Working Paper

Non-Extensitivity versus informative moments for financial models: a unifying framework and empirical results

IWQW Discussion Papers, No. 07/2009

Provided in Cooperation with:
Friedrich-Alexander University Erlangen-Nuremberg, Institute for Economics

Suggested Citation: Herrmann, Klaus (2009) : Non-Extensitivity versus informative moments for financial models: a unifying framework and empirical results, IWQW Discussion Papers, No. 07/2009, Friedrich-Alexander-Universität Erlangen-Nürnberg, Institut für Wirtschaftspolitik und Quantitative Wirtschaftsforschung (IWQW), Erlangen

This Version is available at:
http://hdl.handle.net/10419/29559

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Non-Extensity versus Informative Moments for Financial Models - A Unifying Framework and Empirical Results

Klaus Herrmann
University of Nuremberg

ISSN 1867-6707
Non-Extensitivity versus Informative Moments for Financial Models - A Unifying Framework and Empirical Results

KLAUS HERRMANN

Department of Statistics and Econometrics
University of Erlangen-Nürnberg, Germany
Email: Klaus.Herrmann@wiso.uni-erlangen.de

June 19, 2009

SUMMARY
Information-theoretic approaches still play a minor role in financial market analysis. Nonetheless, there have been two very similar approaches evolving during the last years, one in so-called econophysics and the other in econometrics. Both generalize the notion of GARCH processes in an information-theoretic sense and are able to capture skewness and kurtosis better than traditional models. In this article we present both approaches in a more general framework and compare their performance in some illustrative data sets.

Keywords and phrases: Entropy density; Skewness; Kurtosis, GARCH

JEL C22
PACS 05.45.Tp; 05.90.+m; 89.65.Gh

1 Introduction

Though information-theoretic approaches still play only a minor role in financial market analysis, there have been two recent approaches developing independently one in econophysics and the other in econometrics. Both approaches generalize the notion of either ARCH or GARCH models.

The econometric models for financial markets analysis have been introduced by Rockinger/Jondeau (2002) and been extended and generalized in Bera/Park (2009) and Fischer/Herrmann (2008). These approaches model variance’s motion in time, like traditional GARCH models, but include knowledge of higher moments, such as skewness and kurtosis, in order to capture known stylized facts.

The models in econophysics have been developed mainly in Queirós/Tsallis (2005), Borland (2005) and Queirós (2007). The econophysics approaches model only variance, but derive corresponding maximum entropy distributions by maximizing Tsallis q-entropy

\[ H_T = \int_{D(X)} \frac{1 - f(X)^q}{q - 1} dx, \]

where \( X \) is some random variable, \( D(X) \) its support and \( f(x) \) its density function, instead of the Shannon or Boltzmann-Gibbs entropy for given constraints. The resulting distribution
given only variance constraints is called q-Gaussian and its functional form is equivalent to a generalized Student-t distribution. A deviation of the value found for q from 1, the case when the normal distribution would be recovered, is called non-extensivity. Both approaches have been successfully applied to financial market data as both generalizations allow to capture kurtosis in given variance models. Following Kesavan/Kapur (1989) we give a more general framework for information theoretic models for time-varying moments and derive GARCH models as well as the above mentioned non-extensive approaches as special cases. We extend the existing models and in an application to financial market data we compare these model’s capability to capture financial returns behavior.

2 Models for Time-Varying Moments using the Generalized Principle of Maximum Entropy

In this note we will only consider time-series models that assume the conditional distribution of some random variable $X$ at time $t$ given the information set available at time $t-1$ to be the distribution maximizing some generalized entropy measure as

$$H(f) = -\int_{D(X)} \phi(f(x))dx,$$

for some convex function $\phi$ and for $D(X)$ the random variables support, subject to constraints of $k+1$ conditional expectation values of suitable functions $g_i(\cdot)$ as

$$E(g_0(X)) = E(1) = 1, \quad E(g_i(X)|\mathcal{F}_{t-1}) = m_i(\mathcal{F}_{t-1}), \quad \forall i = 1, \ldots, k,$$

where $a_i$ may be some deterministic functions depending only on the information set available.

The information theoretic interpretation of such models is, that we model only some expectation values’ motion in time and for all information missing to completely determine the corresponding density functions, we maximize entropy. In the case of the Shannon entropy measure such a way of modeling is nothing else but the consequent application of Jaynes (1957) principle of maximum entropy. Using generalized entropy measures is justified by Kesavan/Kapur (1989)’s generalized maximum entropy principle.

