SCHOOL OF ECONOMICS AND MANAGEMENT
FACULTY OF SOCIAL SCIENCES
AARHUS UNIVERSITY

CREATES
Center for Research in Econometric Analysis of Time Series

CREATES Research Paper 2011-9

Option valuation with the simplified component GARCH model

Matt P. Dziubinski

School of Economics and Management
Aarhus University
Bartholins Allé 10, Building 1322, DK-8000 Aarhus C
Denmark

OPTION VALUATION WITH THE SIMPLIFIED COMPONENT GARCH MODEL

MATT P. DZIUBINSKI

ABSTRACT. We introduce the Simplified Component GARCH (SC-GARCH) option pricing model, show and discuss sufficient conditions for non-negativity of the conditional variance, apply it to low-frequency and high-frequency financial data, and consider the option valuation, comparing the model performance with similar models from the literature. Two volatility components in our model allow us to model time structure of volatility.

JEL Classification. G12, C32.

1. INTRODUCTION

In this paper we introduce a discrete-time volatility model in which the conditional variance of the underlying asset follows a particular GARCH process. Our model can be used for option pricing, while two volatility components allow us to model time structure thereof.

The model builds on Engle and Lee (1999), Heston and Nandi (2000) and Christoffersen et al. (2008) (hereafter referred to as "CJOW") models. The model by Engle and Lee (1999) introduced the volatility component model in the GARCH context, while Heston and Nandi (2000) introduced a model with a closed-form solution for the European call option-pricing formulas. The CJOW model is a generalization of the Heston and Nandi model allowing for a time-varying long-run component. Our model is a simplified specification of the CJOW model, which solves the problem of ensuring the non-negativity of the conditional variance.

The paper proceeds as follows. In Section 2 we provide basic definitions and notation. We introduce the model in Section 3, and discuss the

Date: February 15, 2011.

2000 *Mathematics Subject Classification.* Primary 37M10, 62M10, 91B84; Secondary 62P05.

Key words and phrases. Stochastic volatility, volatility components, GARCH, option pricing.

We wish to thank Timo Teräsvirta for a discussion regarding non-negativity conditions. We acknowledge financial support by the Center for Research in Econometric Analysis of Time Series, CREATES, funded by the Danish National Research Foundation. All errors, omissions and mistakes are author's own responsibility.

estimation thereof in Section 4. In Section 5 we present the estimation results. Section 6 is devoted to option pricing, and, finally, Section 7 contains our conclusions.

2. Basic Definitions and Notation

We assume as given a probability space $(\Omega, \mathcal{F}, P)$ and a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}$, where, depending on the context, we shall assume $\mathbb{T} = \mathbb{Z}_+$ or $\mathbb{T} = \mathbb{Z} \cap [0, T]$, $T > 0$ or $\mathbb{T} = \mathbb{Z} \cap [-1, T]$, $T \geq 0$. We refer to P as the *physical probability measure* and we call $(\Omega, \mathcal{F}, \mathbb{F}, P)$ a *filtered physical probability space*. We shall also use probability measure Q on $(\Omega, \mathcal{F})$ and refer to it as the *risk-neutral probability measure*.

A *stochastic process* X on $(\Omega, \mathcal{F}, P)$ is a collection of $\mathbb{R}$-valued random variables $(X_t)_{t\in\mathbb{T}}$, and we denote it by $X = (X_t)_{t\in\mathbb{T}}$.

The process X is said to be *adapted* if $X_t \in \mathcal{F}_t \ \ \forall t \in \mathbb{T}$ (that is, it is $\mathcal{F}_t$ measurable for each $t \in \mathbb{T}$).

The process X is said to be *predictable* if $X_t \in \mathcal{F}_{t-1} \ \ \forall t \in \mathbb{T}$, and we denote this by $X \in \mathcal{P}$.

The process X is said to be $(\mathbb{F}, P)$-*white noise* with mean μ_X and variance σ_X^2, written $X \overset{P}{\sim} WN(\mu_X, \sigma_X^2)$ if and only if, under probability measure P, X has mean $\mu_X \in \mathbb{R}$ and covariance function $\gamma(s,t) = \sigma_X^2 \delta_{|t-s|}$, where $\delta_h := \mathbb{1}_{\{0\}}(h)$ is the Kronecker delta and $\sigma_X^2 \in \mathbb{R}_{++}$.

The process X is said to be $(\mathbb{F}, P)$-*Gaussian white noise* with mean μ_X and variance σ_X^2, written $X \overset{P}{\sim} GWN(\mu_X, \sigma_X^2)$ if and only if $X \overset{P}{\sim} WN(\mu_X, \sigma_X^2)$ and $X_t \overset{P}{\sim} \mathcal{N}(\mu_X, \sigma_X^2) \quad \forall t \in \mathbb{T}$.

First-order partial differential operator with respect to x is denoted ∂_x.

For further details regarding stochastic processes and time series we refer the reader to Protter (2005) and Brockwell and Davis (1991).

3. The Model

We begin by presenting the CJOW model. The advantages of this model are the existence of a (quasi-)closed-form solution for the option pricing formulas, improved ability to model the smirk and the path of spot volatility and, distinctively, the ability to model the volatility term structure – for details, see Christoffersen et al. (2008). A problem with this model is that the volatility components may admit negative values. This leads to a contradiction in the context of conditional variance modeling, as the conditional variance cannot be negative. We propose a more parsimonious model which solves this problem and discuss its relation to the CJOW model. Furthermore, we shall consider the properties of the model and discuss the estimation of its parameters.

