A Socio-Finance Model: Inference and empirical application

Jørgen Vitting Andersen, Ioannis D. Vrontos, Petros Dellaportas, Serge Galam

UNIVERSITÉ PARIS 1
PANTHÉON SORBONNE

Documents de Travail du
Centre d'Economie de la Sorbonne

CES Working Papers

A Socio-Finance Model:
Inference and empirical application

Jorgen Vitting ANDERSEN, Ioannis D. VRONTOS
Petros DELLAPORTAS, Serge GALAM

2015.76

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
http://centredeconomiesorbonne.univ-paris1.fr/
ISSN : 1955-611X

A Socio-Finance Model: Inference and empirical application

Jørgen Vitting Andersen[a], Ioannis D. Vrontos[b], Petros Dellaportas[b], and Serge Galam[c]

[a] CNRS, Centre d'Economie de la Sorbonne, Université Paris 1 Panthéon-Sorbonne, Maison des Sciences Economiques, 106-112 Boulevard de lHôpital 75647 Paris Cedex 13, France

[b] Department of Statistics, Athens University of Economics and Business, Athens, Greece

[c] Cevipof - Center for Political Reseach, Sciences Po and CNRS, 98 rue de l'Université, 75007 Paris, France

Abstract

In this report we show the empirical application of our socio-finance model introduced in Andersen, Vrontos, Dellaportas and Galam (2014).

1 The Socio-Finance model

In the following we remind the reader of the definition of our socio-financial model Andersen, Vrontos, Dellaportas and Galam (2014).

Consider a population of market participants, shown schematically as circles in Figure 1A. We will proceed as in the so-called Galam model of opinion formation (Galam, 2005, 2012, and Biondi, Giannoccolo and Galam, 2012) and for simplicity imagine that people have just two different opinions on the market, which we can characterize as either ‘bullish’ (black circles) or ‘bearish’ (white circles). Letting $B(t)$ denote the proportion of bullishness in a population at time t, the proportion of bearishness is then $1 - B(t)$.

Insert Figure 1 about here

Figure 1A represents the opinions of the participants at the beginning of a given day. During the day people meet in random subgroups of different sizes, as illustrated by the different boxes in Figure 1B, to update their view of the market. Take, for example, the leftmost box in Figure 1B with six persons, two bullish, four bearish, who we can imagine are sitting around a table, or having a conference call, discussing the latest market developments. The outcome of the discussions for the different groups are illustrated in Figure 1C. For simplicity we have illustrated the case where a majority opinion in a given subgroup manages to polarize the opinion of the group by changing the opinion of those who had an opinion belonging to the minority. If we take the afore mentioned group of six persons we can see that after discussing, because of the majority polarizing rule, they have all become bearish. More realistically, we will in the following instead assume that is a certain *probability* for a majority opinion to prevail, and that even under certain conditions a minority could persuade a part of the majority to change their opinion.

For a given group of size k with j agents having a bullish opinion and $k - j$ a bearish opinion, we let $m_{k,j}$ denote the transition probability for all (k) members to adopt the bullish opinion as a result of their meeting. After one update taking into account communications in all groups of size k with j bullish agents, the new probability of finding an agent with a bullish view in the population can therefore be written:

$$B(t+1) = m_{k,j}(t)C_j^k B(t)^j [1 - B(t)]^{k-j} \tag{1}$$

where

$$C_j^k \equiv \frac{k!}{j!(k-j)!} \tag{2}$$

are the binomial coefficients. Notice that the transition probabilities $m_{k,j}$ depend on time, since we assume that they change as the market performance changes (this point will be explained further below).

Taking the sum over different groups of different sizes and different composition of bullishness within each group (see Figure 1B) one obtains a general term, $B(t + 1)$, for the bullishness in a population at time $t + 1$ due to the outcome of meetings of groups with different sizes and different composition of bullishness:

$$B(t+1) = \sum_{k=1}^{L} a_k \sum_{j=0}^{k} m_{k,j}(t) C_j^k B(t)^j [1 - B(t)]^{k-j}, \tag{3}$$

with

$$\sum_{k=1}^{L} a_k = 1,\ a_k \equiv \frac{1}{L},$$

with L denoting the size of the largest group and a_k denoting the weight of the group of size k. The link between communication and its impact on the markets can then be taken into account by assuming that the price return $r(t)$ changes whenever there is a change in the bullishness. The idea is that the bullishness itself is not the relevant factor determining how prices will change. Those feeling bullish would naturally already hold long positions on the market. Rather, when people change their opinion, say becoming more negative about the market, or less bullish, this will increase their tendency to sell. The fact that the absolute sentiment can act as a contrarian idicator (and the change in sentiment as an indicator) for future market returns is well known among practicioners[1]. Assuming the return to be proportional to the percentage change in bullishness, $RB(t)$, as well as economic news, $\eta(t)$, the return $r(t)$ is given by:

$$r(t) = \frac{1}{\lambda} RB(t) + \eta(t),\ \lambda > 0 \tag{4}$$

with $RB(t) = \frac{B(t)-B(t-1)}{B(t-1)}$ the change or 'return' of the bullishness. The variable $\eta(t) = r(t) - \frac{1}{\lambda} RB(t)$ is assumed to be either Gaussian or Student-t distributed with mean zero and a standard deviation that varies as a function of time depending on changes in sentiment. We will assume that the market will react to fundamental economic news represented by η but that the amplitude of the reaction depends on changes in the sentiment $RB(t)$:

$$\sigma(t) = \sigma_0 \exp\left[\frac{1}{\beta} |RB(t)|\right], \sigma_0 > 0, \beta > 0. \tag{5}$$

The influence of the financial market on decision-making can now be included in a natural way by letting the strength of persuasion depend on how the market has performed since the last meeting of the market participants. The idea is that, if for example the market had a dramatic downturn at the close yesterday, then in meetings the next morning, those with a bearish view will be more likely to convince even a bullish majority of their point of view. In the formal description below, this is taken into account by letting the transition probabilities for a change of opinion, i.e., the probabilities of transitions like Figure 1.B → Figure 1.C, depend on the market return over the last period:

$$m_{k,j}(t) = m_{k,j}(t-1) \exp\left[\frac{1}{\alpha} r(t)\right];\ m_{k,j}(t=0) \equiv j/k,\ \alpha > 0 \tag{6}$$

where α defines the scale for which a given return $r(t)$ impacts the transition probabilities. The condition $m_{k,j}(t=0) \equiv j/k$ describes the initially unbiased case where in average no market participant changes opinion.

[1] The Hulbert Stock Newsletter Sentiment Index (HSNSI) is used among practitioners as a contrarian signal for future stock returns, see also http://www.cxoadvisory.com/3265/sentiment-indicators/mark-hulbert/.

2 Estimation and Inference

In this section, we present the inferential method adopted to estimate the parameters of the socio-finance model. The method on which estimation will be based is *maximum likelihood*. Let $\boldsymbol{\theta}$ denotes the parameter vector to be estimated and $r = (r(1), r(2), ..., r(T))$ is the observed sample of size T. The approach will be to calculate the joint probability density

$$f_{R(1),R(2),...,R(T)}\left(r(1), r(2), ..., r(T)|\boldsymbol{\theta}\right) \tag{7}$$

which can be viewed as the probability of having observed this particular sample[2]. Then, the maximum likelihood estimate (MLE) of $\boldsymbol{\theta}$ is the vector $\widehat{\boldsymbol{\theta}}$ that maximizes the joint probability density (7), that is, the vector for which this sample is most likely to have been observed. To find maximum likelihood estimates, first we will have to calculate the likelihood function and then to find values of $\boldsymbol{\theta}$ that maximize this function. In this section, we will present analytically the calculation of the likelihood function under the assumption of a normal and a Student-t distribution for the error process $\eta(t)$, while maximization of this function will be based on numerical optimization algorithms[3].

The likelihood function i.e. the joint probability density of the complete sample $r(1), r(2), ..., r(T)$ for the proposed socio-finance model can be written as

$$\begin{aligned} f\left[r|B(0), m_{k,j}(0), \boldsymbol{\theta}\right] &= f\left[r(T)|r(1), ..., r(T-1), B(0), m_{k,j}(0), \boldsymbol{\theta}\right] \cdot f\left[r(1), ..., r(T-1)|B(0), m_{k,j}(0), \boldsymbol{\theta}\right] \\ &= f\left[r(T)|r(1), ..., r(T-1), B(0), m_{k,j}(0), \boldsymbol{\theta}\right] \cdot \\ &\quad \cdot f\left[r(T-1)|r(1), ..., r(T-2), B(0), m_{k,j}(0), \boldsymbol{\theta}\right] \cdot f\left[r(1), ..., r(T-2)|B(0), m_{k,j}(0), \boldsymbol{\theta}\right] \\ &= ... = f\left[r(1)|B(0), m_{k,j}(0), \boldsymbol{\theta}\right] \cdot \prod_{t=2}^{T} f\left[r(t)|\Phi\left(t-1\right), B(0), m_{k,j}(0), \boldsymbol{\theta}\right], \end{aligned}$$

where $\Phi\left(t-1\right)$ is the information set up to time $t-1$.

Based on this property, that the joint probability can be written as a product of conditional probabilities, we can calculate the likelihood function as follows. First, under the normality assumption for the error process $\eta(t)$. Consider the probability distribution of $r(1)$, the first observation in the sample. Since $\eta(t)$ is Gaussian, $r(1)$ is also Gaussian and the density of the first observation, conditional on $B(0)$ and $m_{k,j}(0) = j/k$, takes the form

$$f\left[r(1)|B(0), m_{k,j}(0), \boldsymbol{\theta}_N\right] = \frac{1}{\sqrt{2\pi}\sqrt{\sigma(1)^2}} \exp\left\{-\frac{1}{2\sigma(1)^2}\left[r(1) - \frac{1}{\lambda}RB(1)\right]^2\right\},$$

[2]With capital letters, $R(t)$, we denote the random variable of the return at time t, while with small letters, $r(t)$, we denote a particular value that the random variable takes at time t.