Following their suggestion, we restrict $\phi$ to the set of differentiable convex functions. A variational approach shows that, under some weak conditions for $\phi$, for the maximum entropy density (if it exists) holds that,

$$\phi'(f(x)) = \sum_{i=0}^{k} \lambda_i g_i(x),$$

(2.1)

$^1$Here we follow the suggestion of Kapur/Kesavan (1989).
where the $\lambda_i$ have to be chosen such that the constraints are fulfilled.\(^2\) If a solution exists for $f$, the $\lambda_i$ can be found by numerically minimizing a concave function of some dual problem. Some generalized entropy measures and corresponding dual problems are given e.g. in Kapur (1994). Table (1) gives some examples:

<table>
<thead>
<tr>
<th>Entropy</th>
<th>$\phi(t)$</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burg</td>
<td>$\ln(t)$</td>
<td>-</td>
</tr>
<tr>
<td>Kapur</td>
<td>$t \cdot \ln(t) + \frac{1}{c}(t + c)\ln(t + c)$</td>
<td>$c &gt; 0$</td>
</tr>
<tr>
<td>Shannon</td>
<td>$t \cdot \ln(t)$</td>
<td>-</td>
</tr>
<tr>
<td>Tsallis</td>
<td>$\frac{1-t^{q}}{1-q}$</td>
<td>$q &gt; 0$, $q \neq 1$</td>
</tr>
</tbody>
</table>

Table 1: Some generalized entropy measures.

A very efficient numerical algorithm for their implementation is given by Rockinger/Jondeau (2002) for the Shannon entropy, but can easily be adjusted to other dual problems.

3 GARCH Models as Models for Time-Varying Moments

Bollerslev (1986)’s original GARCH(p,q) model may be given as

$$x_t|_{\mathcal{F}_{t-1}} = z_t \cdot \sigma_t, \quad z_t \overset{iid}{\sim} \mathcal{N}(0,1),$$

$$\sigma_t = \alpha_0 + \sum_{i=1}^{p} \alpha_{1,i}x_{t-i}^2 + \sum_{i=1}^{q} \alpha_{2,i}\sigma_{t-i}^2,$$

where $x_t$ is the random variable at time $t$ and $\mathcal{F}_{t-1}$ the information set available in the period before.

We can present the same model as a model for time-varying moments as defined above, if we derive the conditional densities of $x_t$ given the information available at $t - 1$, by maximizing Shannon’s entropy measure

$$H_S(f) = \int_{\mathcal{D}(X)} f(x)\ln(f(x))dx$$

subject to the constraints

$$E(1|\mathcal{F}_{t-1}) = 1, \quad E(X|\mathcal{F}_{t-1}) = \mu, \quad E((X-\mu)^2|\mathcal{F}_{t-1}) = \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_{1,i}x_{t-i}^2 + \sum_{i=1}^{q} \alpha_{2,i}\sigma_{t-i}^2.$$

We can derive the Shannon entropy as a generalized measure of entropy, if we set $\phi(t) = t \cdot \ln(t)$. Because

$$\phi'(t) = \frac{\delta(t \cdot \ln(t))}{\delta t} = 1 + \ln(t),$$

\(^2\)Compare e.g. Kapur (1994).
we derive the functional form of the conditional distribution using equation (2.1) as

\[ f(x|\mathbf{\bar{y}}_{t-1}) = \exp(-1 + \lambda_0 + \lambda_1 x + \lambda_2 x^2), \]

where the \( \lambda_i \) are functions of the conditional moment constraints. As we derive \( \lambda \) to meet the above constraints, this distribution is nothing else but the normal distribution with mean \( \mu \) and variance \( \sigma_t^2 \), just as in Bollerslev (1986)'s model.

It is a well known fact, that GARCH-models assuming gaussian innovations do not sufficiently describe financial markets data. Apart from a vast literature concerned with that topic\(^3\), this can be seen by looking at the empirical distribution of GARCH-filtered innovations. After estimating \( \alpha_0, \alpha_{1,1}, ..., \alpha_{1,p}, \alpha_{2,1}, ..., \alpha_{2,q} \), these can be obtained by rewriting GARCH-models in the original notation of Bollerslev (1986), as \( \hat{z}_t = \frac{z_t}{\sigma_t} \).