Assumption 1. *The spot asset price, S (including accumulated interest or dividends) follows (over time steps of length $\Delta \equiv 1$) the following process under the physical probability measure P,*

$$r_{t+1} \equiv \log \frac{S_{t+1}}{S_t} = \mu_{t+1} + \sqrt{v_{t+1}} w_{t+1} \tag{3.1}$$

$$v_{t+1} = x_{t+1} + p_v(v_t - x_t) + i_v u_{v,t} \tag{3.2}$$

$$x_{t+1} = m_x + p_x(x_t - m_x) + i_x u_{x,t} \tag{3.3}$$

with

$$\mu_{t+1} = r_f + \lambda v_{t+1} \tag{3.4}$$

$$u_{v,t} = (w_t^2 - 1) - 2 g_v \sqrt{v_t} w_t \tag{3.5}$$

$$u_{x,t} = (w_t^2 - 1) \tag{3.6}$$

$$w \overset{P}{\sim} GWN(0,1) \tag{3.7}$$

where r_f is the continuously compounded interest rate for the time interval of length Δ, v_t is the conditional variance of the log return between $t-1$ and t, with $v \in \mathcal{P}$.

We use a notation that is closely linked to the interpretation of our model. First, the r process is the logarithmic return of the underlying, with μ being its physical conditional mean, while v is its conditional variance. The market price of risk is denoted by λ. Second, the process x is the long-run volatility component. The short-run volatility component can be written, in the spirit of Engle and Lee (1999), as $s = v - x$. Under weak stationarity (discussed in the sequel) we have $\mathbb{E}[v_{t+1}] = \mathbb{E}[x_{t+1}] = m_x \equiv n_x/(1-p_x)$. Thus, m_x is the unconditional mean of x and v, with n_x being the numerator of $n_x/(1-p_x)$, directly proportional to the unconditional mean level. Third, the u_x and u_v processes serve as mean-zero innovations for x and v, respectively, with the coefficients i_x and i_v measuring the strength of the impact of those innovations. The coefficient p_x measures the persistence of x. Analogously, the persistence of v is measured by $b_v = p_v - i_v g_v^2$. Finally, the source of the randomness w is the $(\mathbb{F}, P)$-*Gaussian white noise* with mean 0 and variance 1, hereafter also referred to as the $(\mathbb{F}, P)$-*standard Gaussian white noise.*

3.1. **Non-Negativity of the Conditional Variance.** First, we shall look at the CJOW model and consider the issues regarding the non-negativity of the conditional variance arising in its application.

3.1.1. *CJOW Model.* First, recall that Christoffersen et al. (2008) model can be rewritten in our notation, replacing (3.6) with

$$u_{x,t} = (w_t^2 - 1) - 2g_x\sqrt{v_t}w_t \tag{3.8}$$

and keeping the remaining equations intact.

The problem with this specification is that there is no guarantee on non-negativity of v – and since v is the conditional variance process, we arrive at a possible contradiction. In order to examine the seriousness of the problem, we perform a simulation study and analyze the behavior of the model.

3.1.2. *CJOW Model – Simulation Study.* We perform a simulation study to examine the behavior of this model – performing a grid search with respect to p_x – searching from 0.0 to 1.0 with a step size of 0.001. We do this both for the original CJOW model and a deterministic version thereof (i.e. the one where the driving noise process is assumed to be identically equal to zero instead of a standard GWN), fixing all the other parameter values to those in Table 1 in Christoffersen et al. (2008) (for convenience, we reproduce it in Table 1) – in addition setting r_f to 1.000×10^{-1}. We choose this particular parameter value, since it is the one used by Christoffersen et al. (2008) to differentiate between the Component and the Persistent Component ($p_x = 1$) models. Furthermore, the reason we consider the unit interval as the parameter range is that for $p_x < 0$ non-negativity issues arise immediately (as we shall show later on), while $p_x > 1$ leads to non-stationarity (in particular, the explosiveness of x and, consequently, v). For purposes of this study, $\mathbb{T} = \mathbb{Z} \cap [0, T]$, $T = 1,000$.

T = 1,000	Simulation
r_f	1.000×10^{-1}
λ	$2.092 \times 10^{+0}$
n_x	8.208×10^{-7}
i_v	1.580×10^{-6}
i_x	2.480×10^{-6}
p_v	6.437×10^{-1}
p_x	9.896×10^{-1}
g_v	$4.151 \times 10^{+2}$
g_x	$6.324 \times 10^{+1}$

Table 1. The coefficient values used for the CJOW model simulation study.

We divide the set of p_x coefficient values into *invalid* and *valid* values, where the *invalid* ones are those that lead to negative values of v. We find that the low parameter values are invalid, while the higher ones are valid – the boundary being at approximately 0.9. This means for all $p_x < 0.9$ in our simulation study there exists a $t(p_x) \in \mathbb{T}$ such that $v_{t(p_x)} < 0$. Note, that in practice this leads to $v_{t(p_x)}^{1/2}$ returning NaN[1] for the IEEE 754[2] conforming architecture. Since commonly applied optimization routines will reject arguments leading to NaNs (or terminate with an error, leading to restarting the optimization with different starting values), this potentially explains the estimate of $p_x = 0.9896$ obtained by Christoffersen et al. (2008), which is very close to 1. Hence, due to this numerical property of the model, one cannot necessarily infer "high persistence" to hold in this case. This is because the high estimate might well be a numerical artifact, as opposed to being an empirical property of the data described by the model.

