Rate Risk with Multivariate GARCH. *Applied Financial Economics* **12**: 241–251.

Silvennoinen A, Teräsvirta T. 2009. Multivariate GARCH Models. In *Handbook of Financial Time Series*, Mikosch T, KreißJP, Davis RA, Andersen TG (eds). Springer Berlin Heidelberg: Berlin, Heidelberg. ISBN 978-3-540-71296-1.

Teräsvirta T. 2009. An Introduction to Univariate GARCH Models. In *Handbook of Financial Time Series*, Mikosch T, KreißJP, Davis RA, Andersen TG (eds). Springer Berlin Heidelberg: Berlin, Heidelberg. ISBN 978-3-540-71296-1.

Tsay RS. 2010. *Analysis of Financial Time Series*. John Wiley & Sons, Inc., 3rd edn. ISBN 9780471690740.

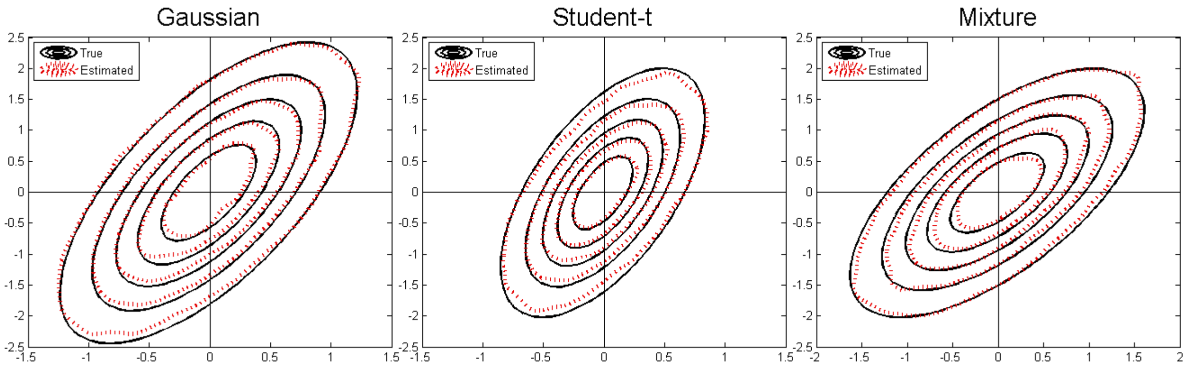
Walker SG. 2007. Sampling the Dirichlet Mixture Model with Slices. *Communications in Statistics - Simulation and Computation* **36**: 45–54. ISSN 0361-0918. DOI: 10.1080/03610910601096262

Yang W, Allen DE. 2004. Multivariate GARCH Hedge Ratios and Hedging Effectiveness in Australian Futures Markets. *Accounting and Finance* **45**: 301–321. ISSN 0810-5391. DOI: 10.1111/j.1467-629x.2004.00119.x

Yilmaz T. 2011. Improving Portfolio Optimization by DCC and DECO GARCH: Evidence from Istanbul Stock Exchange. *International Journal of Advanced Economics and Business Management* **1**: 81–92.

Virbickaite, A., 2012

Figure 1. True and Estimated Contours of the One-Step-Ahead Returns r_{t+1}



Virbickaite, A., 2012

Figure 2. Kernel Smoothing Densities of One-Step-Ahead Errors, their Volatilities and the Volatilities of the Returns