If the assumption of gaussian innovations would hold, the empirical distribution of \( \hat{z}_t \) should be close to a normal distribution. But having a look at some illustrative data sets\(^4\), we find some evidence, that this assumption does not hold, see figure (1).

![Figure 1: GARCH(1,1) innovations' empirical densities versus normal density (dotted line) and results for tests on normality.](image)

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Gold</th>
<th>p-Value</th>
<th>DJIA</th>
<th>p-Value</th>
<th>EurUS</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness</td>
<td>-0.0309</td>
<td>0.0503</td>
<td>-0.0404</td>
<td>0.2504</td>
<td>0.0652</td>
<td>0.0640</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.8532</td>
<td>3.7080</td>
<td>3.8534</td>
<td>3.4602</td>
<td>3.6666</td>
<td>0.0652</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>49.514</td>
<td>1.771e-11</td>
<td>35.124</td>
<td>3.064e-08</td>
<td>50.514</td>
<td>1.074e-11</td>
</tr>
<tr>
<td>( \chi^2 )</td>
<td>11.382</td>
<td>0.2504</td>
<td>34.602</td>
<td>5.6666</td>
<td>16.084</td>
<td>0.0652</td>
</tr>
</tbody>
</table>

Figure 1: GARCH(1,1) innovations’ empirical densities versus normal density (dotted line) and results for tests on normality.

It is generally assumed that financial market GARCH innovations deviate from the normal distribution because of skewness and higher kurtosis. Traditional approaches try to overcome that problem by assuming \( z_t \) to follow some parametric distribution that is flexible enough

---

\(^3\)See e.g. Hansen (1994) for an overview.

\(^4\)See section (6.2) for a more detailed description of the data sets.
to model skewness and kurtosis, such as e.g. the SGT2, EGB2 or the family of stable distributions.\footnote{Compare e.g. Theodossiou (1998), MacDonald (1991) and Rachev/Mittnik (2000).}

This note, of course, is devoted to information theoretic approaches that directly model such features by, e.g. including measures of skewness and kurtosis in the entropy maximization task or, as the non-extensive approaches do, by choosing suitable entropy measures to allow for higher kurtosis.

4 Higher Informative Moments

To the best of our knowledge, the first approach of including knowledge of higher moments in an information theoretic framework was given by Rockinger/Jondeau (2002). In their approach the use third and fourth power moments to measure skewness and kurtosis and use Shannon’s entropy as information measure. But as Fischer/Herrmann (2008) point out, these measures do not allow the derivation of proper maximum Shannon entropy densities for kurtosis values higher than implied by the lower moments. In order to overcome that problem, one has to use measures that are defined as expectation values of functions that grow slower in $x$ than $x^2$.\footnote{Some of the generalized parametric approaches may also find some representation in the above framework - but because of the vast supply of such suggestions, we will restrict this note to the most basic case.} Suggestions for such measures as $m_i = E(g_i(X)|\hat{X}_{t-1})$ can be found in Bera/Park (2009) and Fischer/Herrmann (2008), as e.g.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|}
\hline
$m_3$ & $m_4$ \\
\hline
$E(\tan^{-1}(X))$ & $E(\tan^{-1}(X)^2)$ \\
$E\left(\frac{X}{1+X^2}\right)$ & $E(\log(1 + X^2))$ \\
\hline
\end{tabular}
\caption{Some suggestions for measures of asymmetry $m_3$ and measures of kurtosis $m_4$ as proposed in recent literature.}
\end{table}

where $m_3$ denotes measures of asymmetry and $m_4$ denotes measures of kurtosis.

Following Fischer/Herrmann (2008) and Bera/Park (2009), we can extend Bollerslev GARCH model as model for time-varying moments, assuming for the conditional distributions of $X_t$ given the information set available at $t-1$ the distribution maximizing the Shannon entropy subject to

$$E(X|\hat{X}_{t-1}) = \mu, \quad E((X - \mu)^2|\hat{X}_{t-1}) = \sigma_t,$$

$$E\left(g_3\left(\frac{X - \mu}{\sigma_t}\right)|\hat{X}_{t-1}\right) = m_3, \quad E\left(g_4\left(\frac{X - \mu}{\sigma_t}\right)|\hat{X}_{t-1}\right) = m_4,$$

where we let $\sigma_t$ vary over time as in the original GARCH model, see equation (3.1), while we assume all other moment functions to be constant over time.\footnote{Nonetheless there is of course weak evidence of time-variability of higher moments, compare Fischer/Herrmann (2008) or e.g. Dubauskas/Teresiene (2005).}
5 Non-Extensive Approaches