[3]Note that maximum likelihood estimates for dynamic non-linear models that take into account the heteroscedastic characteristics of financial series are usually obtained by using numerical optimization algorithms such as the scoring algorithm, the method proposed by Mak (1993) and developed further by Mak, Wong and Li (1997), the Berndt, Hall, Hall and Hausman (1974) algorithm, the Broyden, Fletcher, Golfarb and Shanno (BFGS) algorithm (Golfarb, 1970 and Shanno, 1970) which is a quasi-Newton method, or by using mixed-gradient algorithms to accelerate convergence (see, for example, Fiorentini, Calzolari, and Panattoni, 1996, Vrontos, Dellaportas and Politis, 2003, Diamantopoulos and Vrontos (2010), Vrontos (2013).

where $\boldsymbol{\theta}_N = (\lambda,\, \sigma_0,\, \beta,\, \alpha)'$ denotes the parameter vector to be estimated under the normality assumption. Next, conditioning on $r(1)$, the density of the second observation $r(2)$ is

$$f\left[r(2)|r(1), B(0), m_{k,j}(0), \boldsymbol{\theta}_N\right] = \frac{1}{\sqrt{2\pi}\sqrt{\sigma(2)^2}} \exp\left\{-\frac{1}{2\sigma(2)^2}\left[r(2) - \frac{1}{\lambda}RB(2)\right]^2\right\}.$$

Proceeding in this fashion, the conditional density of the $t-th$ observation can be calculated as

$$f\left[r(t)|r(t-1), ..., r(1), B(0), m_{k,j}(0), \boldsymbol{\theta}_N\right] = \frac{1}{\sqrt{2\pi}\sqrt{\sigma(t)^2}} \exp\left\{-\frac{1}{2\sigma(t)^2}\left[r(t) - \frac{1}{\lambda}RB(t)\right]^2\right\}.$$

Therefore, the likelihood of the complete sample can be written as

$$f\left[r|B(0), m_{k,j}(0), \boldsymbol{\theta}_N\right] = (2\pi)^{-\frac{T}{2}} \prod_{t=1}^{T} \left[\sigma(t)^2\right]^{-1/2} \exp\left\{-\frac{1}{2}\sum_{t=1}^{T}\frac{1}{\sigma(t)^2}\left[r(t) - \frac{1}{\lambda}RB(t)\right]^2\right\}, \tag{8}$$

The log-likelihood function, denoted $L_N\left(r|\boldsymbol{\theta}_N\right)$, can be written as

$$L_N\left(r|\boldsymbol{\theta}_N\right) = -\frac{T}{2}\ln\left(2\pi\right) - \frac{1}{2}\sum_{t=1}^{T}\left[\ln\sigma(t)^2\right] - \frac{1}{2}\sum_{t=1}^{T}\frac{1}{\sigma(t)^2}\left[r(t) - \frac{1}{\lambda}RB(t)\right]^2. \tag{9}$$

Clearly, the vector of $\boldsymbol{\theta}_N$ that maximizes the conditional likelihood (8) is identical to the vector that maximizes the conditional log-likelihood (9).

Although the normal distribution is the most commonly used in applications, there is empirical evidence that the distribution of financial time series has usually fat tails, even after taking into account the volatility clustering phenomenon. In other words, the normality assumption of standardised residuals of estimated financial models is usually rejected in most financial applications. A solution to this problem is to specify a distribution that accounts for fat tails and deviations from normality such as the Student-t (see, for example, Bollerslev, 1987) or a Generalized Error distribution (Nelson, 1991). Under the Student-t distribution with v degrees of freedom for the error process $\eta(t)$, the likelihood for the socio-finance model for the complete sample can be written as

$$\begin{aligned} f_{ST}\left[r|B(0), m_{k,j}(0), \boldsymbol{\theta}_{ST}\right] \;=\; & \left[\Gamma\left(\frac{v+1}{2}\right)\right]^{T}\left[\Gamma\left(\frac{v}{2}\right)\right]^{-T}\left[\pi\left(v-2\right)\right]^{-T/2} \\ & \prod_{t=1}^{T}\left[\sigma(t)^2\right]^{-1/2}\prod_{t=1}^{T}\left[1 + \frac{\left(r(t) - \frac{1}{\lambda}RB(t)\right)^2}{\left(v-2\right)\sigma(t)^2}\right]^{-(v+1)/2}, \end{aligned} \tag{10}$$

where $\Gamma\left(.\right)$ is the gamma function. The log-likelihood can be written as

$$\begin{aligned} L_{ST}\left(r|\boldsymbol{\theta}_{ST}\right) \;=\; & T\ln\Gamma\left(\frac{v+1}{2}\right) - T\ln\Gamma\left(\frac{v}{2}\right) - \frac{T}{2}\ln\left[\pi\left(v-2\right)\right] \\ & -\frac{1}{2}\sum_{t=1}^{T}\left[\ln\sigma(t)^2\right] - \frac{v+1}{2}\sum_{t=1}^{T}\ln\left[1 + \frac{\left(r(t) - \frac{1}{\lambda}RB(t)\right)^2}{\left(v-2\right)\sigma(t)^2}\right], \end{aligned} \tag{11}$$

where $\boldsymbol{\theta}_{ST} = (\lambda,\, \sigma_0,\, \beta,\, \alpha,\, v)'$ denotes the parameter vector to be estimated under the Student-t distribution.

In order to avoid the positivity restrictions for the parameters, $\lambda > 0$, $\sigma_0 > 0$, $\beta > 0$ and $\alpha > 0$ we use the logarithmic transformation, so that $\lambda^* = \ln\left(\lambda\right)$, $\sigma_0^* = \ln\left(\sigma_0^2\right)$, $\beta^* = \ln\left(\beta\right)$ and $\alpha^* = \ln\left(\alpha\right)$.

For the degrees of freedom parameter v we use the transformation $v^* = \ln(v-2)$. Thus the parameter vector to estimate is $\boldsymbol{\theta}_N^* = (\lambda^*,\ \sigma_0^*,\ \beta^*,\ \alpha^*)'$ under conditional normality, and $\boldsymbol{\theta}_{ST}^* = (\lambda^*,\ \sigma_0^*,\ \beta^*,\ \alpha^*,\ v^*)'$ under the Student-t error distribution. The technique of reparametrizing the model parameters is very useful in order to ensure that a numerical optimization algorithm always provides parameter values within certain specified boundaries. An attractive feature of the method of maximum likelihood estimation is its invariance to one-to-one transformations of the parameters of the log-likelihood. That is, the maximum likelihood solution is invariant under transformation of parameters. Finally, due to the highly non-linear nature of the proposed socio-finance model, maximization of the conditional log-likelihood with respect to the model parameters is achieved by using numerical optimization algorithms. In this study, we maximize the log-likelihood function by applying the optimization functions '*fminsearch*' and/or '*fminunc*' of Matlab.

3 Simulation Study

In this section we conduct a simulation study concerning the proposed socio-finance model. The aim of this study is to assess the performance of the inferential method, based on maximum likelihood, to estimate the model parameters and the proportion of bullishness. We conduct a series of simulation experiments considering different sample sizes of time series, i.e. $T = 2000$ and $T = 5000$ data, and various parameter values. Different starting values of the model parameters are used to ensure that the maximization/minimization algorithm converges to the true simulated parameters. We simulate data from the socio-finance model (1-6) under the normality assumption for the error process. However, we estimate the model parameters using both normal and Student-t errors. Under the assumption of Student-t errors, we expect that the estimated parameter values will be close to the true simulated values, while the estimated degrees of freedom will be large enough, since the data are simulated from a normal distribution.

Insert Table 1 - Table 6 about here

We present three simulation experiments; in the first two simulation scenarios we simulate $T = 2000$ and $T = 5000$ data points, respectively, using $\lambda = 1.1$, $\sigma_0 = 0.01$, $\beta = 0.001$ and $a = 400$ as the 'true' parameter values. In the third simulation scenario we simulate $T = 2000$ data points using $\lambda = 2.8$, $\sigma_0 = 0.02$, $\beta = 0.04$ and $a = 2.65$ as the 'true' parameter values. Table 1 and Table 2, together with Table 3 and Table 4 present the estimation results for the first and second simulation scenario ($T = 2000$ and $T = 5000$) under the normal (Table 1 and Table 3) and the Student-t (Table 2 and Table 4) distribution for the error process, respectively, while Table 5 and Table 6 present the corresponding estimates for the third simulation ($T = 2000$). The first column of Table 1 - Table 6 shows the 'true' parameter values $\boldsymbol{\theta} = (\lambda,\ \sigma_0,\ \beta,\ \alpha)'$ used to simulate the data, while column 2 presents the 'true' transformed parameter values $\boldsymbol{\theta}^* = (\lambda^*,\ \sigma_0^*,\ \beta^*,\ \alpha^*)'$. The maximum likelihood estimates $\widehat{\boldsymbol{\theta}}^*$ of the transformed parameters $\boldsymbol{\theta}^*$ and their standard errors are presented in column 3 and 4, respectively, while in column 5 are given the corresponding estimated parameter values $\widehat{\boldsymbol{\theta}}$. Finally, in the last column we present the value of the gradient of the log-likelihood function evaluated at the parameter estimates $\widehat{\boldsymbol{\theta}}^*$. Different starting values for the model parameters have been used in the maximum likelihood estimation procedure (see Panel A-B). The results presented in Tables 1-6 indicate that the maximum likelihood inferential procedure provide estimates that are close to the 'true' parameter values, taking into account the estimates and the corresponding standard errors. The results provide evidence of convergence of the maximization/minimization algorithm since the estimates found by using different starting values are very similar, and the value of the gradient of the log-likelihood evaluated at the parameter estimates is near zero. Similar results are obtained when we simulate $T = 5000$ data points. Inspection of these results indicates that the maximization algorithm converges to almost identical parameter estimates using different starting values and their corresponding standard errors are smaller than those obtained by the simulated data based on T=2000 data points. This seems reasonable since the sample size increases.

Insert Figure 2-Figure 7 about here

Next, we examine whether the proposed algorithm can estimate adequately the proportion of bullishness across time and the time-varying conditional volatilities. To this end, in Figure 2 - Figure 7, we present the simulated prices $P(t)$, and the corresponding returns $r(t)$, as well as the simulated bullishness proportions $B(t)$ across time, for $T = 2000$ and $T = 5000$ data point for the three simulation experiments, respectively. We also present the estimated bullishness proportions $\widehat{B}(t)$ and the estimated conditional volatilities $\widehat{\sigma}(t)$ which are based on the parameter estimates of the socio-finance model. Looking at Figures 2-7(b) and Figure 2-7(d) which illustrate the 'true' and the estimated bullishness proportions, respectively, we can observe that these proportions are almost identical. Thus, the proposed inferential procedure manages to identify and detect correctly the proportions of a bullish view on the simulated prices. Finally, looking at Figure 2-7(c) and Figure 2-7(e), which illustrate the simulated returns and the estimated conditional volatilities, respectively, we can see that periods of lower or higher deviation in the return series can be detected by the socio-finance volatility estimates. Therefore, the proposed model is able to capture the volatility clustering phenomenon of the return series.

4 Application to the European Union Bank index

In this section we present an empirical application of the proposed socio-finance model to the European Union Banks five-year index. The idea is to consider a very volatile market to study abrupt and large changes in market performance and map the corresponding evolution in sentiments. The data consists of 1505 daily prices over the 1/1/2008-10/7/2013 period. We compute and analyse the returns of the EU banks 5-year index. Figure 8(b) presents the EU bank return series, which shows that the volatility of the return series changes over time. There is also high kurtosis (14.7) in the return series. Thus, there is evidence for fat tails and volatility clustering phenomenon in the EU bank index return series, that the socio-finance model is inveted to deal with.