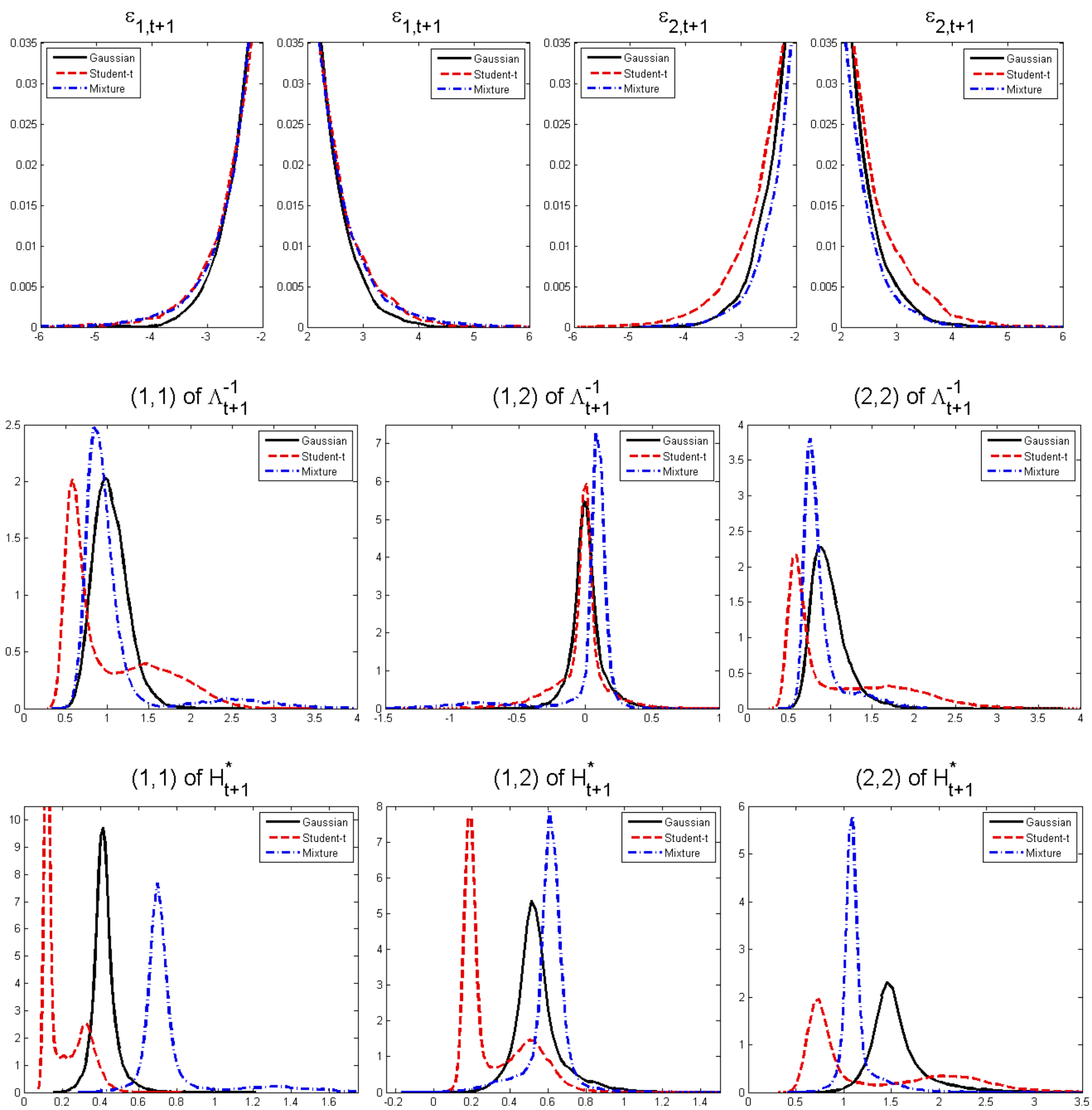
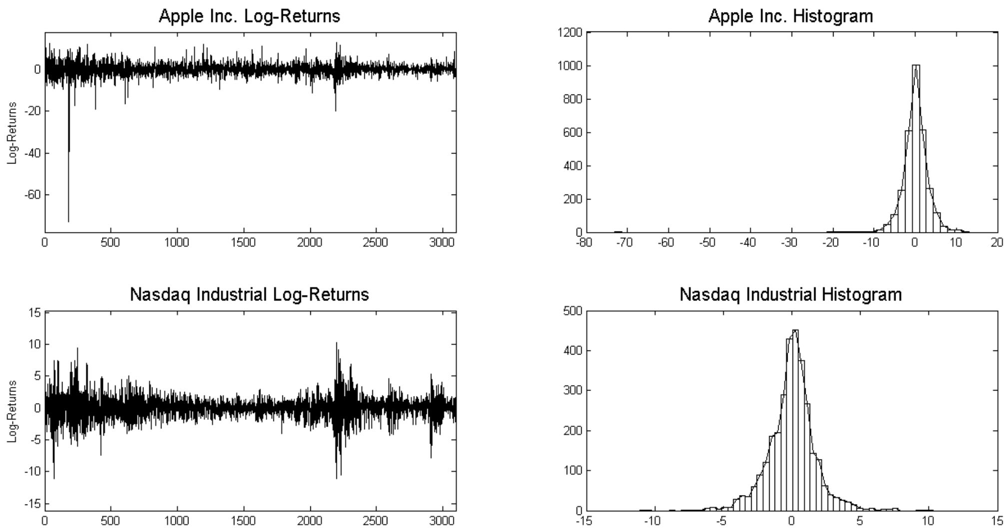