Information theoretic models in econophysics have been developed mainly in Queirós/Tsallis (2005), Borland (2005) and Queirós (2007). Their approaches generalize GARCH models in the above discussed form by assuming the conditional density functions to maximize the generalized entropy measure suggested by Tsallis (1988)\(^9\)

\[
H_T = \int_{D(X)} \frac{1 - f(X)^q}{q - 1} \, dx, \quad \text{s.t.} \quad E(X|\tilde{y}_{t-1}) = \mu, \quad E\left(\left(X - \mu\right)^2|\tilde{y}_{t-1}\right) = \sigma_t, \quad (5.1)
\]

with different suggestions for \(\sigma_t\)’s motion in time.

Contrary to the econometric approaches these models capture kurtosis not as information explicitly included, but by flexibly varying \(q\). As the resulting conditional distributions of these models are

\[
f_{ME,T} = \left(\frac{q - 1}{q} \sum_{i=0}^{k} \lambda_i g_i(x)\right)^{\frac{1}{q-1}} = \left(\frac{q - 1}{q} \left(\lambda_0 + \lambda_1 x + \lambda_2 (x - \mu)^2\right)\right)^{\frac{1}{q-1}},
\]

they can be interpreted as some generalized t-distribution, where \(q\) is a parameter that drives the degrees of freedom, \(\nu\). For fixed mean and variance a finite \(\nu\) corresponds to a distribution with higher kurtosis than the normal distribution. Because distributions with a value for \(q\) different than 1 are called non-extensive\(^{10}\), we refer to these approaches as non-extensive approaches. As the degree of this non-extensitivity increases with the kurtosis implied, these models capture kurtosis by their non-extensitivity.

6 Application to Financial Market Data

We compare three models, equivalent in their flexibility, by applying them to three time series typical for financial market data from three different markets. We will compare these models by likelihood and likelihood based goodness-of-fit measures, by the distance of their empirical innovations’ distribution to the theoretical model and by their capability to explain their empirical quantiles in the tails.

6.1 Models

For the class of the econometric models (ECO) we follow a suggestion by Bera/Park (2009) and maximize the Shannon entropy subject to

\[
E(X|\tilde{y}_{t-1}) = \mu, \quad E\left(\left(X - \mu\right)^2|\tilde{y}_{t-1}\right) = \sigma_t, \quad (6.1)
\]

\(^9\)Tsallis suggestion is also known as the entropy of Havrda/Charvat.

\(^{10}\)Compare e.g. Tsallis (1988).
and include knowledge of higher moments using

\[ E\left(\tan^{-1}\left(\frac{X - \mu}{\sigma_t}\right) | \tilde{F}_{t-1}\right) = m_3, \quad E\left(ln\left(1 + \left(\frac{X - \mu}{\sigma_t}\right)^2\right) | \tilde{F}_{t-1}\right) = m_4, \quad (6.2) \]

as in some previous study models using these moments performed best.

For the class of non-extensive models (NEX) we maximize the Tsallis entropy given in equation (5.1) subject to equation (6.1) and, contrary to the proposal in econophysics, additionally introduce the skewness moment of equation (6.2). The later is included to make sure, that the non-extensive models exhibit the same flexibility as the econometric models in terms of skewness.

In order to show the flexibility of the proposed generalized framework, we suggest a third model (KAP), where we maximize the entropy measure suggested by Kapur (1994), subject to equation (6.1) and the skewness moment of equation (6.2), where the flexibility in terms of kurtosis shall be introduced by flexibly varying \(c\).

For all models we assume the time-model for variance as in a traditional GARCH(1,1)\(^{11}\) as

\[ \sigma_t = \alpha_0 + \alpha_1 x_{t-1}^2 + \alpha_2 \sigma_{t-i}^2 \]

and assume all higher moments or parameters to be constant over time.

6.2 Data

In order to sample different kinds of financial market indices, we chose the daily returns between January 1st 2003 and March 20th 2009 for the gold price, the Dow Jones Industrial Average, both from yahoo.finance.com, and for the euro-US-dollar exchange rate from www.ecb.int.

