

Signal Extraction Can Generate Volatility Clusters

From IID Shocks

Prasad V. Bidarkota*

Department of Economics, Kansas State University, Manhattan, KS 66506-4001, USA

J. Huston McCulloch

Department of Economics, The Ohio State University, Columbus, OH 43210, USA

Abstract: We develop a framework in which information about firm value is noisily observed. Investors are then faced with a signal extraction problem. Solving this would enable them to probabilistically infer the fundamental value of the firm and, hence, price its stocks. If the innovations driving the fundamental value of the firm and the noise that obscures this fundamental value in observed data come from non-Gaussian thick-tailed probability distributions, then the implied stock returns could exhibit volatility clustering. We demonstrate the validity of our hypothesis with a simulation study.

Key phrases: stock returns; volatility clusters; GARCH processes; signal extraction; thick-tailed distributions; simulations.

JEL Codes: C22, E31, C53

November 17, 2002

* Corresponding author. Address for correspondence: Department of Economics, 245 Waters Hall, Kansas State University, Manhattan, KS 66506-4001, USA;
Tel: +1-(785)-532-4578; Fax: +1-(785)-532-6919; E-mail: pbidarko@ksu.edu

Introduction

It has now been well established in the empirical finance literature that returns on many financial assets exhibit the phenomenon of volatility clustering (see, for instance, Pagan and Schwert, 1990). By this we understand that large shocks in asset returns tend to be followed by large shocks (of either sign) and small shocks tend to be followed by small shocks (Mandelbrot, 1963). The autoregressive conditional heteroskedasticity (ARCH) and related class of models (Engle, 1982) have been developed to capture this type of phenomenon in asset returns.

Since then, several papers have attempted to characterize more extensively both the univariate statistical properties of volatility dynamics as well as its relationship with other economic variables. Several other papers have also attempted to provide an understanding of the underlying economic mechanism that might generate such features of returns volatility. See, for instance, Peng and Xiong (2001) for a brief discussion of this literature.

However, although twenty years have passed since the publication of the seminal paper on ARCH by Engle (1982), there is still no widely accepted economic explanation for why returns exhibit the basic phenomenon of volatility clustering. An early idea in French and Roll (1986) relates volatility to the arrival of information and the reaction of traders to this information. Bookstaber and Pomerantz (1989) develop a model of market volatility based on this idea, assuming that information arrives in ‘discrete packets’ and that it takes time for the market to digest this information and react to it. An extension of this work is the recent paper by Peng and Xiong (2001), wherein the effort required to process newly arriving information (assumed constant in Bookstaber and Pomerantz,

1989) is endogenized, subject to capacity constraints on the information processing capabilities of investors. The idea that market participants face information processing capacity constraints originates with Sims (2002).

In his paper, Sims (2002) argues that outcomes resulting from information flow constraints would resemble those from a situation where market participants face a signal extraction problem.

In this study, we address the basic question: why does the volatility of returns on risky assets vary over time, and more specifically, why does this volatility exhibit clustering over time? We assume that investors do not observe the fundamental value of a firm but only observe noisy data that contain signals about firm performance. They are then faced with a signal extraction problem; a problem of trying to filter the observed noisy data in order to extract the fundamental value of the firm. Investors then use that extracted information to price stocks. Our main contention here is that if the innovations driving the fundamental value of the firm and the noise that obscures this fundamental value come from non-Gaussian thick-tailed probability distributions, then the implied stock returns could exhibit volatility clustering. This is true even though the inherent exogenous process driving the fundamental value of the firm over time as well as the noise in the accounting data that obscures the fundamental value do not exhibit this phenomenon.

This is important because one could posit instead (as in Bookstaber and Pomerantz, 1989, or Peng and Xiong, 2001) that information regarding the fundamental value of a firm arrives in clusters. If that is true, then it is perhaps not surprising to find returns exhibiting volatility clustering. Since we have little reason to believe that

information usually arrives in clusters, the challenge is to demonstrate, using a model where exogenous shocks arrive in an independently and identically distributed (iid) fashion, that returns exhibit clusters of volatility.

In this sense, we view our work as close in spirit, and methodology, to that of den Haan and Spear (1998). They provide an explanation for volatility clustering in real interest rates using an equilibrium economic model with heterogeneous agents and incomplete markets driven by iid disturbances that display no volatility clustering. As in that study, we attempt to validate our model that provides a mechanism for volatility clustering in returns on risky assets by comparing the characteristics of simulated returns data implied by our model with the well-documented characteristics of returns data observed in real financial markets.

It is difficult to provide intuition here for the exact mechanism at work in our model that makes this phenomenon happen. We therefore postpone an elaboration on this issue to the penultimate section. Prior to that, we formally set out in section 2 the information framework of our model and the associated signal extraction problem. In section 3, we discuss how to obtain stock prices and returns in our model. In section 4, we examine simulated stock returns implied by our model to see whether or not they display volatility clusters. In section 5, we provide intuition for our simulation results. The final section concludes with a summary and some observations on our study.