In addition, as we change the sample size T, the boundary value increases as the sample size increases. A possible interpretation of this finding is that as the model runs for a longer time (i.e., as we have more draws in the generated sample) the chance of drawing at least one negative value increases. However, this is not solely due to Gaussianity of w, because we obtain similar result for the deterministic version of the model (i.e. even for a bias forecast) – in fact, the boundary is higher for the deterministic case than the stochastic one.

3.1.3. *CJOW Model – Discussion.* We shall now proceed as follows: assuming the CJOW model, we rewrite (3.2) and (3.3), substituting (3.5) and (3.8), respectively:

$$v_{t+1} = x_{t+1} + p_v(v_t - x_t) + i_v \left((w_t^2 - 1) - 2g_v\sqrt{v_t}w_t\right) \tag{3.9}$$

$$x_{t+1} = n_x + p_x x_t + i_x \left((w_t^2 - 1) - 2g_x\sqrt{v_t}w_t\right) \tag{3.10}$$

where

$$n_x = m_x(1 - p_x). \tag{3.11}$$

Rearranging terms, we obtain

$$\begin{aligned} x_{t+1} &= n_x + p_x x_t - 2i_x g_x\sqrt{v_t}w_t + i_x \left((w_t^2 - 1)\right) && (3.12)\\ &= n_x - i_x + p_x x_t + i_x(w_t - g_x\sqrt{v_t})^2 - i_x g_x^2 v_t. && (3.13)\end{aligned}$$

[1]The term NaN stands for "Not a Number." Here it results from applying the square root function to argument outside its domain, due to attempt to take the square root of a negative number.

[2]IEEE Standard 754 is a floating-point arithmetic standard, the most common floating-point representation of real numbers today on computers – for further reference, see IEEE Task P754 (2008).

Now, assume $p_x > 0$ and $v > 0$. Consider two cases with respect to i_x. If we assume $i_x < 0$, we have, in (3.13), that $-i_x g_x^2 v_t > 0$ and $-i_x > 0$ – this, however, results in $i_x(w_t - g_x\sqrt{v_t})^2 < 0$. On the other hand, if we assume $0 < i_x$ (and we may also want $i_x < n_x < n_x$, so that $n_x - i_x > 0$), then $-i_x g_x^2 v_t < 0$. Hence, we conclude that $P(\exists t \in \mathbb{T} : x_t < 0) > 0$. However, since x is the long-run volatility component, it should remain non-negative over time.

Furthermore, even if we assume $x = 0$, we obtain

$$
\begin{aligned}
v_{t+1} &= p_v v_t + i_v \left((w_t^2 - 1) - 2g_v\sqrt{v_t}w_t\right) && (3.14)\\
&= p_v v_t + i_v \left(w_t^2 - 2g_v\sqrt{v_t}w_t + g_v^2 v_t - g_v^2 v_t - 1\right) && (3.15)\\
&= -i_v + p_v v_t + i_v \left((w_t - g_v\sqrt{v_t})^2 - g_v^2 v_t\right) && (3.16)\\
&= -i_v + b_v v_t + i_v \left(w_t - g_v\sqrt{v_t}\right)^2 && (3.17)
\end{aligned}
$$

where

$$b_v = p_v - i_v g_v^2. \tag{3.18}$$

Now, for all $i_v \neq 0$, $P(\exists t \in \mathbb{T} : v_t < 0) > 0$. In fact, we can obtain the result for an arbitrary $t \in \mathbb{T}$, using (3.17) and the fact that $w \overset{P}{\sim} GWN(0,1)$:

$$
\begin{aligned}
&P(v_{t+1} < 0 | v_t > 0) && (3.19)\\
&= P(-i_v + b_v v_t + i_v \left(w_t - g_v\sqrt{v_t}\right)^2 < 0 | v_t > 0) && (3.20)\\
&= P(\left(w_t - g_v\sqrt{v_t}\right)^2 < (i_v - b_v v_t)/i_v | v_t > 0) && (3.21)\\
&= P(-(i_v - b_v v_t)/i_v < w_t - g_v\sqrt{v_t} < (i_v - b_v v_t)/i_v | v_t > 0) && (3.22)\\
&= P(-\frac{(i_v - b_v v_t)}{i_v} + g_v\sqrt{v_t} < w_t < \frac{(i_v - b_v v_t)}{i_v} + g_v\sqrt{v_t} \,|\, v_t > 0) > 0, && (3.23)
\end{aligned}
$$

as long as the interval $\left(-(i_v - b_v v_t)/i_v + g_v\sqrt{v_t}, (i_v - b_v v_t)/i_v + g_v\sqrt{v_t}\right)$ is non-empty.

An analogous result can be obtained for the x process. However, in order to show the possibility of the negative conditional variance, the existence result is sufficient.

We conclude that assuming a non-zero skewness parameter g_x leads to a model that can result in negative values for the volatility components.

3.1.4. *A Solution.* To mend this problem, we shall now introduce a specification which allows us to derive a sufficient conditions for the volatility components to stay non-negative, given $x_0 > 0$ and $v_0 > 0$. Assume that $g_x = 0$. This eliminates the asymmetry from the long-run component x, and we obtain

$$x_{t+1} = m_x + p_x(x_t - m_x) + i_x u_{x,t} \tag{3.24}$$

with

$$u_{x,t} = (w_t^2 - 1). \tag{3.25}$$

Rearranging (3.24) and substituting (3.25) we obtain

$$\begin{aligned} x_{t+1} &= m_x(1 - p_x) + p_x x_t + i_x(w_t^2 - 1) & (3.26)\\ &= n_x - i_x + p_x x_t + i_x w_t^2 & (3.27) \end{aligned}$$

where

$$n_x = m_x(1 - p_x). \tag{3.28}$$

Note, that (3.27) follows the GMACH(1, 1) model by Yang and Bewley (1995)[3]. Now, in order to obtain non-negative values of x, we need $n_x > i_x > 0$ and $p_x > 0$. Furthermore, under weak stationarity (for which we also need $|p_x| < 1$) we have

$$\mathbb{E}[x_{t+1}] = m_x \equiv \frac{n_x}{1 - p_x}. \tag{3.29}$$

This motivates our previous notation m_x for the unconditional mean of x.