Some descriptive statistics for the data are given in table (3):\(^{12}\)

<table>
<thead>
<tr>
<th></th>
<th>Gold</th>
<th>DJIA</th>
<th>EurUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-1.363e-03</td>
<td>2.316e-05</td>
<td>-1.657e-04</td>
</tr>
<tr>
<td>Stan.Dev.</td>
<td>3.479e-02</td>
<td>1.286e-02</td>
<td>6.499e-03</td>
</tr>
<tr>
<td>(\hat{m}_3)</td>
<td>-1.596e-01</td>
<td>-5.802e-02</td>
<td>7.545e-02</td>
</tr>
<tr>
<td>(\hat{m}_4)</td>
<td>6.961</td>
<td>14.301</td>
<td>7.254</td>
</tr>
<tr>
<td>Observations</td>
<td>1606</td>
<td>1603</td>
<td>1632</td>
</tr>
</tbody>
</table>

Table 3: Some descriptive statistics for the illustrative data sets.

\(^{11}\)As there is empirical evidence that \(p = 1\) and \(q = 1\) sufficiently describe financial returns behavior, compare e.g. Bera/Higgins (1993).

\(^{12}\)\(m_3\) (\(m_4\)) denotes the third (fourth) standardized power moment.
6.3 Empirical Results

We use numerical optimization routines to implement a maximum likelihood estimation of the model parameters. Estimates, log-likelihood and standard errors (in brackets) are given in figure (2):

<table>
<thead>
<tr>
<th>Model</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$\beta_0$</th>
<th>$\gamma_0$</th>
<th>LogL</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH</td>
<td>$7.2937e-06$</td>
<td>$0.04064$</td>
<td>$0.95317$</td>
<td>-</td>
<td>-</td>
<td>3268.252</td>
</tr>
<tr>
<td></td>
<td>(3.5099e-06)</td>
<td>(0.00753)</td>
<td>(0.00832)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECO</td>
<td>$7.6282e-06$</td>
<td>$0.04101$</td>
<td>$0.95246$</td>
<td>$0.00409$</td>
<td>$0.50789$</td>
<td>3283.845</td>
</tr>
<tr>
<td></td>
<td>(4.1548e-06)</td>
<td>(0.00874)</td>
<td>(0.00991)</td>
<td>(0.00393)</td>
<td>(0.00534)</td>
<td></td>
</tr>
<tr>
<td>NEX</td>
<td>$7.028e-06$</td>
<td>$0.03936$</td>
<td>$0.95468$</td>
<td>$0.00537$</td>
<td>$0.80672$</td>
<td>3284.064</td>
</tr>
<tr>
<td></td>
<td>(3.9615e-06)</td>
<td>(0.00824)</td>
<td>(0.00922)</td>
<td>(0.00385)</td>
<td>(0.03794)</td>
<td></td>
</tr>
<tr>
<td>KAP</td>
<td>$7.6789e-06$</td>
<td>$0.04158$</td>
<td>$0.95197$</td>
<td>$0.0052$</td>
<td>$3.88308$</td>
<td>3283.851</td>
</tr>
<tr>
<td></td>
<td>(4.1579e-06)</td>
<td>(0.00886)</td>
<td>(0.01)</td>
<td>(0.00375)</td>
<td>(1.01096)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$\beta_0$</th>
<th>$\gamma_0$</th>
<th>LogL</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH</td>
<td>$7.0443e-07$</td>
<td>$0.07095$</td>
<td>$0.92296$</td>
<td>-</td>
<td>-</td>
<td>5241.281</td>
</tr>
<tr>
<td></td>
<td>(2.7215e-07)</td>
<td>(0.01010)</td>
<td>(0.01025)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECO</td>
<td>$7.0784e-07$</td>
<td>$0.07065$</td>
<td>$0.92299$</td>
<td>$-0.00004$</td>
<td>$0.50851$</td>
<td>5255.646</td>
</tr>
<tr>
<td></td>
<td>(3.1362e-07)</td>
<td>(0.01147)</td>
<td>(0.01158)</td>
<td>(0.00399)</td>
<td>(0.00531)</td>
<td></td>
</tr>
<tr>
<td>NEX</td>
<td>$7.3838e-07$</td>
<td>$0.07024$</td>
<td>$0.92337$</td>
<td>$-0.00114$</td>
<td>$0.81873$</td>
<td>5252.580</td>
</tr>
<tr>
<td></td>
<td>(3.1597e-07)</td>
<td>(0.01143)</td>
<td>(0.01149)</td>
<td>(0.00393)</td>
<td>(0.04088)</td>
<td></td>
</tr>
<tr>
<td>KAP</td>
<td>$7.3096e-07$</td>
<td>$0.07018$</td>
<td>$0.92328$</td>
<td>$-0.00284$</td>
<td>$4.0709$</td>
<td>5255.698</td>
</tr>
<tr>
<td></td>
<td>(3.1737e-07)</td>
<td>(0.01148)</td>
<td>(0.01163)</td>
<td>(0.00384)</td>
<td>(1.08256)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$\beta_0$</th>
<th>$\gamma_0$</th>
<th>LogL</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH</td>
<td>$1.1797e-07$</td>
<td>$0.03380$</td>
<td>$0.96342$</td>
<td>-</td>
<td>-</td>
<td>6062.748</td>
</tr>
<tr>
<td></td>
<td>(7.5857e-08)</td>
<td>(0.00584)</td>
<td>(0.00592)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECO</td>
<td>$1.4706e-07$</td>
<td>$0.03371$</td>
<td>$0.96254$</td>
<td>$-0.00144$</td>
<td>$0.50931$</td>
<td>6076.519</td>
</tr>
<tr>
<td></td>
<td>(9.2357e-08)</td>
<td>(0.00675)</td>
<td>(0.00702)</td>
<td>(0.00386)</td>
<td>(0.00524)</td>
<td></td>
</tr>
<tr>
<td>NEX</td>
<td>$1.6122e-07$</td>
<td>$0.03127$</td>
<td>$0.96441$</td>
<td>$-0.00272$</td>
<td>$0.81997$</td>
<td>6076.861</td>
</tr>
<tr>
<td></td>
<td>(9.0189e-08)</td>
<td>(0.00638)</td>
<td>(0.00674)</td>
<td>(0.00379)</td>
<td>(0.03741)</td>
<td></td>
</tr>
<tr>
<td>KAP</td>
<td>$1.4455e-07$</td>
<td>$0.03434$</td>
<td>$0.96211$</td>
<td>$-0.00236$</td>
<td>$3.63563$</td>
<td>6075.98</td>
</tr>
<tr>
<td></td>
<td>(9.2727e-08)</td>
<td>(0.00682)</td>
<td>(0.00706)</td>
<td>(0.00369)</td>
<td>(0.99112)</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2: Empirical results for our models for some illustrative data sets.