Figure 3. Log-Returns and Histograms of Apple Inc. and NASDAQ Ind. Index



Virbickaite, A., 2012

Figure 4. The Right and Left Tails of Kernel Smoothing Densities of Predictive Returns r_{t+1}

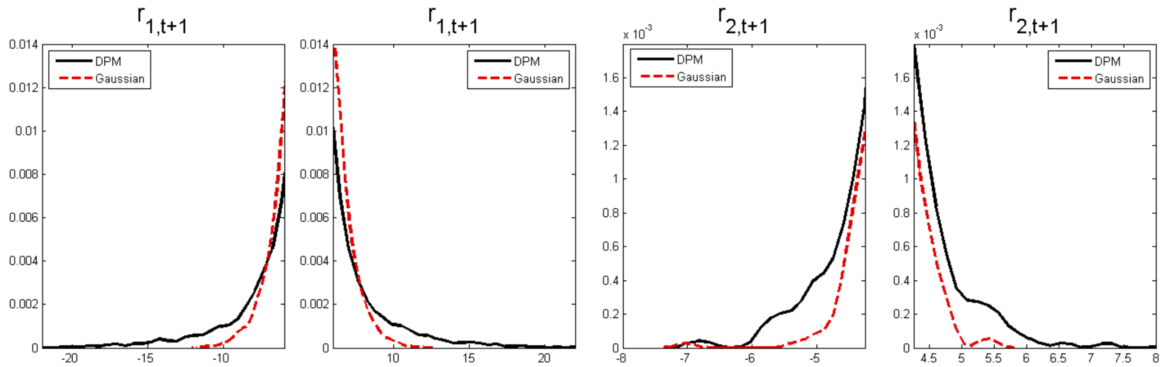
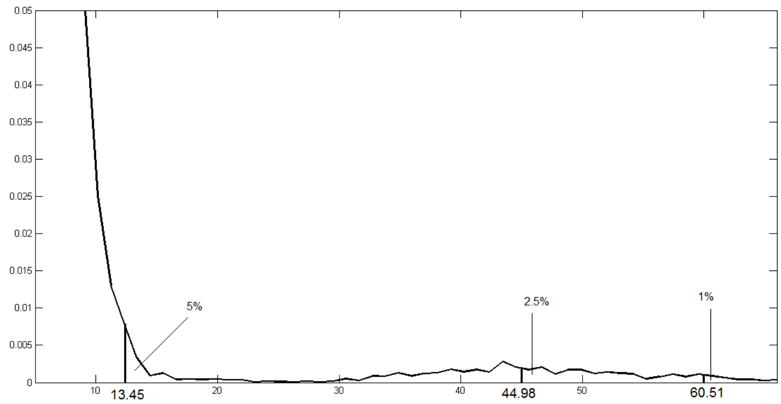
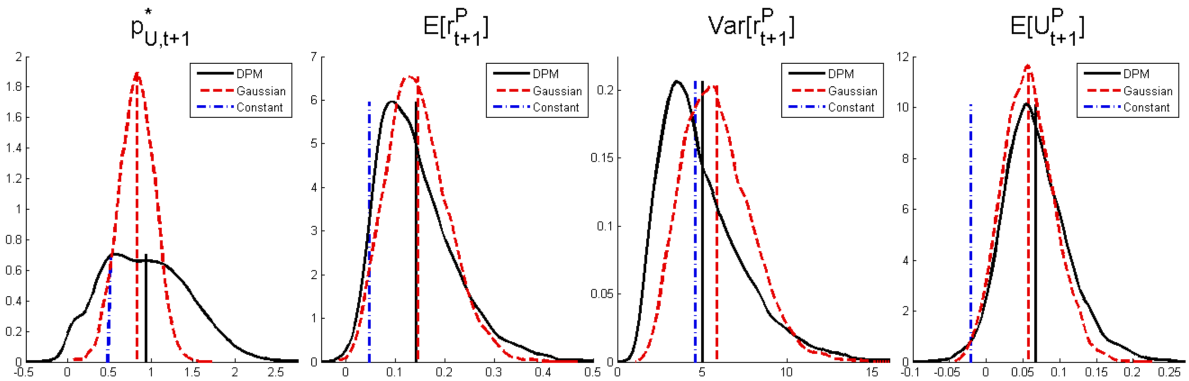