Inserting (3.27) and (3.5) into (3.2) yields:

$$\begin{aligned} v_{t+1} &= x_{t+1} + p_v(v_t - x_t) + i_v u_{v,t} & (3.30)\\ &= n_x - i_x + p_x x_t + i_x w_t^2 \\ &+ p_v(v_t - x_t) + i_v u_{v,t} & (3.31)\\ &= n_x - i_x + p_x x_t + i_x w_t^2 \\ &+ p_v(v_t - x_t) + i_v \left((w_t^2 - 1) - 2g_v\sqrt{v_t}w_t\right). & (3.32) \end{aligned}$$

Rearranging terms and using $p_v = b_v + i_v g_v^2$ we have

$$\begin{aligned} v_{t+1} = n_x &- i_x + p_x x_t + i_x w_t^2 \\ &+ (b_v + i_v g_v^2)(v_t - x_t) + i_v \left(w_t^2 - 1 - 2g_v\sqrt{v_t}w_t\right) & (3.33)\\ &= (n_x - i_x - i_v) + (p_x - b_v - i_v g_v^2)x_t + i_x w_t^2 \\ &+ b_v v_t + i_v \left(w_t^2 - 2g_v\sqrt{v_t}w_t + i_v g_v^2\right) & (3.34)\\ &= (n_x - i_x - i_v) + (p_x - p_v)x_t + b_v v_t \\ &+ i_x w_t^2 + i_v \left(w_t - g_v\sqrt{v_t}\right)^2 & (3.35) \end{aligned}$$

Now, assuming $b_v > 0$, $i_v > 0$ and $i_x > 0$, we have a sufficient condition for non-negativity of v, which is $(n_x - i_x - i_v) + (p_x - p_v)x > 0$. Since we have already established conditions for non-negativity of x, we need to ensure that in addition to them, $(n_x - i_x - i_v) > 0$, $(p_x - p_v) > 0$

[3]This is similar to assuming $c = 0$ in the model by Heston and Nandi (2000) in a way that we also obtain GMACH(1, 1) dynamics.

and $b_v > 0$. Thus, the joint sufficient conditions for non-negativity of the volatility components v and x are as follows:

$$p_x \leq 1, b_v > 0,\ i_v > 0,\ i_x > 0 \tag{3.36}$$
$$n_x > i_x + i_v \tag{3.37}$$
$$p_x > p_v > i_v g_v^2 > 0 \tag{3.38}$$

Restrictions in (3.36) are analogous to those in Engle and Lee (1999). Note, that similarly to Engle and Lee (1999) we also assume that the weak stationarity restriction $p_x < 1$ holds. The economic interpretation of (3.37) is that the mean long-term volatility level has to be sufficiently high relative to the strength of the innovation impact (recall from (3.29) that n_x is the numerator of the unconditional mean, i.e. $m_x \equiv n_x/(1 - p_x)$). The interpretation of (3.38), which can also be stated as $p_x > b_v > 0$, is that the persistence of the long-run component has to be higher than the one of the short-run component and that the impact of the innovation(s) to the short-run component cannot be as strong as to outweigh the persistence.

Hereafter we shall denote our parameter vector by

$$\theta := (r_f,\ \lambda,\ n_x,\ i_v,\ i_x,\ p_v,\ p_x,\ g_v)^\mathsf{T}$$

and the restricted parameter space

$$\tilde{\Theta} := \{\theta \subseteq \Theta :\ (3.36) - (3.38)\},$$

where $\Theta \subseteq \mathbb{R}^p$, $p = 8$.

4. Maximum Likelihood Estimation

We shall now derive a Maximum Likelihood Estimator (MLE)[4] for our model. For notational convenience we assume that the sample includes an observation for $t = 0$. Hereafter we shall assume that non-negativity conditions (3.36)–(3.38) hold.

First, note that by the assumptions (3.1)–(3.7) we have that $w \overset{P}{\sim} GWN(0, 1)$ and $v \in \mathcal{P}$. Using this and (3.1) yields $r_t|\mathcal{F}_{t-1} \overset{P}{\sim} \mathcal{N}(\mu_t, v_t)$. Hence, the conditional probability density function (PDF) of $r_t|\mathcal{F}_{t-1}$ is

$$f(r_t|\mathcal{F}_{t-1}) = \frac{1}{\sqrt{2\pi v_t}} \exp\left(-\frac{(r_t - \mu_t)^2}{2v_t}\right). \tag{4.1}$$

[4]The kind of MLE we derive is called the conditional MLE in Hayashi (2000) – see pp. 547–549.