Insert Figure 8 about here

First, we apply the proposed socio-finance model under the assumption of the conditional normal distribution for the error process $\eta(t)$ to the EU banks returns. Table 7 (Panel A) presents the maximum likelihood estimates of the socio-finance model under normal errors. Specifically Table 7 presents the estimates $\widehat{\boldsymbol{\theta}}^*$ of the transformed parameters $\boldsymbol{\theta}^*$ and their standard errors (columns 1 and 2, respectively), the corresponding estimated parameter values $\widehat{\boldsymbol{\theta}}$ (column 3), and the value of the gradient of the log-likelihood function evaluated at the parameter estimates $\widehat{\boldsymbol{\theta}}^*$ (last column). Looking at the parameter estimates of Table 7 (Panel A) we observe that the transformed parameter $\widehat{\lambda}^*$ in the mean equation, the transformed parameters σ_0^* and $\widehat{\beta}^*$ in the variance equation, and parameter α^* in the transition probabilities equation, are all statistically significant. These results show that the change in the bullishness proportion is able to explain the price return $r(t)$ as well as the time-varying volatility of the return series.

Insert Table 7 about here

Next, we estimate the proportion of bullishness across time and the time-varying conditional volatilities under the assumption of conditional normal distribution for the error process $\eta(t)$. In Figure 8, we present the EUBanks prices $P(t)$, and the corresponding returns $r(t)$, as well as the estimated bullishness proportions $\widehat{B}(t)$ and the estimated conditional volatilities $\widehat{\sigma}(t)$, which are based on the parameter estimates of the socio-finance model. Comparing the EUBanks price evolution Figure 8(a) and the estimated bullishness proportions Figure 8(c), one observes that an increase (decrease) on the price of EUBanks index can be affected by the estimated proportions of a bullish view, which reflects the opinion of different groups about the movement of the index. Finally, comparing the observed volatility of EUBanks index Figure 8(b) and the estimated volatility from the model Figure 8(d), one notices that periods of lower or higher deviation in the return series is indeed detected by volatility estimates of the socio-finance model. Therefore, the proposed model is able to capture the volatility clustering phenomenon of the return series.

Having estimated the model parameters, one can examine the appropriateness of the assumption of conditional normality. To this end, we apply the Jarque-Bera and the Kolmogorov test of normality to the standardized residuals ($\frac{\widehat{\eta}(t)}{\widehat{\sigma}(t)}$). These tests show that the null hypothesis of normality is rejected ($Jarque - Bera\ p - value$ = 0.004, $Kolmogorov\ p - value$ = 0.000), thus the standardised residual series exhibit deviations from normality. Therefore, the normality assumption is violated[4] and it is worth estimating a socio-finance model based on the Student-t error distribution.

Insert Figure 9 about here

[4] Due to the heteroskedastic characteristics of the EUBanks return series, we have also estimated the well-known GARCH(1,1) model with normal errors to the EUBanks return series and examine the appropriateness of the assumption of conditional normality under this alternative model specification. The corresponding p-values of the Jarque-Bera and the Kolmogorov test of normality to the standardized residuals of the GARCH(1,1) model was 0.001 and 0.000, respectively, indicating that the GARCH(1,1) model with normal errors can not capture the fat tails of the return's distribution.

Next, we apply the proposed socio-finance model based on Student-t errors. Table 7 (Panel B) presents the estimates $\widehat{\boldsymbol{\theta}}^*$ of the transformed parameters $\boldsymbol{\theta}^*$ and their standard errors (columns 1 and 2, respectively), the corresponding estimated parameter values $\widehat{\boldsymbol{\theta}}$ (column 3), and the value of the gradient of the log-likelihood function evaluated at the parameter estimates $\widehat{\boldsymbol{\theta}}^*$ (last column). Looking at the parameter estimates of Table 7 (Panel B) we observe that the parameter $\widehat{\lambda}^*$ in the mean equation, the parameters σ_0^* and $\widehat{\beta}^*$ in the variance equation, the parameter α^* in the transition probabilities equation, and the degrees of freedom v^*, are all statistically significant. Under the Student-t model, the degrees of freedom parameter, v, is estimated to be 2.312, indicating heavy tails. Applying the Kolmogorov test of a Student-t distribution to the standardized residuals $(\frac{\widehat{\eta}(t)}{\widehat{\sigma}(t)})$, the null hypothesis of Student-t errors is not rejected ($Kolmogorov\ p-value = 0.272$), indicating that the Student-t socio-finance model provides an appropriate modelling approach than the corresponding normal model[5].

Insert Figure 10 about here

We also estimate the proportion of bullishness across time and the time-varying conditional volatilities under the assumption of conditional Student-t distribution for the error process $\eta(t)$. In Figure 9, we present the EUBanks prices $P(t)$, and the corresponding returns $r(t)$, as well as the estimated bullishness proportions $\widehat{B}(t)$ and the estimated conditional volatilities $\widehat{\sigma}(t)$, which are based on the parameter estimates of the Student-t socio-finance model. Looking at Figure 9(a) and 9(c) we observe that the bullishness proportion explain the prices of the index, and looking at Figure 9(b) and 9(d) we can see that periods of lower or higher deviation in the return series can be detected by the socio-finance volatility estimates, i.e. we arrive at similar conclusions with those taken by the normal socio-finance model. However, by comparing the volatilities taken by the socio-finance normal model [see Figure 8(d)], and those taken by the socio-finance Student-t model [Figure 9(d)] we observe that the estimated volatilities taken by the Student-t socio-finance model are higher than the coresponding estimated volatilities of the socio-finance normal model. This result can explain the better fit of the Student-t model, which seems to give a larger volatility signal when periods of higher deviation in the return series is coming. Similar conclusion can be drawn by comparing the estimated volatilities of Normal-GARCH [Figure 10(c)], and Student-t GARCH [Figure 10(d)] with those of the Student-t socio-finance model [Figure 9(d)]. This explains the inadequacy of GARCH-type models in this empirical application.

To conclude, the above findings show that the change in the bullishness proportion is able to explain the price return $r(t)$ as well as the time-varying volatility of the return series, while the assumption of a Student-t error distribution is an appropriate choice to capture the fat tail property of the EUBanks return series. Therefore, the proposed socio-finance model with Student-t errors provides a reasonable modelling approach for the EUBanks return series.

[5] We have also estimated the GARCH(1,1) model with Student-t errors to the EUBanks return series as an alternative modelling approach. The corresponding p-value of the Kolmogorov test of a Student-t distribution to the standardized residuals of the GARCH(1,1) model was 0.028, indicating that the GARCH(1,1) model with Student-t errors can not capture adequately the tails of the return's distribution.

5 Discussion

In this paper we have proposed a socio-finance Student-t model for the analysis of financial return series. The aim of our analysis was, first, to develop a new class of models that might be usefull for the empirical analysis of financial series, and, second, to apply the model to real data in order to show the benefits of this modelling approach. Furthermore, we have proposed a classical approach to inference based on maximum likelihood to estimate the model parameters.

The proposed modelling approach is particularly useful in cases where we believe that communication of different groups of the population about the movement of the market can affect the returns and/or the volatility of the financial assets as well as in cases where the distribution of returns is characterised by large kurtosis, fat tails, or in general deviates from normality. In those cases, the normality assumption may not be adequate, while the socio-finance Student-t modelling approach provides more reliable results. We have applied our proposed approach to the EUBanks index and found evidence that the proposed model can explain the movement of the prices of the index and its time-varying volatility via the change in bullishness proportion.

We believe that the proposed socio-finance model can be used as an alternative reliable modelling approach to empirical applications. Clearly, many interesting questions remain open and various topics for future research arise in the context of this new class of models.

.

Acknowledgments

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7-SSH/2007-2013) under grant agreement no 320270 “SYRTO”.

References

Andersen, J.V., Vrontos, I.D., Dellaportas, P. and S. Galam (2014). Communication Impacting Financial Markets, *Working paper*, available at: http://papers.ssrn.com/so13/papers.cfm?abstract_id=2428745.

Berndt, E.K., Hall, B.H., Hall, R.E., and J.A. Hausman (1974), Estimation and Inference in Nonlinear Structural Models, *Annals of Economic and Social Measurement,* 3/4, 653-665.

Biondi, Y., P. Giannoccolo and S. Galam (2012), Formation of share market prices under heterogeneous beliefs and common knowledge, *Physica A,* 391, 5532-5545.

Bollerslev, T. (1987), A conditionally heteroskedastic time series model for speculative prices and rates of return, *Review of Economics and Statistics*, 69, 542-547.

Diamantopoulos, K. and I.D. Vrontos (2010), A Student-t Full Factor Multivariate GARCH Model, *Computational Economics*, 35, 1, 63-83.

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

Galam, S. (2005), Local dynamics vs social mechanisms: a unifying frame, *Europhysics Letters,* 70, 705-711.

Galam, S. (2012), Sociophysics: A Physicist's Modeling of Psycho-political Phenomena, Springer, New York.

Goldfarb, D. (1970), A family of variable-metric methods derived by variational means, *Mathematics of Computation*, 24, 23-26.

Mak, T.K. (1993), Solving Non-linear estimation equations, *Journal of Royal Statistical Society*, B, 55, 945-955.

Mak, T.K., Wong, H., and W.K. Li (1997), Estimation of nonlinear time series with conditional heteroscedastic variances by iteratively weighted least squares, *Journal of Computational Statistics and Data Analysis*, 24, 169-178.

Nelson, D. (1991), Conditional heteroskedasticity in asset returns: A new approach, *Econometrica*, 59, 347-370.

Shanno, D.F. (1970), Conditioning on Quasi-Newton methods for function minimization, *Mathematics of Computation*, 24, 647-656.

Vrontos, I.D., Dellaportas, P. and Politis D.N. (2003), A full-factor multivariate GARCH model. *Econometrics Journal*, 6, 312-334.

Vrontos, I.D. (2012), Evidence for hedge fund predictability from a multivariate Student's t full-factor GARCH model, *Journal of Applied Statistics*, 39, 6, 1295-1321.

Table 1: Maximum likelihood estimation results of the first simulation experiment ($T = 2000$) using normal errors.