2. Information Framework and the Signal Extraction Problem

Section 2.1 outlines the information that investors in our model observe and a general framework they use for filtering that information. Section 2.2 describes briefly

the solution to the signal extraction problem. Section 2.3 demonstrates the behavior of the filter density within a simulation setup.

2.1. Information Framework

Suppose that x_t is the logarithm of the unobserved fundamental value of the firm and that y_t is an observable series that reflects x_t with noise. For instance, y_t could include, among other things, the accounting data of the firm, news reports on firm performance, and relevant macroeconomic data. Then, we have:

$$y_t = x_t + \varepsilon_t \quad (1)$$

Here, ε_t is the noise in the observed data that obscures the (logarithm of the) fundamental value of the firm (per share) at time t .

Although investors do not observe the fundamental value of the firm x_t , they are able to infer it probabilistically from the noisy observed data through a filtering (or signal extraction) process. In order to make filtering operational, investors need a model for the law of motion governing the dynamics of how the fundamental value of the firm evolves over time. Assume that investors use a simple random walk without drift as the governing law of motion for x_t :

$$x_t = x_{t-1} + \eta_t \quad (2)$$

Using Equations (1) and (2), investors perform a filtering (or signal extraction) procedure on the noisy observed data that enables them to infer:

$$p\{x_t | Y_t\}$$

where $Y_t \equiv \{y_t, y_{t-1}, \dots, y_1\}$ is the entire history of noisily observed data available to date. Here, $p\{A|B\}$ denotes the conditional probability density of event A given that event B has occurred.

2.2. Non-Gaussian Signal Extraction

When the disturbances ε_t and η_t in Equations (1) and (2) are both non-Gaussian, this is a non-Gaussian filtering situation. Appendix A describes the non-Gaussian probability distributions used in this paper. Under non-Gaussian filtering, the exact probability distribution of the filter density $p\{x_t | Y_t\}$ is also non-Gaussian and is given by the Sorenson-Alspach (1971) recursive formulae (see Harvey (1992), p.162-165). Appendix B reproduces these recursive formulae and also provides further details on non-Gaussian filtering. In general, the filter density cannot be fully described by its mean and variance alone. The entire distribution can be approximated by numerically evaluating the density at a set of abscissa for x_t . Appendix B provides details on numerical evaluation of the filter density.

Having obtained the filter density $p\{x_t | Y_t\}$ on a set of grid points for x_t , we can numerically compute moments of the filter density. We discuss in section 3 how these moments can be used to determine stock prices and stock returns.

2.3. Non-Gaussian Filter Density

In this subsection, we demonstrate that if the observational noise and signal shock, ε_t and η_t in Equations (1) and (2) above, are drawn from thick-tailed non-

Gaussian probability distributions, then the filter density $p\{x_t | Y_t\}$ can exhibit volatility clustering even though the shocks themselves are independently and identically distributed (iid).

To illustrate this phenomenon, we undertake a simulation study. We draw random numbers for ε_t in Equation (1) from the symmetric stable distribution $S_\alpha(0,1)$ and η_t in Equation (2) from the symmetric stable distribution $S_\alpha(0,\kappa)$, where κ is the signal-to-noise scale ratio.¹ Assuming that the initial value of x_t in Equation (2) is zero, that is $x_0 = 0$, we then use the simulated η_t series to generate a sequence $\{x_t, t = 1, 2, \dots, T\}$ using Equation (2). We use the simulated ε_t series and Equation (1) to generate a sequence $\{y_t, t = 1, 2, \dots, T\}$. In Figure 1, we plot the simulated shocks η_t and ε_t along with the raw observable data y_t .

With the simulated sequence $\{y_t, t = 1, 2, \dots, T\}$, we estimate the following model:

$$y_t = x_t + \varepsilon_t, \quad \varepsilon_t \sim S_\alpha(0, c) \quad (3)$$

$$x_t = x_{t-1} + \eta_t, \quad \eta_t \sim S_\alpha(0, \rho c). \quad (4)$$

¹ Appendix A provides a brief description of symmetric stable distributions and McCulloch (1996a) a comprehensive survey on the financial applications of these distributions. For generating random numbers from the symmetric stable distribution $S_\alpha(0,1)$, we use the GAUSS program written by J. Huston McCulloch and archived at <http://www.econ.ohio-state.edu/jhm/jhm.html>. For the simulations we use $\alpha = 1.8$ and $\kappa = 10$.

Estimation is done by maximum likelihood. The likelihood function is given in Equation (B4) of Appendix B.

Parameter estimates are presented in the first row of Table 1. The characteristic exponent α is estimated to be higher than the true value at 1.93, but the scale parameter c and the signal-to-noise scale ratio ρ are both estimated to be lower than their true value at 6.39 and 1.50, respectively.