Using (3.1) again and simplifying we obtain

$$f(r_t|\mathcal{F}_{t-1}) = \frac{1}{\sqrt{2\pi v_t}} \exp\left(-\frac{(\sqrt{v_t}w_t)^2}{2v_t}\right) \tag{4.2}$$

$$= \frac{1}{\sqrt{2\pi v_t}} \exp\left(-\frac{v_t w_t^2}{2v_t}\right) \tag{4.3}$$

$$= \frac{1}{\sqrt{2\pi v_t}} \exp\left(-\frac{w_t^2}{2}\right). \tag{4.4}$$

Now, we can formulate our MLE in terms of an M-estimator. If the PDFs are parametrized by a parameter vector θ then the M-estimator using the log-likelihood of the sample over $t = 0, 1, ..., N$ can be written as:

$$\hat{\theta} = \arg\max_{\theta \in \tilde{\Theta}} Q_n(\theta) \tag{4.5}$$

$$Q_n(\theta) = \frac{1}{N} \sum_{t=0}^{N} \ell_t(\theta) \tag{4.6}$$

$$\ell_t(\theta) = \log f(r_t|\mathcal{F}_{t-1}; \theta) \tag{4.7}$$

$$= -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(v_t) - \frac{1}{2}w_t^2. \tag{4.8}$$

Note, that using proportional and monotonic transformations, we can state our problem for the purposes of minimization as follows:

$$\hat{\theta} = \arg\min_{\theta \in \tilde{\Theta}} \tilde{Q}_n(\theta) \tag{4.9}$$

$$\tilde{Q}_n(\theta) = \sum_{t=0}^{N} l_t(\theta) \tag{4.10}$$

$$l_t(\theta) = \log(v_t) + w_t^2. \tag{4.11}$$

For the numerical details, including objective function computation algorithm and analytical gradient formulas, we refer the reader to Dziubinski (2010).

5. Estimation – Results

Due to the results in Dziubinski (2010) we choose Conditionally-Uniform Feasible Grid Search (CUFGS) with Feasible Sequential Quadratic Programming (FSQP) to estimate the models. The FSQP allows us to solve the constrained optimization problem (4.11), while coupling it with CUFGS enables us to widen the search space and thus increasing the chance of convergence.

We use the S&P 500 index data to calculate the (log) returns. We fit our model to both daily (source: Yahoo Finance, period 1/3/1950–7/22/2009) and high-frequency (5-minute) data (source: Price-Data.com S&P 500, period 4/21/1982–12/6/2007 from Price-Data.com). For the purposes of research reproducibility, we use the same starting values as the ones in the column "Estimation Starting Values" in Table 1 in Dziubinski (2010).[5]

We have considered three methods of obtaining the standard errors – the OPG method, the numerical Hessian, and the sandwich estimator. Since there are numerical issues present when inverting the numerical Hessian (even if it is obtained using analytical first derivatives), we choose to report the OPG standard errors.

As an alternative, one could also use analytical Hessian. In fact, Fiorentini et al. (1996) and Hafner and Herwartz (2008) report, in the context of GARCH estimation, that the analytical Hessian significantly outperforms the approximation. However, in our model, this comes at a cost of calculating $8^2 = 64$ derivatives (or, ensuring that the estimates $\hat{\theta}$ remain in $\tilde{\Theta}$ and using $\tilde{Q}_n \in C^2(\tilde{\Theta})$ with symmetry due to Young's Theorem, $\frac{8(8+1)}{2} = 36$ derivatives). Also, bootstrapping the errors is a possibility.

Note that looking at the estimates for two data sets sampled at different frequencies is a way to empirically investigate temporal aggregation properties of our model. See (Zivot, 2009, Section 3.4) for a discussion of temporal aggregation in a context of GARCH models.

The estimates obtained using the FSQP-AL CUFGS optimization algorithm appear in Table 2. The problems with the large standard errors (causing insignificance) were practically not encountered in case of the FSQP optimization (where we used sandwich estimation and only used OPG or Hessian errors in case of numerical problems; the only problem was with g_v standard error in low-frequency data). This confirms our belief that the choice of the optimization method matters a great deal. Unsurprisingly, as in the similar models in the literature, there are still some issues with estimating λ and g_v. The persistence seems to be slightly lower in case of the low-frequency data (coefficients p_v and p_x) – note, however, that equality of p_v and p_x means that CUFGS yielded p_x to be optimal at the lower corner solution. Furthermore, p_v constitutes the lower bound for p_x generated by CUFGS. This suggests that further research regarding the grid search might be of importance.

[5]Note, that as Zivot (2009) reports, a poor choice of starting values can lead to an ill-behaved log-likelihood and cause convergence problems, which is why we use the starting values that satisfy the non-negativity conditions.

The results for the FSQP-NL CUFGS optimization algorithm in Table 3 are mostly similar to those discussed above. It can be seen (looking at the estimates and associated standard errors) that g_v seems to be estimated more accurately in this case.

	Daily Data	5-minute Data
T	$14,984$	$523,068$
$\hat{r_f}$	5.604×10^{-4} (1.048×10^{-4})	2.685×10^{-10} (1.548×10^{-6})
$\hat{\lambda}$	-7.024×10^{-1} $(1.392 \times 10^{+0})$	$-4.392 \times 10^{+0}$ $(1.475 \times 10^{+0})$
$\hat{n_x}$	4.744×10^{-6} (5.998×10^{-7})	1.047×10^{-7} (8.741×10^{-9})
$\hat{i_v}$	2.320×10^{-6} (4.376×10^{-7})	2.742×10^{-8} (1.785×10^{-9})
$\hat{i_x}$	2.396×10^{-6} (9.433×10^{-7})	7.714×10^{-8} (6.401×10^{-9})
$\hat{p_v}$	9.375×10^{-1} (3.896×10^{-3})	9.157×10^{-1} (6.563×10^{-3})
$\hat{p_x}$	9.375×10^{-1} (7.828×10^{-3})	9.157×10^{-1} (6.652×10^{-3})
$\hat{g_v}$	$2.183 \times 10^{+2}$ (NaN)	$1.595 \times 10^{+1}$ $(1.247 \times 10^{+2})$
$\tilde{Q}_n(\hat{\theta})$	$-1.29423 \times 10^{+5}$	$-6.7172 \times 10^{+6}$

TABLE 2. The estimates (standard errors in parentheses) obtained for the S&P 500 data using FSQP-AL CUFGS optimization algorithm.