Panel A: Starting values $\widehat{\lambda}^*$=-0.693, $\widehat{\sigma}_0^*$=-6.438, $\widehat{\beta}^*$=-3.507, $\widehat{\alpha}^*$=5.704									
True initial		True trans.			MLEs trans.	StdErr	MLEs		Gradient
λ	1.1	λ^*	0.095	$\widehat{\lambda}^*$	0.344	1.346	$\widehat{\lambda}$	1.4103	-0.0008
σ_0	0.01	σ_0^*	-9.210	$\widehat{\sigma}_0^*$	-9.119	0.051	$\widehat{\sigma}_0$	0.0105	0.0004
β	0.001	β^*	-6.908	$\widehat{\beta}^*$	-6.845	0.349	$\widehat{\beta}$	0.0011	-0.0002
α	400	α^*	5.991	$\widehat{\alpha}^*$	5.995	0.371	$\widehat{\alpha}$	401.33	0.0006
Panel B: Starting values $\widehat{\lambda}^*$=-2.996, $\widehat{\sigma}_0^*$=-13.816, $\widehat{\beta}^*$=-6.908, $\widehat{\alpha}^*$=6.215									
True initial		True trans.			MLEs trans.	StdErr	MLEs		Gradient
λ	1.1	λ^*	0.095	$\widehat{\lambda}^*$	0.321	1.290	$\widehat{\lambda}$	1.3780	-0.0145
σ_0	0.01	σ_0^*	-9.210	$\widehat{\sigma}_0^*$	-9.119	0.051	$\widehat{\sigma}_0$	0.0105	0.1039
β	0.001	β^*	-6.908	$\widehat{\beta}^*$	-6.844	0.348	$\widehat{\beta}$	0.0011	-0.0109
α	400	α^*	5.991	$\widehat{\alpha}^*$	5.993	0.370	$\widehat{\alpha}$	400.79	-0.0361

True initial: denotes the 'True' simulated parameter values $\boldsymbol{\theta}$, True trans.: denotes the 'True' transformed simulated parameter values $\boldsymbol{\theta}^*$, MLEs trans: denote the maximum likelihood estimates $\widehat{\boldsymbol{\theta}}^*$ of the transformed parameters $\boldsymbol{\theta}^*$, StdErr: denotes the standard errors of the transformed parameters $\boldsymbol{\theta}^*$, MLEs: denote the maximum likelihood estimates $\widehat{\boldsymbol{\theta}}$ of the parameters $\boldsymbol{\theta}$, Gradient: denotes the value of the gradient of the log-likelihood evaluated at the parameter estimates.

Table 2: Maximum likelihood estimation results of the first simulation experiment ($T = 2000$) using Student-t errors.

Panel A: Starting values $\widehat{\lambda}^*$=-0.693, $\widehat{\sigma}_0^*$=-6.438, $\widehat{\beta}^*$=-3.507, $\widehat{\alpha}^*$=5.704, $\widehat{v}^*$=1.609									
True initial		True trans.			MLEs trans.	StdErr	MLEs		Gradient
λ	1.1	λ^*	0.095	$\widehat{\lambda}^*$	0.297	1.286	$\widehat{\lambda}$	1.3454	0.0005
σ_0	0.01	σ_0^*	-9.210	$\widehat{\sigma}_0^*$	-9.122	0.053	$\widehat{\sigma}_0$	0.0105	0.0010
β	0.001	β^*	-6.908	$\widehat{\beta}^*$	-6.847	0.360	$\widehat{\beta}$	0.0011	-0.0021
α	400	α^*	5.991	$\widehat{\alpha}^*$	5.994	0.382	$\widehat{\alpha}$	400.81	-0.0020
				$\widehat{v}^*$	3.530	0.782	$\widehat{v}$	36.14	0.0003
Panel B: Starting values $\widehat{\lambda}^*$=-2.996, $\widehat{\sigma}_0^*$=-13.816, $\widehat{\beta}^*$=-6.908, $\widehat{\alpha}^*$=6.215, $\widehat{v}^*$=1.099									
True initial		True trans.			MLEs trans.	StdErr	MLEs		Gradient
λ	1.1	λ^*	0.095	$\widehat{\lambda}^*$	0.298	1.281	$\widehat{\lambda}$	1.3466	-0.0008
σ_0	0.01	σ_0^*	-9.210	$\widehat{\sigma}_0^*$	-9.122	0.053	$\widehat{\sigma}_0$	0.0105	0.0085
β	0.001	β^*	-6.908	$\widehat{\beta}^*$	-6.847	0.360	$\widehat{\beta}$	0.0011	-0.0056
α	400	α^*	5.991	$\widehat{\alpha}^*$	5.993	0.382	$\widehat{\alpha}$	400.79	-0.0051
				$\widehat{v}^*$	3.530	0.781	$\widehat{v}$	36.12	-0.0017

True initial: denotes the 'True' simulated parameter values $\boldsymbol{\theta}$, True trans.: denotes the 'True' transformed simulated parameter values $\boldsymbol{\theta}^*$, MLEs trans: denote the maximum likelihood estimates $\widehat{\boldsymbol{\theta}}^*$ of the transformed parameters $\boldsymbol{\theta}^*$, StdErr: denotes the standard errors of the transformed parameters $\boldsymbol{\theta}^*$, MLEs: denote the maximum likelihood estimates $\widehat{\boldsymbol{\theta}}$ of the parameters $\boldsymbol{\theta}$, Gradient: denotes the value of the gradient of the log-likelihood evaluated at the parameter estimates.

Table 3: Maximum likelihood estimation results of the second simulation experiments ($T = 5000$) using normal errors.

Panel A: Starting values $\widehat{\lambda}^*$=-0.693, $\widehat{\sigma}_0^*$=-6.438, $\widehat{\beta}^*$=-3.507, $\widehat{\alpha}^*$=5.704									
True initial		True trans.		MLEs trans.		StdErr	MLEs		Gradient
λ	1.1	λ^*	0.095	$\widehat{\lambda}^*$	0.174	0.806	$\widehat{\lambda}$	1.1903	-0.0009
σ_0	0.01	σ_0^*	-9.210	$\widehat{\sigma}_0^*$	-9.189	0.028	$\widehat{\sigma}_0$	0.0101	-0.0601
β	0.001	β^*	-6.908	$\widehat{\beta}^*$	-6.888	0.042	$\widehat{\beta}$	0.0010	0.0658
α	400	α^*	5.991	$\widehat{\alpha}^*$	5.975	0.038	$\widehat{\alpha}$	393.32	0.4501
Panel B: Starting values $\widehat{\lambda}^*$=-2.996, $\widehat{\sigma}_0^*$=-13.816, $\widehat{\beta}^*$=-6.908, $\widehat{\alpha}^*$=6.215									
True initial		True trans.		MLEs trans.		StdErr	MLEs		Gradient
λ	1.1	λ^*	0.095	$\widehat{\lambda}^*$	0.174	0.808	$\widehat{\lambda}$	1.1905	0.0045
σ_0	0.01	σ_0^*	-9.210	$\widehat{\sigma}_0^*$	-9.189	0.028	$\widehat{\sigma}_0$	0.0101	0.0308
β	0.001	β^*	-6.908	$\widehat{\beta}^*$	-6.888	0.041	$\widehat{\beta}$	0.0010	0.0112
α	400	α^*	5.991	$\widehat{\alpha}^*$	5.975	0.036	$\widehat{\alpha}$	393.32	-0.2521

True initial: denotes the 'True' simulated parameter values $\boldsymbol{\theta}$, True trans.: denotes the 'True' transformed simulated parameter values $\boldsymbol{\theta}^*$, MLEs trans: denote the maximum likelihood estimates $\widehat{\boldsymbol{\theta}}^*$ of the transformed parameters $\boldsymbol{\theta}^*$, StdErr: denotes the standard errors of the transformed parameters $\boldsymbol{\theta}^*$, MLEs: denote the maximum likelihood estimates $\widehat{\boldsymbol{\theta}}$ of the parameters $\boldsymbol{\theta}$, Gradient: denotes the value of the gradient of the log-likelihood evaluated at the parameter estimates.

Table 4: Maximum likelihood estimation results of the second simulation experiments ($T = 5000$) using Student-t errors.

Panel A: Starting values $\widehat{\lambda}^*$=-0.693, $\widehat{\sigma}_0^*$=-6.438, $\widehat{\beta}^*$=-3.507, $\widehat{\alpha}^*$=5.704, $\widehat{\upsilon}^*$=1.609									
True initial		True trans.			MLEs trans.	StdErr	MLEs		Gradient
λ	1.1	λ^*	0.095	$\widehat{\lambda}^*$	0.174	0.797	$\widehat{\lambda}$	1.1903	-0.0005
σ_0	0.01	σ_0^*	-9.210	$\widehat{\sigma}_0^*$	-9.189	0.028	$\widehat{\sigma}_0$	0.0101	0.0093
β	0.001	β^*	-6.908	$\widehat{\beta}^*$	-6.888	0.042	$\widehat{\beta}$	0.0010	-0.0043
α	400	α^*	5.991	$\widehat{\alpha}^*$	5.975	0.038	$\widehat{\alpha}$	393.32	0.8404
				$\widehat{\upsilon}^*$	5.685	3.498	$\widehat{\upsilon}$	296.29	-0.0071
Panel B: Starting values $\widehat{\lambda}^*$=-2.996, $\widehat{\sigma}_0^*$=-13.816, $\widehat{\beta}^*$=-6.908, $\widehat{\alpha}^*$=6.215, $\widehat{\upsilon}^*$=2.303									
True initial		True trans.			MLEs trans.	StdErr	MLEs		Gradient
λ	1.1	λ^*	0.095	$\widehat{\lambda}^*$	0.174	0.808	$\widehat{\lambda}$	1.189	-0.0021
σ_0	0.01	σ_0^*	-9.210	$\widehat{\sigma}_0^*$	-9.189	0.028	$\widehat{\sigma}_0$	0.0101	0.0475
β	0.001	β^*	-6.908	$\widehat{\beta}^*$	-6.888	0.042	$\widehat{\beta}$	0.0010	0.0287
α	400	α^*	5.991	$\widehat{\alpha}^*$	5.975	0.038	$\widehat{\alpha}$	393.32	0.3934
				$\widehat{\upsilon}^*$	5.681	3.635	$\widehat{\upsilon}$	295.19	-0.0025

True initial: denotes the 'True' simulated parameter values $\boldsymbol{\theta}$, True trans.: denotes the 'True' transformed simulated parameter values $\boldsymbol{\theta}^*$, MLEs trans: denote the maximum likelihood estimates $\widehat{\boldsymbol{\theta}}^*$ of the transformed parameters $\boldsymbol{\theta}^*$, StdErr: denotes the standard errors of the transformed parameters $\boldsymbol{\theta}^*$, MLEs: denote the maximum likelihood estimates $\widehat{\boldsymbol{\theta}}$ of the parameters $\boldsymbol{\theta}$, Gradient: denotes the value of the gradient of the log-likelihood evaluated at the parameter estimates.

Table 5: Maximum likelihood estimation results of the third simulation experiments ($T = 2000$) using normal errors.