Figure 5. The right Tail of the Posterior Distribution of $H_{t+1}^{*(1,1)}$



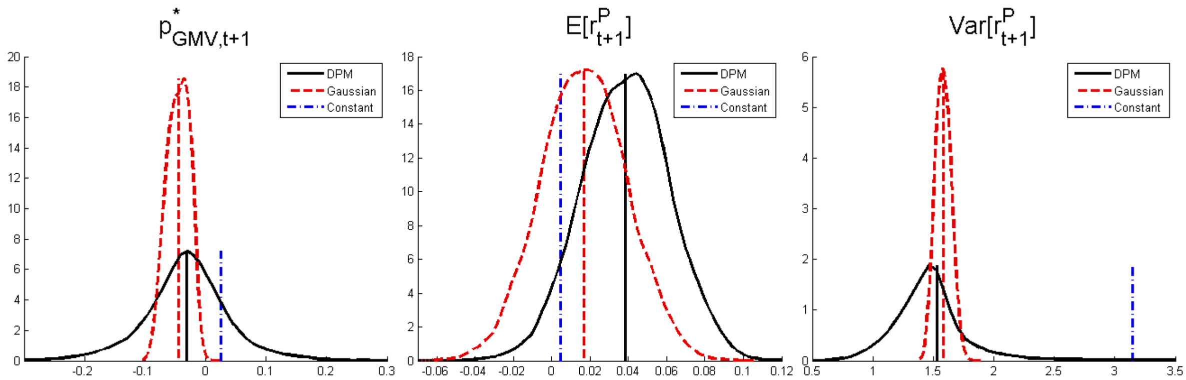
Virbickaite, A., 2012

Figure 6. Posterior Distributions of Utility-Based Portfolio Optimal Weights, Expected Returns, Variance and Expected Utility