	Daily Data	5-minute Data
T	$14,984$	$523,068$
$\hat{r_f}$	7.028×10^{-4} (1.174×10^{-4})	2.906×10^{-29} (1.540×10^{-6})
$\hat{\lambda}$	-4.007×10^{-1} $(1.488 \times 10^{+0})$	$-3.913 \times 10^{+0}$ $(1.465 \times 10^{+0})$
$\hat{n_x}$	4.040×10^{-6} (8.989×10^{-7})	1.049×10^{-7} (2.714×10^{-10})
$\hat{i_v}$	3.474×10^{-6} (1.910×10^{-8})	2.018×10^{-10} (8.008×10^{-5})
$\hat{i_x}$	5.336×10^{-7} (1.097×10^{-7})	1.047×10^{-7} (8.008×10^{-5})
$\hat{p_v}$	9.460×10^{-1} (4.777×10^{-3})	9.155×10^{-1} (1.350×10^{-1})
$\hat{p_x}$	9.460×10^{-1} (4.312×10^{-2})	9.155×10^{-1} (1.784×10^{-4})
$\hat{g_v}$	$1.022 \times 10^{+2}$ $(4.339 \times 10^{+0})$	$5.031 \times 10^{+2}$ $(9.509 \times 10^{+1})$
$\tilde{Q}_n(\hat{\theta})$	$-1.29362 \times 10^{+5}$	$-6.71722 \times 10^{+6}$

TABLE 3. The estimates (standard errors in parentheses) obtained for the S&P 500 data using FSQP-NL CUFGS optimization algorithm.

6. Option Pricing

We shall consider option pricing under our model. In general, there are several approaches to look at:

(1) CJOW risk-neutralization – using the conditional moment generating function(MGF), based on Christoffersen et al. (2008)
(2) Monte-Carlo, Empirical Martingale Simulation (EMS), based on Duan and Simonato (1998)
(3) Monte-Carlo, Empirical Martingale Correction (EMC), based on Chorro et al. (2010)
(4) alternative risk-neutralization method,
(5) different model specification and derive the pertinent non-negativity conditions:
    (a) change the source of randomness w so that it follows a distribution with positive support,
    (b) change the v and x specification, e.g. formulate the equations in log terms in the spirit of an EGARCH model.

The pricing formulas using the analytical methods might be harder to derive (and would often be infeasible in the case of exotic options). The disadvantage of the Monte-Carlo-based pricing methods might be slower performance and, besides, they might need further adjustment to ensure the martingale property, see Duan and Simonato (1998) and Chorro et al. (2010).

6.1. **CJOW-MGF Approach.** As our model is a simplification of the Christoffersen et al. (2008) model we may, in principle, consider using the option-pricing formulas presented there. In practice, however, a difficulty arises in attempts to apply them. In order to perform the option valuation one needs to derive the moment generating function (MGF) for the component GARCH process (provided in Appendix A of Christoffersen et al. (2008)), specify the dynamics under the risk-neutral measure Q (provided in Appendix B of Christoffersen et al. (2008)) and proceed with the option-valuation formula (given in section 4.4 of Christoffersen et al. (2008)). The problem arises in the second step, the risk-neutralization.

Following Christoffersen et al. (2008) we need $\mathbb{E}^Q[exp(r_{t+1})] = exp(r_f)$, which requires that

$$r_{t+1} \equiv \log \frac{S_{t+1}}{S_t} = \mu^Q_{t+1} + \sqrt{v_{t+1}} w^Q_{t+1} \tag{6.1}$$

with

$$\mu^Q_{t+1} = r_f - \frac{1}{2} v_{t+1}. \tag{6.2}$$

This in turn implies that

$$w^Q_{t+1} = w_{t+1} + (\lambda + \frac{1}{2})\sqrt{v_{t+1}}. \tag{6.3}$$

We also want to ensure the equality of the conditional variances under the two measures:

$$\mathbb{V}^P[r_{t+1}|\mathcal{F}_t] = \mathbb{V}^Q[r_{t+1}|\mathcal{F}_t]. \tag{6.4}$$

We therefore need to have equal variance innovations under the two measures – that is

$$(w_t - g_i\sqrt{v_t})^2 = (w_t^Q - g_i^Q\sqrt{v_t})^2, \quad i = v, x. \tag{6.5}$$

This can be achieved by defining the risk-neutral parameters:

$$g_i^Q = g_i + \lambda + \frac{1}{2}, \quad i = v, x. \tag{6.6}$$

Now, the problem is that in our specification we have

$$g_x = 0. \tag{6.7}$$

Thus

$$g_x^Q = \lambda + \frac{1}{2}. \tag{6.8}$$

This means

$$g_x^Q = 0 \iff \lambda = -\frac{1}{2}. \tag{6.9}$$

Hence, without restricting the market price of risk λ to a value which is not particularly realistic, we cannot ensure that the Q-dynamic is going to remain such that we stay within our class of models (where we can apply the sufficient conditions for non-negativity of the conditional variance).