Panel A: Starting values $\widehat{\lambda}^*$=0.095, $\widehat{\sigma}_0^*$=–9.210, $\widehat{\beta}^*$=–6.908, $\widehat{\alpha}^*$=3.689									
True initial		True trans.		MLEs trans.		StdErr	MLEs		Gradient
λ	2.8	λ^*	1.029	$\widehat{\lambda}^*$	0.847	0.169	$\widehat{\lambda}$	2.332	-0.0003
σ_0	0.02	σ_0^*	-7.824	$\widehat{\sigma}_0^*$	-7.877	0.048	$\widehat{\sigma}_0$	0.019	-0.0010
β	0.04	β^*	-3.219	$\widehat{\beta}^*$	-3.339	0.106	$\widehat{\beta}$	0.036	0.0013
α	2.65	α^*	0.975	$\widehat{\alpha}^*$	1.033	0.107	$\widehat{\alpha}$	2.811	0.0970
Panel B: Starting values $\widehat{\lambda}^*$=-0.693, $\widehat{\sigma}_0^*$=-6.438, $\widehat{\beta}^*$=-3.507, $\widehat{\alpha}^*$=4.094									
True initial		True trans.		MLEs trans.		StdErr	MLEs		Gradient
λ	2.8	λ^*	1.029	$\widehat{\lambda}^*$	0.847	0.169	$\widehat{\lambda}$	2.332	0.0062
σ_0	0.02	σ_0^*	-7.824	$\widehat{\sigma}_0^*$	-7.877	0.048	$\widehat{\sigma}_0$	0.019	0.0039
β	0.04	β^*	-3.219	$\widehat{\beta}^*$	-3.339	0.106	$\widehat{\beta}$	0.036	-0.0038
α	2.65	α^*	0.975	$\widehat{\alpha}^*$	1.033	0.107	$\widehat{\alpha}$	2.811	0.0976

True initial: denotes the 'True' simulated parameter values $\boldsymbol{\theta}$, True trans.: denotes the 'True' transformed simulated parameter values $\boldsymbol{\theta}^*$, MLEs trans: denote the maximum likelihood estimates $\widehat{\boldsymbol{\theta}}^*$ of the transformed parameters $\boldsymbol{\theta}^*$, StdErr: denotes the standard errors of the transformed parameters $\boldsymbol{\theta}^*$, MLEs: denote the maximum likelihood estimates $\widehat{\boldsymbol{\theta}}$ of the parameters $\boldsymbol{\theta}$, Gradient: denotes the value of the gradient of the log-likelihood evaluated at the parameter estimates.

Table 6: Maximum likelihood estimation results of the third simulation experiments ($T = 2000$) using Student-t errors.

Panel A: Starting values $\widehat{\lambda}^*$=0.095, $\widehat{\sigma}_0^*$=-9.210, $\widehat{\beta}^*$=-6.908, $\widehat{\alpha}^*$=3.689, $\widehat{v}^*$=1.609									
True initial		True trans.			MLEs trans.	StdErr	MLEs		Gradient
λ	2.8	λ^*	1.029	$\widehat{\lambda}^*$	0.839	0.169	$\widehat{\lambda}$	2.315	-0.0019
σ_0	0.02	σ_0^*	-7.824	$\widehat{\sigma}_0^*$	-7.874	0.049	$\widehat{\sigma}_0$	0.019	-0.0028
β	0.04	β^*	-3.219	$\widehat{\beta}^*$	-3.336	0.109	$\widehat{\beta}$	0.036	0.0002
α	2.65	α^*	0.975	$\widehat{\alpha}^*$	1.036	0.110	$\widehat{\alpha}$	2.818	-0.0031
				$\widehat{v}^*$	3.627	0.798	$\widehat{v}$	39.59	0.0003
Panel B: Starting values $\widehat{\lambda}^*$=–0.693, $\widehat{\sigma}_0^*$=-6.438, $\widehat{\beta}^*$=-3.507, $\widehat{\alpha}^*$=2.996, $\widehat{v}^*$=1.609									
True initial		True trans.			MLEs trans.	StdErr	MLEs		Gradient
λ	2.8	λ^*	1.029	$\widehat{\lambda}^*$	0.839	0.169	$\widehat{\lambda}$	2.315	-0.0002
σ_0	0.02	σ_0^*	-7.824	$\widehat{\sigma}_0^*$	-7.874	0.049	$\widehat{\sigma}_0$	0.019	0.0003
β	0.04	β^*	-3.219	$\widehat{\beta}^*$	-3.336	0.109	$\widehat{\beta}$	0.036	0.0000
α	2.65	α^*	0.975	$\widehat{\alpha}^*$	1.036	0.110	$\widehat{\alpha}$	2.818	-0.0001
				$\widehat{v}^*$	3.627	0.798	$\widehat{v}$	39.60	0.0009

True initial: denotes the 'True' simulated parameter values $\boldsymbol{\theta}$, True trans.: denotes the 'True' transformed simulated parameter values $\boldsymbol{\theta}^*$, MLEs trans: denote the maximum likelihood estimates $\widehat{\boldsymbol{\theta}}^*$ of the transformed parameters $\boldsymbol{\theta}^*$, StdErr: denotes the standard errors of the transformed parameters $\boldsymbol{\theta}^*$, MLEs: denote the maximum likelihood estimates $\widehat{\boldsymbol{\theta}}$ of the parameters $\boldsymbol{\theta}$, Gradient: denotes the value of the gradient of the log-likelihood evaluated at the parameter estimates.

Table 7: Maximum likelihood estimation results of the EUBanks index return series under the assumption of conditional normal and Student-t distribution.

Panel A: Normal errors					
	MLEs trans.	StdErr	MLEs		Gradient
$\widehat{\lambda}^*$	1.020	0.295	$\widehat{\lambda}$	2.774	-0.0017
$\widehat{\sigma}_0^*$	-7.742	0.047	$\widehat{\sigma}_0$	0.021	-0.2271
$\widehat{\beta}^*$	-3.231	0.221	$\widehat{\beta}$	0.039	-0.0866
$\widehat{\alpha}^*$	0.975	0.246	$\widehat{\alpha}$	2.652	0.0556
Panel B: Student-t errors					
	MLEs trans.	StdErr	MLEs		Gradient
$\widehat{\lambda}^*$	0.730	0.204	$\widehat{\lambda}$	2.075	0.0028
$\widehat{\sigma}_0^*$	-7.143	0.452	$\widehat{\sigma}_0$	0.028	0.0924
$\widehat{\beta}^*$	-3.690	0.170	$\widehat{\beta}$	0.025	-0.0267
$\widehat{\alpha}^*$	1.214	0.203	$\widehat{\alpha}$	3.366	0.8350
$\widehat{v}^*$	-1.164	0.572	$\widehat{v}$	2.312	0.0738

MLEs trans: denote the maximum likelihood estimates $\widehat{\boldsymbol{\theta}}^*$ of the transformed parameters $\boldsymbol{\theta}^*$, StdErr: denotes the standard errors of the transformed parameters $\boldsymbol{\theta}^*$, MLEs: denote the maximum likelihood estimates $\widehat{\boldsymbol{\theta}}$ of the parameters $\boldsymbol{\theta}$, Gradient: denotes the value of the gradient of the log-likelihood evaluated at the parameter estimates.

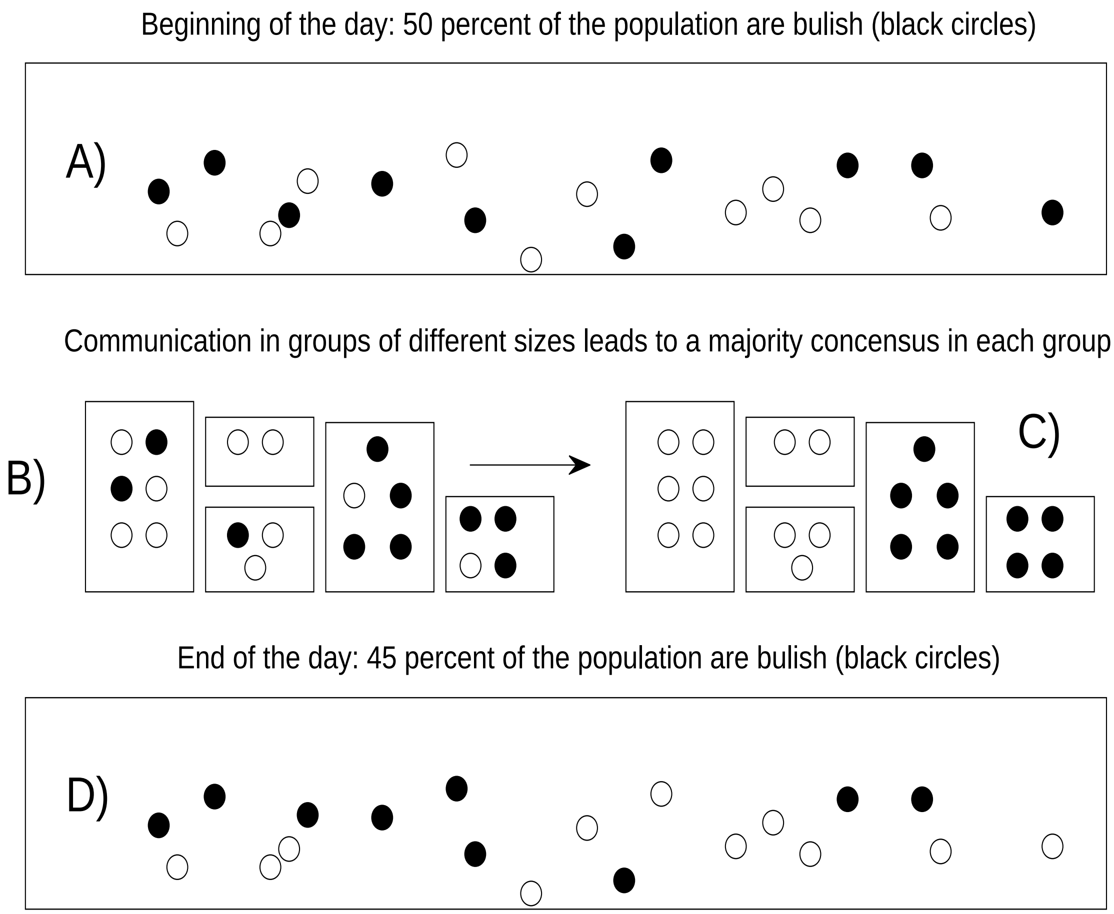

Figure 1: Changing the 'bullishness' in a population via communications in subgroups. A) At the beginning of a given day t a certain percentage $B(t)$ of bullishness, B) During the day communication takes place in random subgroups of different sizes, C) Illustrates the extreme case of complete polarization $m_{k,j} = \pm 1$ created by a majority rule in opinion. In general $m_{k,j} \simeq j/k$ corresponds to the neutral case where in average the opinion remains unchanged within a subgroup of size k, D) Due to the communication in different subgroups the "bullishness" at the end of the day is different from the beginning of the day.