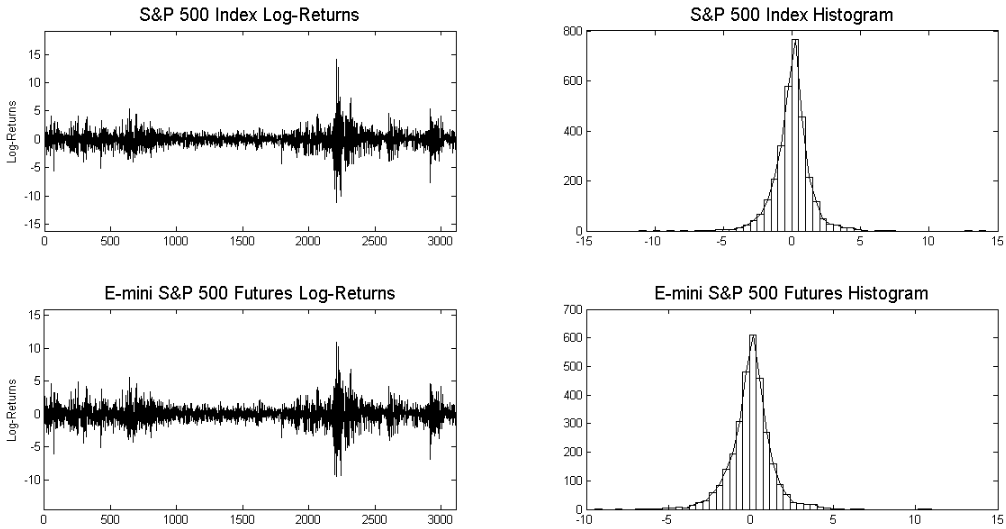
Virbickaite, A., 2012

Figure 7. Posterior Distributions of GMV Portfolio Optimal Weights, Expected Returns and Variance



Virbickaite, A., 2012

Figure 8. Log-Returns and Histograms of S&P 500 Index and E-mini Futures Contracts



Virbickaite, A., 2012

Figure 9. The Right and Left Tails of Kernel Smoothing Densities of Predictive Returns r_{t+1}

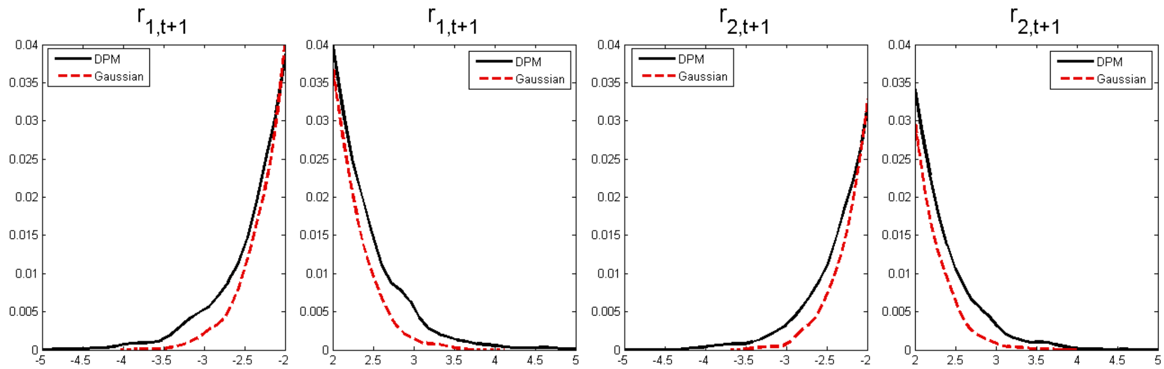


Figure 10. Kernel Smoothing Densities of the Elements One-Step-Ahead Covariance Matrix for DPM Model for S&P 500 and E-mini Data