6.2. **Empirical Martingale Simulation (EMS).** EMS is a variance-reduction method ensuring the martingale property to be used with Monte Carlo pricing. The problem from our point of view is, however, that the EMS relies on the formulation of the model under the Q measure – that is, a prior risk-neutralization. However, our model (similarly to CJOW) is stated under the P measure, so analytical risk-neutralization would be required. But then, the one available method (CJOW, discussed above) is not applicable if we want to stay within our class of models. This excludes the EMS from any further considerations.

6.3. **Empirical Martingale Correction (EMC).** The Empirical Martingale Correction method is, in fact, inspired by the EMS, see Chorro et al. (2010). The fundamental difference is that it is applicable to the models stated under the P measure, such as ours. In this method, we make no assumption on the risk-neutralization (i.e. the shape of the pricing kernel, involving Radon-Nikodym derivative $\frac{dQ}{dP}$), and we compute prices for options with time to maturity $(T-t)$ by simulating sampled paths of the stochastic model under the historical measure P.

To rule out arbitrage opportunities, we directly impose risk neutrality constraints. The i^{th} sampled historical final price for the underlying is denoted by $S_{T,i}$.

The Empirical Martingale Correction works such that the previously sampled prices are replaced by:

$$\tilde{S}_{T,i} = \frac{S_{T,i}}{\frac{1}{N}\sum_{i=1}^{N} S_{T,i}} S_t e^{r(T-t)}. \tag{6.10}$$

The sampled average of $\tilde{S}_{T,i}$ is exactly equal to $S_t e^{r(T-t)}$, that is, the risk neutral conditional expectation. With this approach, we only shift the historical distribution in a way that prevents arbitrage opportunities by implicitly changing the drift of this distribution. Chorro et al. (2010) compare this approach with the affine Stochastic Discount Factor (SDF) methodology in Cochrane (2002) and find that the prices obtained by these two methods are close to each other.

6.4. **Option Pricing Results.** Finally, we compare the option pricing results in our model with those in the Black-Scholes-Merton model and Heston-Nandi GARCH(1, 1) model (HN). In this section we use daily data only. For comparison, we consider option pricing under the SCGARCH model using the estimates obtained using FSQP-AL and FSQP-NL (applying CUFGS in both cases). We estimate the HN model using fOptions R package – for details, see Wuertz (2007). The estimation results are shown in Table 4.

N = 14,984	Daily Data
λ	$3.451 \times 10^{+0}$
ω	1.139×10^{-281}
α	3.671×10^{-6}
β	9.005×10^{-1}
γ	$1.196 \times 10^{+2}$
Log-Likelihood	88559.65
Persistence	0.953
Variance	7.806764×10^{-5}

TABLE 4. The estimates (standard errors in parentheses) obtained for the S&P 500 daily data: the HN model.

Next, we present the pricing results across moneyness and maturity in Tables 5–6 and 7–8. The prices seem reasonable. Depending on whether the FSQP-AL or FSQP-NL optimization was used, they either fall between HN and BSM or further correct in-the-money (ITM)/out-of-the-money (OTM) and time-to-maturity mispricing effects relative to BSM model. Considering the volatility components in the context of GARCH option pricing thus yields a viable approach of practical interest.

$T = 0.5$	SCGARCH (AL)	SCGARCH (NL)
$S = 80$	3.441502×10^{-2}	7.102035×10^{-2}
$S = 100$	$5.710897 \times 10^{+0}$	$5.853388 \times 10^{+0}$
$S = 120$	$2.343247 \times 10^{+1}$	$2.343515 \times 10^{+1}$

TABLE 5. The option prices ($), strike $K = 100$, maturity $T = 0.5$ years $= 0.5 \times 252$ trading days.

$T = 0.5$	HN	BSM
$S = 80$	4.248666×10^{-2}	9.48591×10^{-2}
$S = 100$	$5.981560 \times 10^{+0}$	$5.79014 \times 10^{+0}$
$S = 120$	$2.352170 \times 10^{+1}$	$2.340176 \times 10^{+1}$

TABLE 6. The option prices ($), strike $K = 100$, maturity $T = 0.5$ years $= 0.5 \times 252$ trading days.

$T = 1.0$	SCGARCH (AL)	SCGARCH (NL)
$S = 80$	6.200905×10^{-1}	6.764589×10^{-1}
$S = 100$	$9.260039 \times 10^{+0}$	$9.301633 \times 10^{+0}$
$S = 120$	$2.679264 \times 10^{+1}$	$2.690852 \times 10^{+1}$

TABLE 7. The option prices ($), strike $K = 100$, maturity $T = 1.0$ years $= 252$ trading days.

$T = 1.0$	HN	BSM
$S = 80$	7.172656×10^{-1}	8.225735×10^{-1}
$S = 100$	$9.608113 \times 10^{+0}$	$9.322803 \times 10^{+0}$
$S = 120$	$2.701091 \times 10^{+1}$	$2.680135 \times 10^{+1}$

TABLE 8. The option prices ($), strike $K = 100$, maturity $T = 1.0$ years $= 252$ trading days.

One could add that another empirically interesting exercise would be to use the option prices in estimation, similarly to what Christoffersen et al. (2008) suggested. Note, however, that in our model we do not use analytical option pricing formulas, but instead apply a Monte-Carlo method. On a single-core CPU (central processing unit) this method is too slow to be used in this application. A very promising approach, however, would be to use parallelized many-core GPU (graphics processing unit) computation – since MC is a so-called *embarrassingly parallel problem*, that would yield very significant performance improvements. In particular, in an application of MC pricing involving path-dependent options, Joshi (2010) demonstrates that it is possible to get accuracy of 2×10^{-4} in less than a fiftieth of a second, concluding that "GPU technology has rendered the Monte Carlo pricing of Asian options sufficiently fast that there is no longer any need for analytic approximations." This approach would also make possible to investigate forecasting properties of the model.