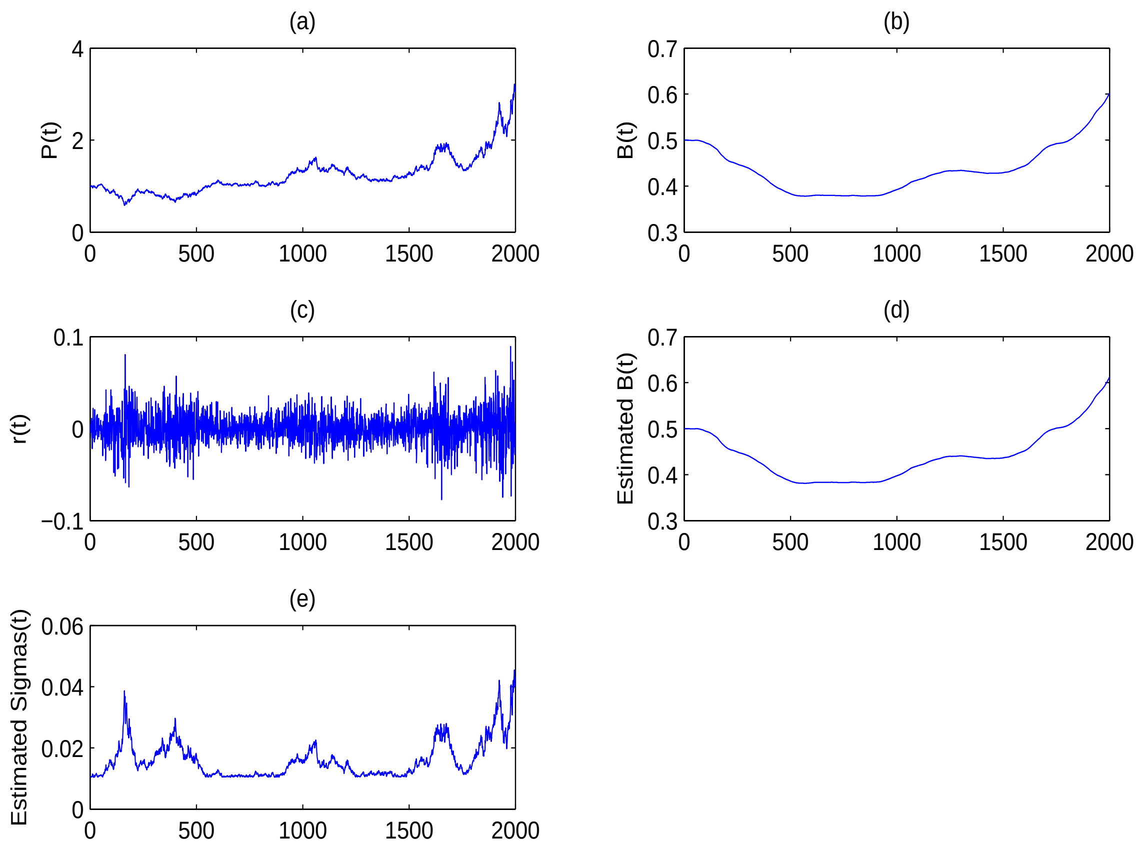

Figure 2: This figure presents simulated price data ($T = 2000$, first simulation experiment), returns and bullishness proportions, as well as the corresponding estimated conditional volatilities and bullishness proportions based on *normal* errors. (a) Simulated prices $P(t)$, (b) Simulated bullishness proportions $B(t)$, (c) Simulated price returns $r(t)$, (d) Estimated bullishness proportions $\widehat{B}(t)$, (e) Estimated conditional volatilities $\widehat{\sigma}(t)$.

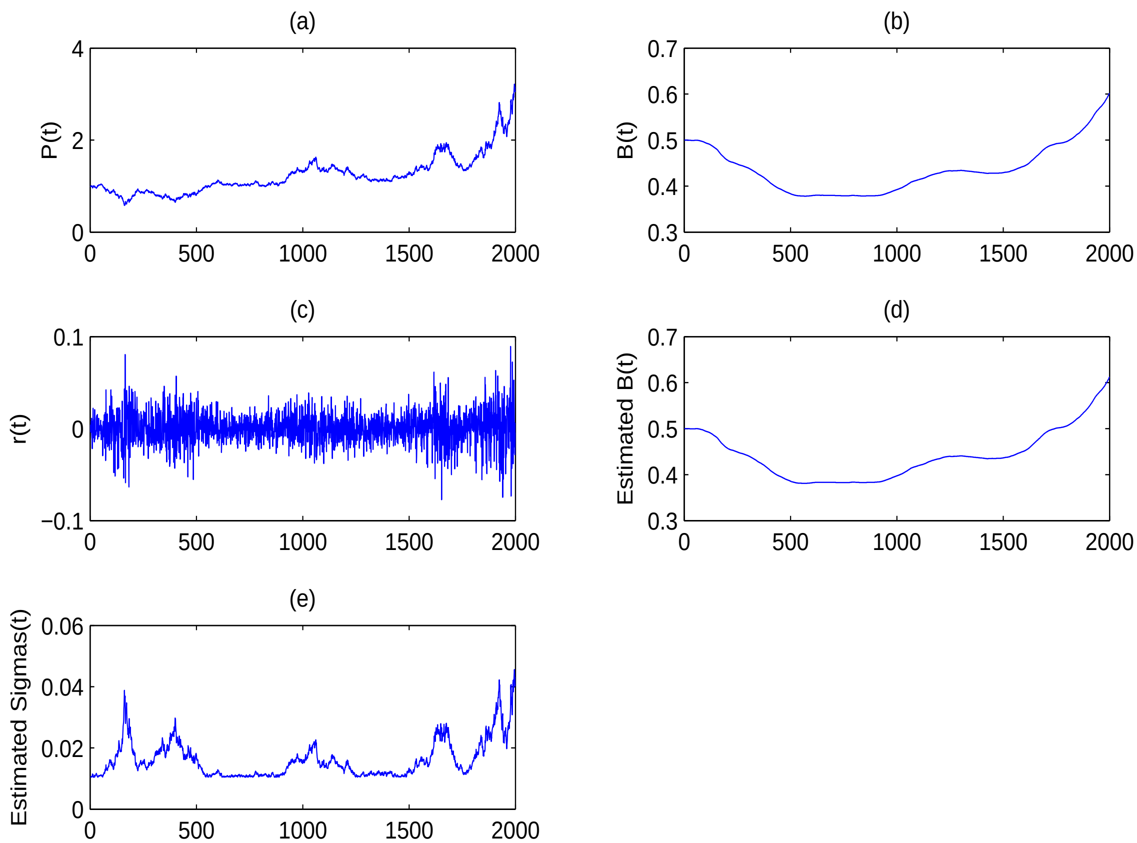

Figure 3: This figure presents simulated price data ($T = 2000$, first simulation experiment), returns and bullishness proportions, as well as the corresponding estimated conditional volatilities and bullishness proportions based on *Student-t* errors. (a) Simulated prices $P(t)$, (b) Simulated bullishness proportions $B(t)$, (c) Simulated price returns $r(t)$, (d) Estimated bullishness proportions $\widehat{B}(t)$, (e) Estimated conditional volatilities $\widehat{\sigma}(t)$.

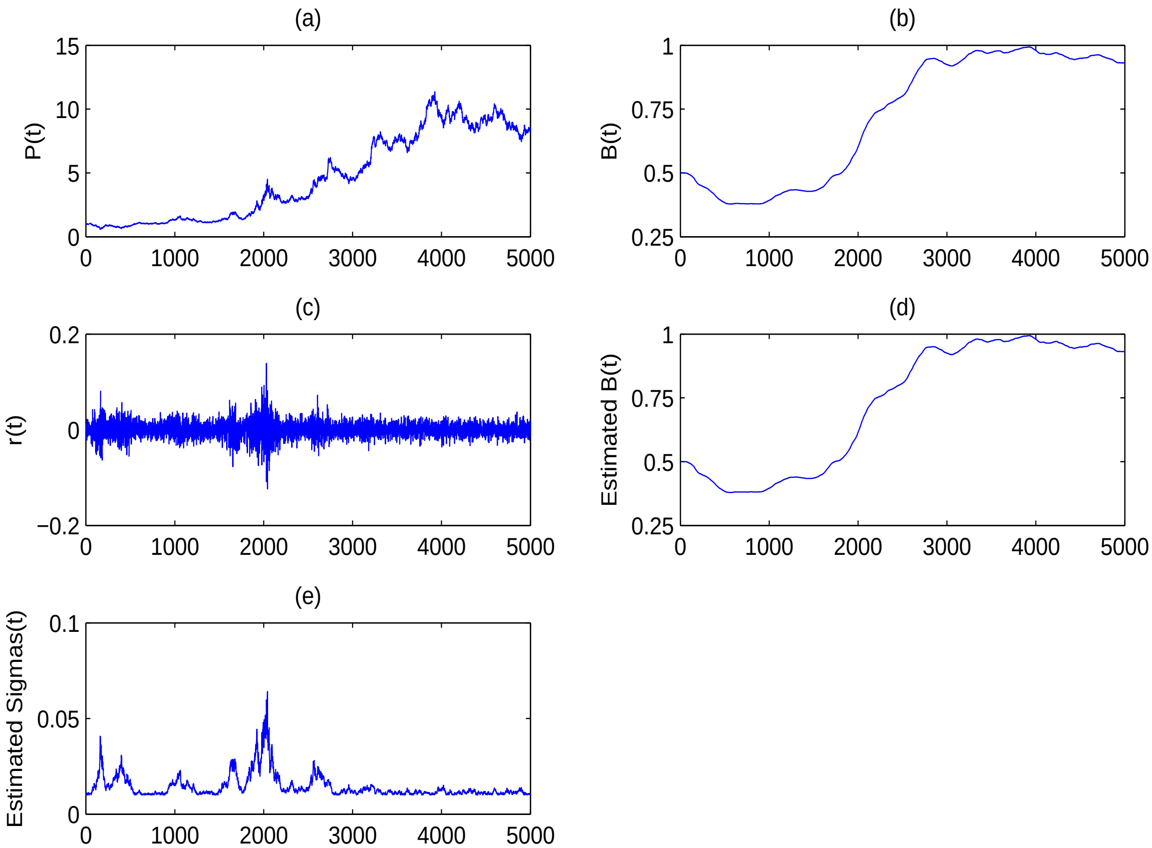

Figure 4: This figure presents simulated price data ($T = 5000$, second simulation experiment), returns and bullishness proportions, as well as the corresponding estimated conditional volatilities and bullishness proportions based on *normal* errors. (a) Simulated prices $P(t)$, (b) Simulated bullishness proportions $B(t)$, (c) Simulated price returns $r(t)$, (d) Estimated bullishness proportions $\widehat{B}(t)$, (e) Estimated conditional volatilities $\widehat{\sigma}(t)$.