We find for all models similar estimates for variance motion in time as well as similar values for skewness. The kurtosis parameters differ of course, as all models have their own way to capture kurtosis. For all models and all data sets we find significant non-normal kurtosis.\(^{13}\)

\(^{13}\)The value for $m_4$ (as used in the “ECO” model) implied by the normal distribution is about 0.53345.
Figure (3) gives an overview over the model’s fit, where LogL denotes log-likelihood, AIC the Akaike information criterion, BIC the Bayesian information criterion, KS the Kolgomorov-Smirnov distance and $\chi^2$ the $\chi^2$-test statistic with 10 classes.

### Gold Price

<table>
<thead>
<tr>
<th>Model</th>
<th>LogL</th>
<th>AIC</th>
<th>BIC</th>
<th>KS</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH</td>
<td>3268.252</td>
<td>-6530.503</td>
<td>-6514.359</td>
<td>1.3004</td>
<td>19.560</td>
</tr>
<tr>
<td>ECO</td>
<td>3283.845</td>
<td>-6557.690</td>
<td>-6530.782</td>
<td>0.5238</td>
<td>4.666</td>
</tr>
<tr>
<td>NEX</td>
<td>3284.064</td>
<td>-6558.129</td>
<td>-6531.221</td>
<td>0.5626</td>
<td>4.396</td>
</tr>
<tr>
<td>KAP</td>
<td>3283.851</td>
<td>-6557.703</td>
<td>-6530.795</td>
<td>0.4707</td>
<td>4.037</td>
</tr>
</tbody>
</table>