7. Conclusions

This paper presents a discrete-time volatility model in which the underlying follows a process with conditional variance driven by the new Simplified Component GARCH process. It is a more parsimonious model than the CJOW one and allows us to derive sufficient conditions for non-negativity of the conditional variance.

Maximum likelihood estimation of the model is discussed.

We provide an empirical illustration, applying the model to the S&P 500 index data. The results are consistent with our economic intuition.

We propose an option pricing method consistent with our model

The performance of the pricing method across moneyness and maturity is compared with that of the Heston-Nandi GARCH and Black-Scholes-Merton models. The results of the comparison are favorable to our model.

Several of the future research directions and possible extensions to this work are worth consideration – regarding to the advanced grid-generation techniques and optimization algorithms and applications of GPUs allowing for more advanced pricing and forecasting applications. We provide a number of approaches to achieve that in the respective sections of this paper.

References

Brockwell, P. J. and R. A. Davis (1991): *Time Series: Theory and Methods*, New York: Springer Verlag, 2nd ed.

Chorro, C., D. Guegan, and F. Ielpo (2010): "Martingalized historical approach for option pricing," *Finance Research Letters*, 7, 24–28.

Christoffersen, P., K. Jacobs, C. Ornthanalai, and Y. Wang (2008): "Option valuation with long-run and short-run volatility components," *Journal of Financial Economics*, 90, 272–297.

Cochrane, J. (2002): *Asset Pricing*, Princeton University Press.

Duan, J.-C. and J.-G. Simonato (1998): "Empirical Martingale Simulation for Asset Prices," *Management Science*, 44, 1218–1233.

Dziubinski, M. P. (2010): "Conditionally-Uniform Feasible Grid Search (CUFGS)," *CREATES Working Paper*.

Engle, R. F. and G. G. J. Lee (1999): "A Permanent and Transitory Component Model of Stock Return Volatility," in *Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W.J. Granger*, University Press, 475–497.

Fiorentini, G., G. Calzolari, and L. Panattoni (1996): "Analytic Derivatives and the Computation of GARCH Estimates," *Journal of Applied Econometrics*, 11, 399–417.

Hafner, C. and H. Herwartz (2008): "Analytical quasi maximum likelihood inference in multivariate volatility models," *Metrika*, 67, 219–239.

Hayashi, F. (2000): *Econometrics*, Princeton University Press.

Heston, S. and S. Nandi (2000): "A closed-form GARCH option valuation model," *Review of Financial Studies*, 13, 585–625.

IEEE Task P754 (2008): *IEEE 754-2008, Standard for Floating-Point Arithmetic*.

Joshi, M. (2010): "Graphical Asian Options," *Wilmott Journal*, 2, 97–107.

Protter, P. E. (2005): *Stochastic Integration and Differential Equations*, Springer, 2nd ed.

Wuertz, D. (2007): "fOptions: Financial Software Collection-fOptions. R package version 260.72," in *Rmetrics*.

Yang, M. and R. Bewley (1995): "Moving average conditional heteroskedastic processes," *Economics Letters*, 49, 367 – 372.

Zivot, E. (2009): "Practical Issues in the Analysis of Univariate GARCH Models," in *Handbook of Financial Time Series*, Springer, 113–155.

(Matt P. Dziubinski) CREATES
School of Economics and Management
University of Aarhus
Building 1322, DK-8000 Aarhus C
Denmark

E-mail address, Matt P. Dziubinski: matt@creates.au.dk

Research Papers 2011

CREATES
Center for Research in Econometric Analysis of Time Series

2010-71:	Nektarios Aslanidis and Isabel Casas : Modelling asset correlations during the recent financial crisis: A semiparametric approach
2010-72:	Søren Johansen and Katarina Juselius: An invariance property of the common trends under linear transformations of the data
2010-73:	Peter Sandholt Jensen and Allan H. Würtz: Estimating the effect of a variable in a high-dimensional regression model
2010-74:	Peter R. Hansen, Asger Lunde and Valeri Voev: Realized Beta GARCH: A Multivariate GARCH Model with Realized Measures of Volatility and CoVolatility
2010-75:	Laurent A.F. Callot: A Bootstrap Cointegration Rank Test for Panels of VAR Models
2010-76:	Peter R. Hansen, Asger Lunde and James M. Nason: The Model Confidence Set
2011-01:	Cristina Amado and Timo Teräsvirta: Modelling Volatility by Variance Decomposition
2011-02:	Timo Teräsvirta: Nonlinear models for autoregressive conditional heteroskedasticity
2011-03:	Roxana Halbleib and Valeri Voev: Forecasting Covariance Matrices: A Mixed Frequency Approach
2011-04:	Mark Podolskij and Mathieu Rosenbaum: Testing the local volatility assumption: a statistical approach
2011-05:	Michael Sørensen: Prediction-based estimating functions: review and new developments
2011-06:	Søren Johansen: An extension of cointegration to fractional autoregressive processes
2011-07:	Tom Engsted and Stig V. Møller: Cross-sectional consumption-based asset pricing: The importance of consumption timing and the inclusion of severe crises
2011-08:	Tommaso Proietti and Stefano Grassi: Bayesian stochastic model specification search for seasonal and calendar effects
2011-09:	Matt P. Dziubinski: Option valuation with the simplified component GARCH model