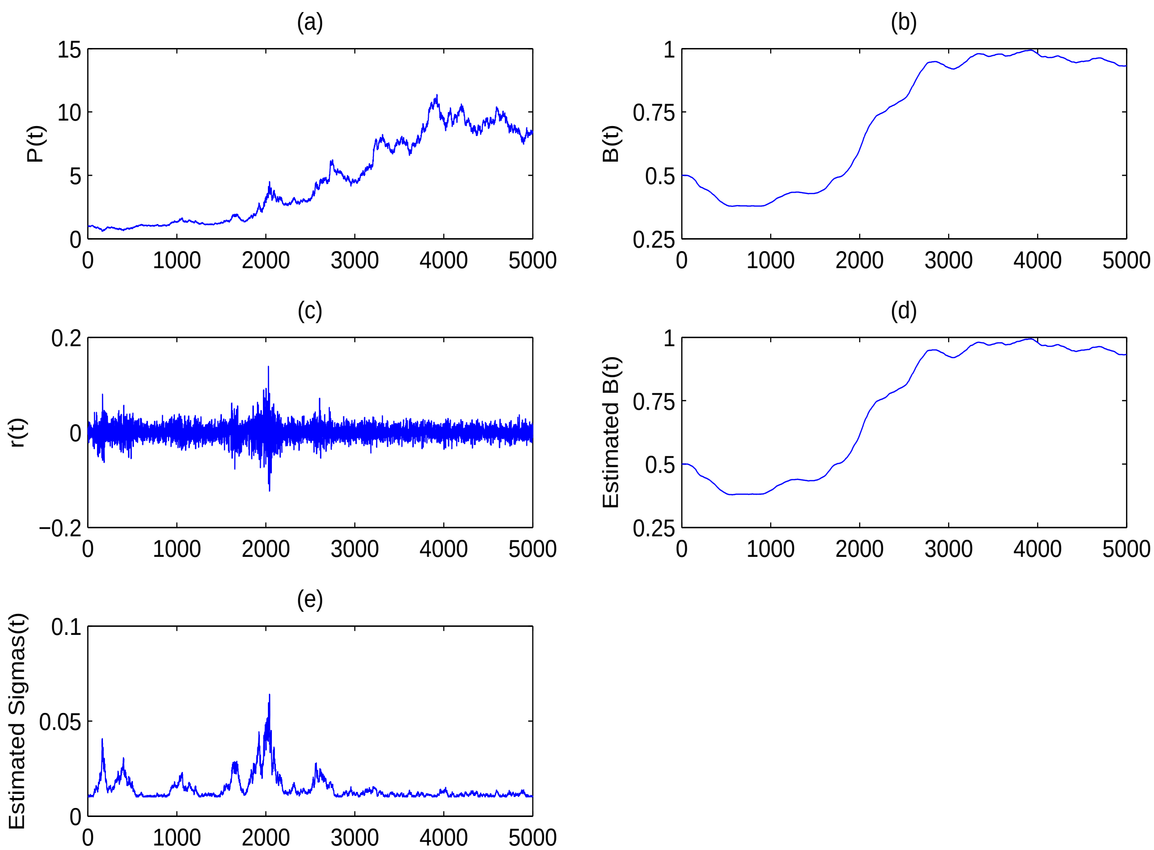

Figure 5: This figure presents simulated price data ($T = 5000$, second simulation experiment), returns and bullishness proportions, as well as the corresponding estimated conditional volatilities and bullishness proportions based on *Student-t* errors. (a) Simulated prices $P(t)$, (b) Simulated bullishness proportions $B(t)$, (c) Simulated price returns $r(t)$, (d) Estimated bullishness proportions $\widehat{B}(t)$, (e) Estimated conditional volatilities $\widehat{\sigma}(t)$.

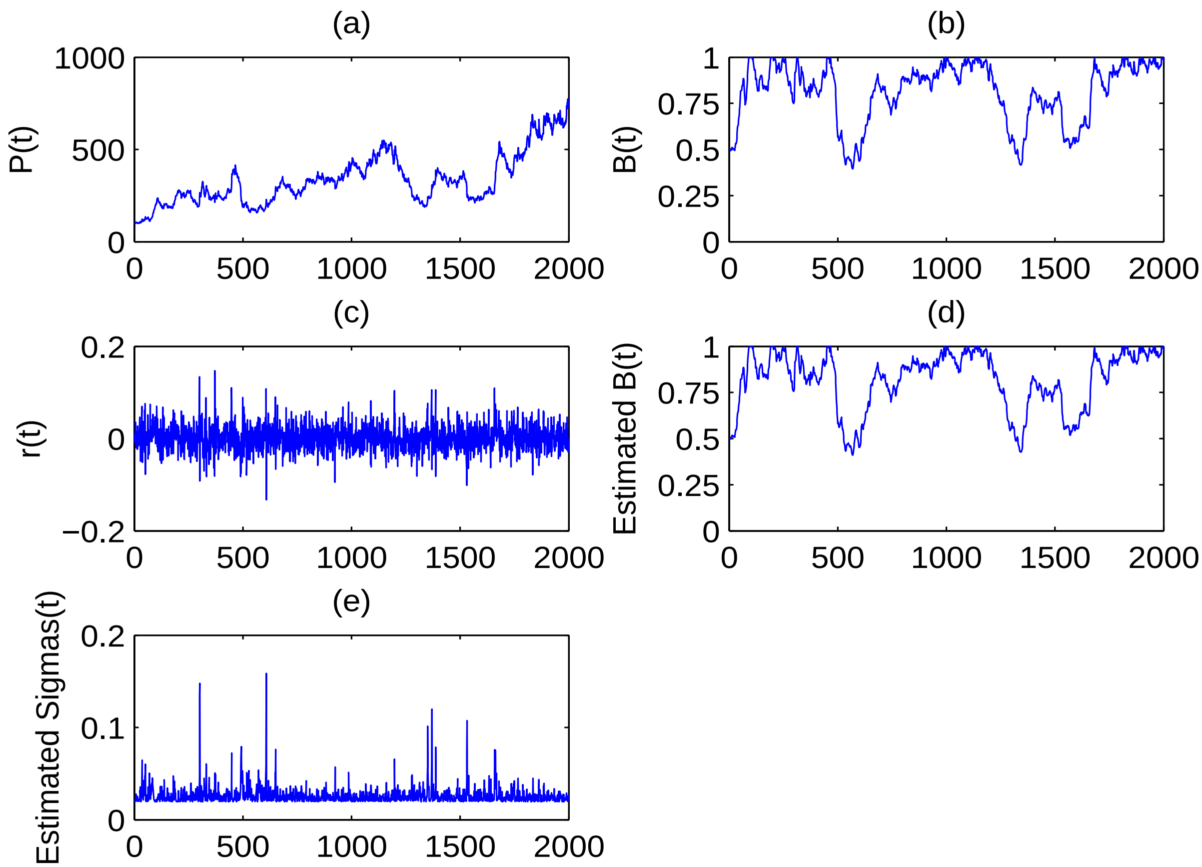

Figure 6: This figure presents simulated price data ($T = 2000$, third simulation experiment), returns and bullishness proportions, as well as the corresponding estimated conditional volatilities and bullishness proportions based on *normal* errors. (a) Simulated prices $P(t)$, (b) Simulated bullishness proportions $B(t)$, (c) Simulated price returns $r(t)$, (d) Estimated bullishness proportions $\widehat{B}(t)$, (e) Estimated conditional volatilities $\widehat{\sigma}(t)$.

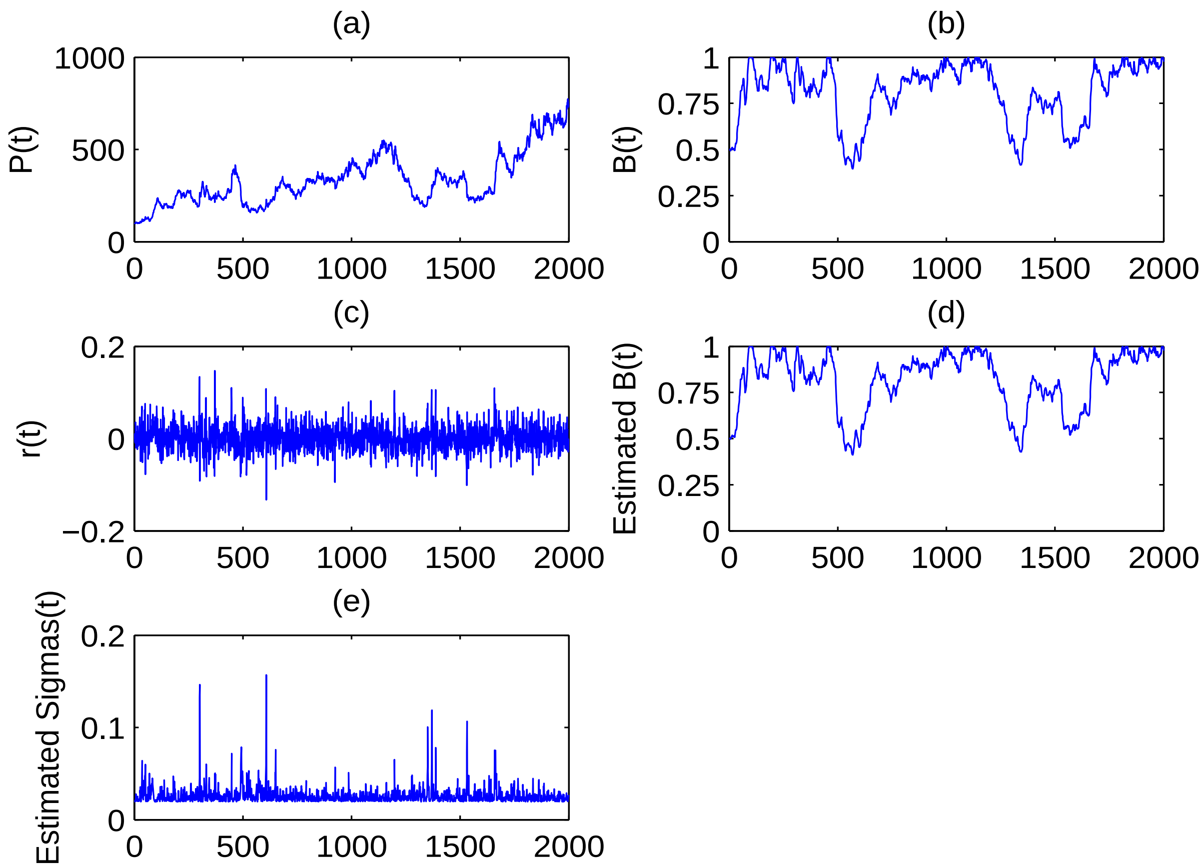

Figure 7: This figure presents simulated price data ($T = 2000$, third simulation experiment), returns and bullishness proportions, as well as the corresponding estimated conditional volatilities and bullishness proportions based on *Student-t* errors. (a) Simulated prices $P(t)$, (b) Simulated bullishness proportions $B(t)$, (c) Simulated price returns $r(t)$, (d) Estimated bullishness proportions $\widehat{B}(t)$, (e) Estimated conditional volatilities $\widehat{\sigma}(t)$.