### Dow Jones Industrial Average

<table>
<thead>
<tr>
<th>Model</th>
<th>LogL</th>
<th>AIC</th>
<th>BIC</th>
<th>KS</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH</td>
<td>5241.281</td>
<td>-10476.56</td>
<td>-10460.42</td>
<td>2.0280</td>
<td>42.770</td>
</tr>
<tr>
<td>ECO</td>
<td>5255.646</td>
<td>-10501.29</td>
<td>-10474.39</td>
<td>1.6563</td>
<td>23.776</td>
</tr>
<tr>
<td>NEX</td>
<td>5252.58</td>
<td>-10495.16</td>
<td>-10468.26</td>
<td>1.7200</td>
<td>29.956</td>
</tr>
<tr>
<td>KAP</td>
<td>5255.698</td>
<td>-10501.40</td>
<td>-10474.50</td>
<td>1.4537</td>
<td>21.874</td>
</tr>
</tbody>
</table>

### Euro-USD Exchange Rate

<table>
<thead>
<tr>
<th>Model</th>
<th>LogL</th>
<th>AIC</th>
<th>BIC</th>
<th>KS</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH</td>
<td>6062.748</td>
<td>-12119.50</td>
<td>-12103.30</td>
<td>1.2970</td>
<td>21.442</td>
</tr>
<tr>
<td>ECO</td>
<td>6076.47</td>
<td>-12142.94</td>
<td>-12116.05</td>
<td>0.6817</td>
<td>5.747</td>
</tr>
<tr>
<td>NEX</td>
<td>6076.811</td>
<td>-12143.62</td>
<td>-12116.73</td>
<td>0.7017</td>
<td>9.754</td>
</tr>
<tr>
<td>KAP</td>
<td>6075.93</td>
<td>-12141.86</td>
<td>-12114.97</td>
<td>0.6125</td>
<td>5.217</td>
</tr>
</tbody>
</table>

Figure 3: Some goodness-of-fit measures for the applied models.

All generalizations of the GARCH model behave well - except for the Dow Jones Industrial Average, where for the $\chi^2$-test would reject all suggested models. In the sense of likelihood-based measures we can not derive a uniformly "best" model. The Kapur model outperforms all other models if we use only KS and $\chi^2$ as criteria.

We conclude that the inclusion of knowledge of higher moments such as skewness and kurtosis significantly improves our model - not only in the sense of likelihood, but also in likelihood based measure that penalize additional parameters as AIC or BIC and also in the sense of distance-based measures of fit, such as $\chi^2$ or KS.

Comparing the generalized models with each other, we find very similar results. This is of course due to the fact, that all models basically rely on the same additional information and only differ in the way of measuring this information.
7 Summary

The information-theoretic approaches to time series models considered in econometrics and econophysics may both be interpreted as special cases of models for time-varying moments using the generalized maximum entropy principle. Using this interpretation we derive a third model. Using three illustrative data sets typical for financial markets, we show that all suggested models from this class exhibit similar flexibility in capturing financial returns behavior.

References


Diskussionspapiere 2009
Discussion Papers 2009

01/2009  **Addison, John T. and Claus Schnabel**: Worker Directors: A German Product that Didn’t Export?

02/2009  **Uhde, André and Ulrich Heimeshoff**: Consolidation in banking and financial stability in Europe: Empirical evidence

03/2009  **Gu, Yiquan and Tobias Wenzel**: Product Variety, Price Elasticity of Demand and Fixed Cost in Spatial Models

04/2009  **Schlüter, Stephan**: A Two-Factor Model for Electricity Prices with Dynamic Volatility

05/2009  **Schlüter, Stephan and Fischer, Matthias**: A Tail Quantile Approximation Formula for the Student t and the Symmetric Generalized Hyperbolic Distribution

06/2009  **Ardelean, Vlad**: The impacts of outliers on different estimators for GARCH processes: an empirical study

Diskussionspapiere 2008
Discussion Papers 2008

01/2008  **Grimm, Veronika and Gregor Zoettl**: Strategic Capacity Choice under Uncertainty: The Impact of Market Structure on Investment and Welfare

02/2008  **Grimm, Veronika and Gregor Zoettl**: Production under Uncertainty: A Characterization of Welfare Enhancing and Optimal Price Caps

03/2008  **Engelmann, Dirk and Veronika Grimm**: Mechanisms for Efficient Voting with Private Information about Preferences

04/2008  **Schnabel, Claus and Joachim Wagner**: The Aging of the Unions in West Germany, 1980-2006

05/2008  **Wenzel, Tobias**: On the Incentives to Form Strategic Coalitions in ATM Markets

06/2008  **Herrmann, Klaus**: Models for Time-varying Moments Using Maximum Entropy Applied to a Generalized Measure of Volatility