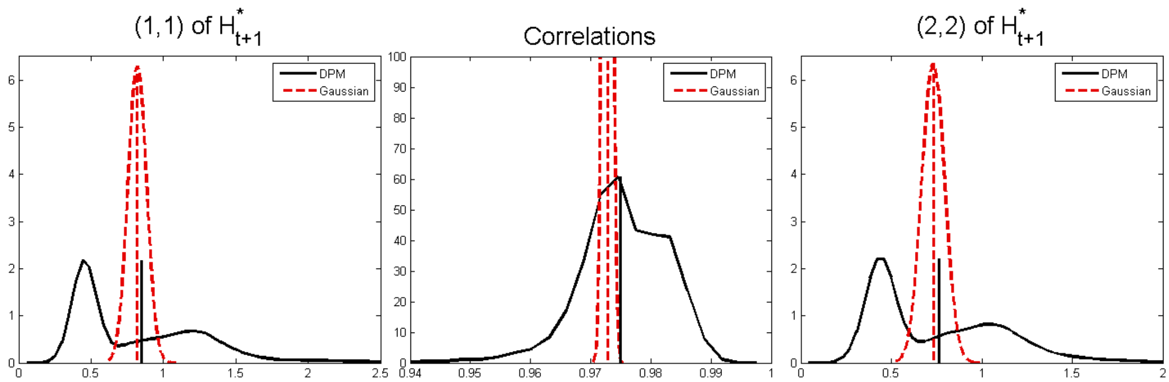
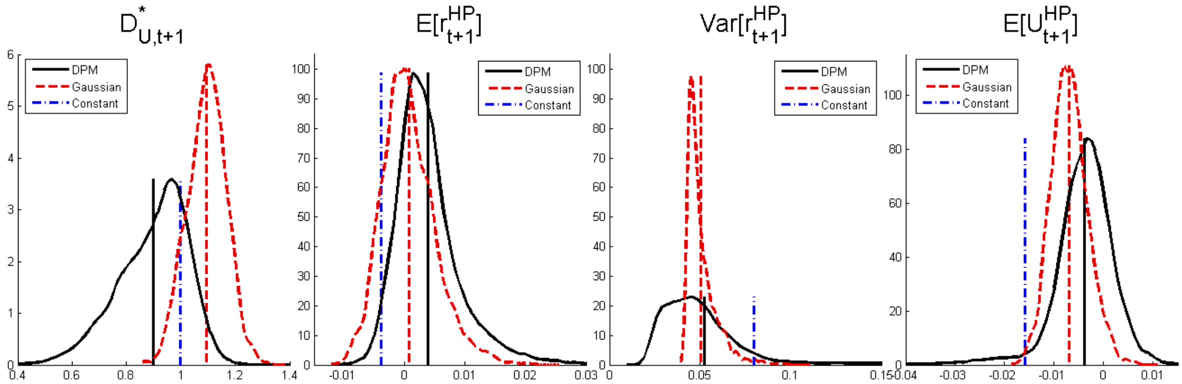
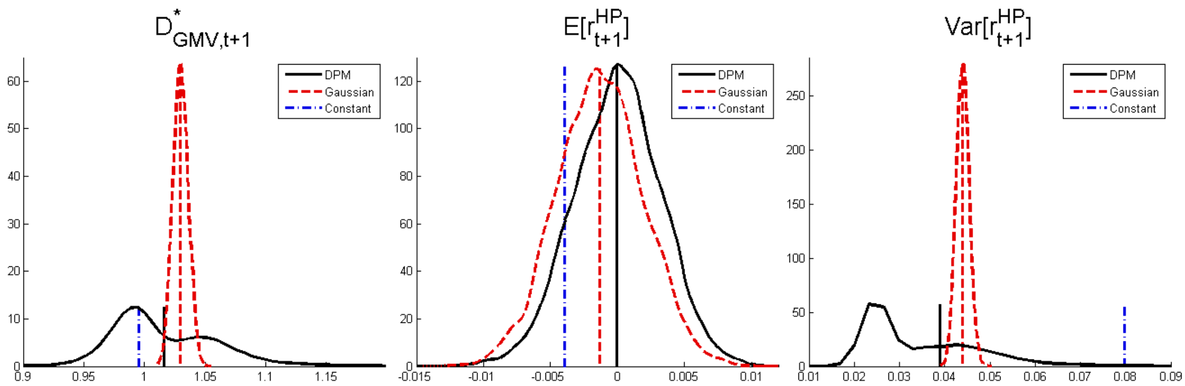


Figure 11. Posterior Distributions of Utility-Based Total Hedged Portfolio Optimal Ratio, Expected Return, Variance and Expected Utility



Virbickaite, A., 2012

Figure 12. Posterior Distributions of GMV-Based Total Hedged Portfolio Hedge Ratio, Expected Return and Variance



Virbickaite, A., 2012

Figure 13. Kernel Smoothing Densities of Optimal Number of Contracts