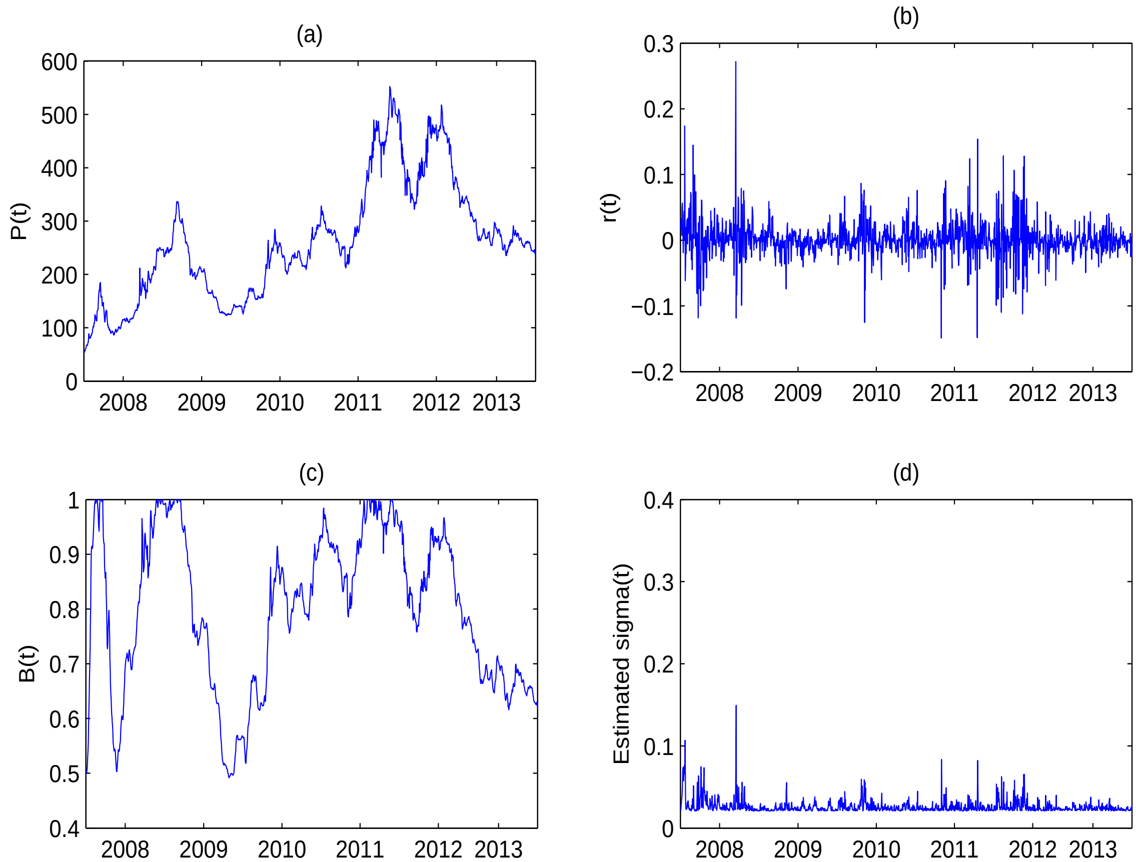

Figure 8: This figure presents EUBanks index prices and returns, as well as the corresponding estimated conditional volatilities and bullishness proportions under the assumption of conditional *normal* distribution. (a) EUBanks index price $P(t)$, (b) EUBanks index returns $r(t)$, (c) Estimated bullishness proportions $\widehat{B}(t)$, (d) Estimated conditional volatilities $\widehat{\sigma}(t)$.

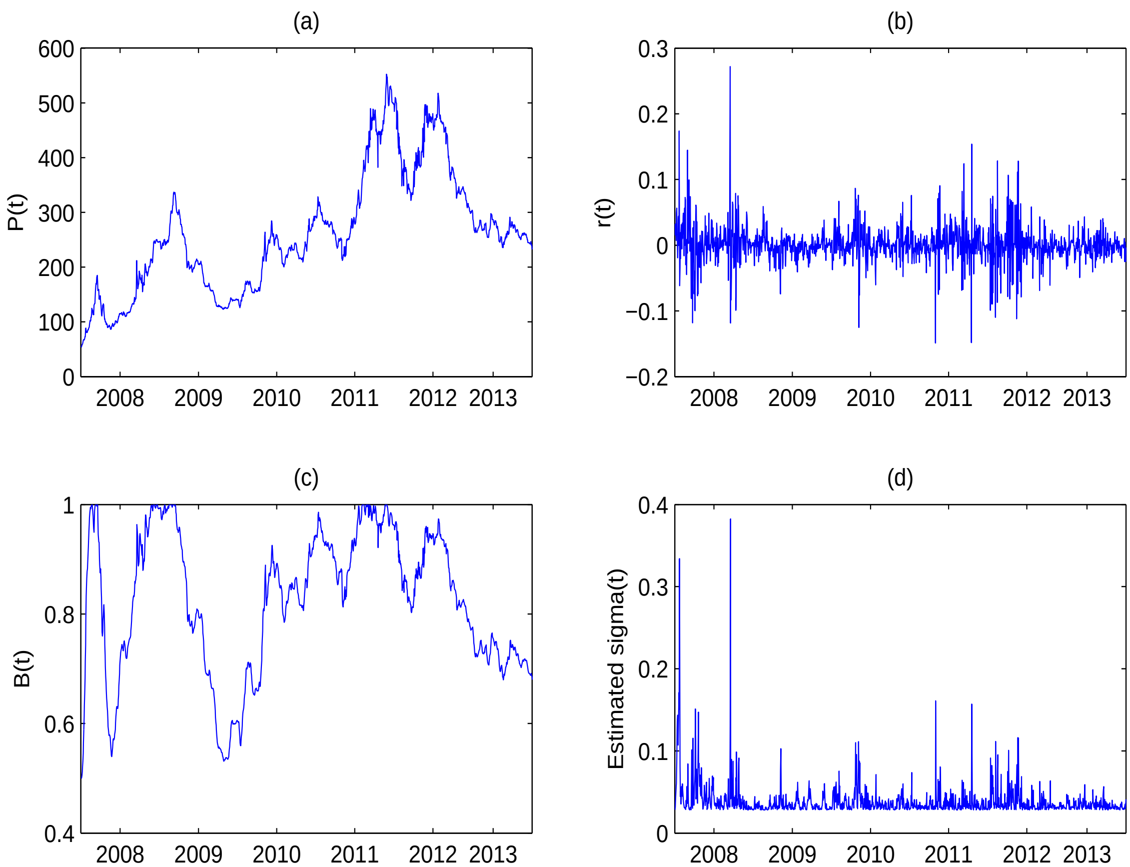

Figure 9: This figure presents EUBanks index prices and returns, as well as the corresponding estimated conditional volatilities and bullishness proportions under the assumption of conditional *Student-t* distribution. (a) EUBanks index price $P(t)$, (b) EUBanks index returns $r(t)$, (c) Estimated bullishness proportions $\widehat{B}(t)$, (d) Estimated conditional volatilities $\widehat{\sigma}(t)$.

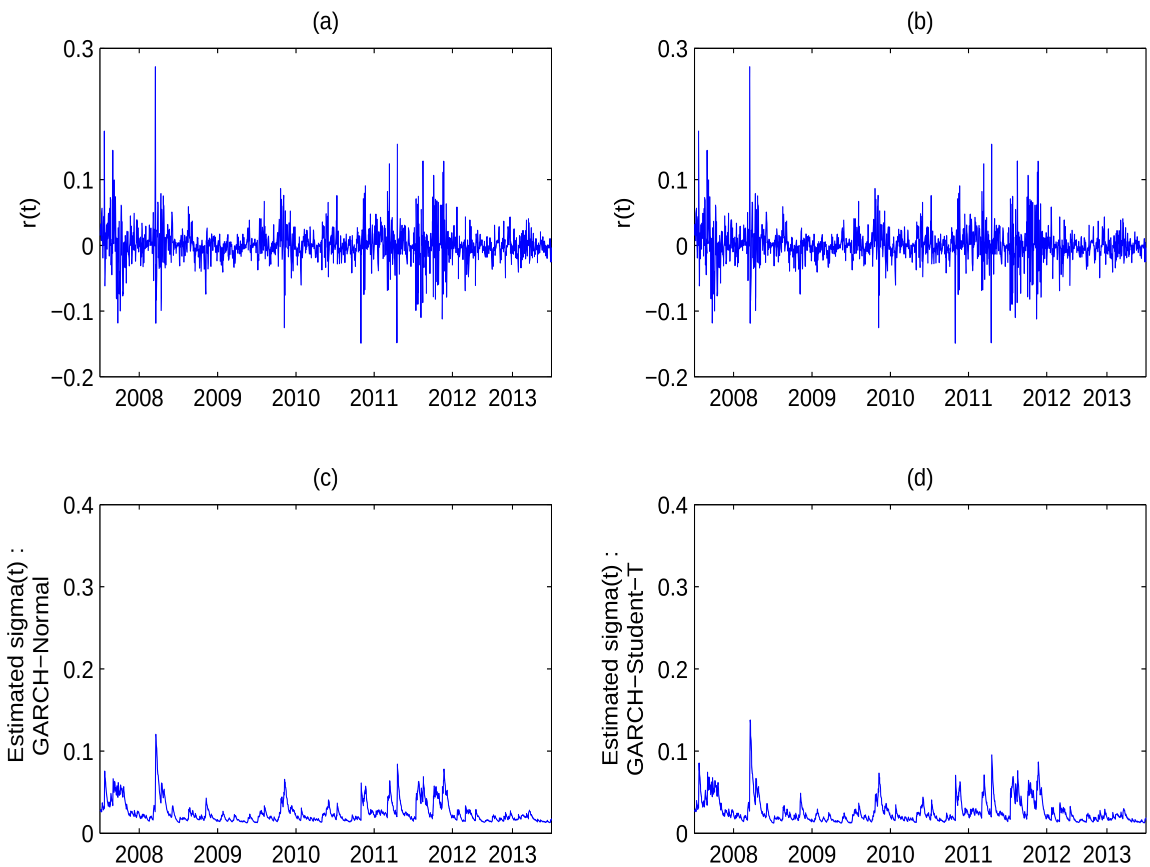

Figure 10: This figure presents EUBanks index returns, as well as the corresponding estimated conditional volatilities based on a GARCH(1,1) model under the assumption of conditional *normal* and *Student-t* distribution. (a) EUBanks index returns $r(t)$, (b) EUBanks index returns $r(t)$, (c) Estimated conditional volatilities $\widehat{\sigma}(t)$ based on a GARCH(1,1) model and *normal* errors, (d) Estimated conditional volatilities $\widehat{\sigma}(t)$ based on a GARCH(1,1) model and *Student-t* errors.