In Figure 2, we plot the estimated mean and standard deviation of the filter density $p\{x_t | Y_t\}$.² Looking at this figure and Figure 1 closely, it is clear that the filter mean tracks the observable data y_t quite well. Also, the filter standard deviation jumps up whenever a big realization (positive or negative) of either shock η_t or ε_t occurs. It is hard to tell from Figure 2, however, whether the jump in the filter standard deviation lingers or not after a big shock has occurred.

In order to ascertain whether jumps in the filter standard deviation persist over time or not, we plot in Figure 3 the sample autocorrelations and partial autocorrelations of the squared filter errors, defined as the squared differences between the filter means and the x_t series that they estimate. It is obvious from this figure that the squared filter errors are indeed autocorrelated. This is indicative of volatility clustering in the filter density, since the innovations to x_t are iid.

² It can be shown that so long as $\alpha > 1/2$, the filter density $p\{x_t | Y_t\}$ has finite variance for $t \geq 2$ in the case of the local level model given in Equations (1) and (2), despite the infinite variances of the stable shocks.

3. Stock Prices and Returns with Signal Extraction

In this section, we discuss how to compute stock prices and returns in our model economy outlined in section 2.

In our framework, the mean of the filter density must follow a martingale. To see this, note that we have from the Law of Iterated Expectations:

$$E[E(x_{t+1} | Y_{t+1}) | Y_t] = E(x_{t+1} | Y_t). \quad (5)$$

From Equation (2), we have:

$$E(x_{t+1} | Y_t) = E(x_t | Y_t). \quad (6)$$

Therefore,

$$E[E(x_{t+1} | Y_{t+1}) | Y_t] = E(x_t | Y_t). \quad (7)$$

In our model economy, the precise price of the asset will depend on how much systematic risk is in the filter distribution uncertainty. If this risk is entirely idiosyncratic, the price will be the expected fundamental value (in levels). However, if investors are concerned, say, that accounting rules may distort the value of all firms in some common way whose magnitude is unknown, or if some of the relevant data is macroeconomic data, the risk may be perceived as systematic, and then will be priced. We do not know exactly how much this gets priced, but we can just say that the market price will "reflect" (if not equal) the mean of the filter density (even in logs). Calling this the quasi-price, the quasi-returns then are just the changes in the mean of the filter density (abstracting from expected returns, dividends, and changing risk premia).

In Appendix C, we formally derive asset prices and returns using a simple asset-pricing model, within a Gaussian setting and using a constant relative risk aversion (CRRA) utility function. Equation (C14) in the appendix gives an exact analytical

formula for the stock price in such a setting, which can be seen to be a function of the filter mean and filter variance (apart from investors' preference parameters). Equation (C15) gives an exact formula for stock returns. These are made up of two components. The first is just the changes in the mean of the filter density. The second is a linear function of changes in the variance of the filter density. Thus, the quasi-returns are simply the first component of returns in this setting.

In the next section, we examine simulated quasi-returns implied by our simple model to see whether they exhibit volatility clustering using a variety of formal techniques.

4. Examination of Quasi-Stock Returns

Section 4.1 reports some preliminary statistics on these returns. Section 4.2 estimates a standard GARCH model for these returns. Section 4.3 modifies the standard GARCH model by assuming non-Gaussian innovations.

4.1. Preliminary Study of Quasi-Returns

We continue with the simulation study begun in section 2.3. There, we performed signal extraction on simulated observed data, and obtained the filter mean and standard deviation for time periods $t = 1, 2, \dots, 5001$. The simulated data and the moments of the filter density are plotted in Figures 1 and 2, and were discussed in section 2.3.

From these 5001 filter means, we compute 5000 quasi-returns (as changes in the filter means), referred to simply as returns in the rest of the paper for convenience. We discard the first 3000 returns so as to ensure that any effects from the startup of the filter

are fully eliminated. In what follows, we evaluate the characteristics of the remaining 2000 returns in order to verify whether or not they exhibit volatility clusters.

In Figure 4 we plot the implied stock returns. The simple model that we have set up in section 2 is designed only to provide an understanding of why returns on risky assets exhibit volatility clustering. Without identifying the observable data y_t in our framework with concrete information from real financial markets, we do not know what process to use to generate artificial data for y_t in our simulations. For the purposes of figuring out whether or not our model generates volatility clustering, this is not a drawback. However, this also means that the only dimension along which we should test to see whether our model-implied returns are similar to observed returns on stocks is in their volatility clustering features. Consequently, for the purposes of validating our model, it is immaterial what the mean of implied returns is, as well as the autocorrelations of levels of returns. It is also immaterial whether or not implied returns exhibit fat tails, as has been well documented in the literature.

4.2. A GARCH Model of Quasi>Returns

To formally investigate whether the implied returns from the filtering mechanism exhibit volatility clustering, we estimate a GARCH model for these returns. This model takes the following form:

$$r_t = \mu + \zeta_t, \quad \zeta_t \sim c_t z_t, \quad z_t \sim \text{iid } N(0, \sqrt{2}) \quad (8)$$

$$c_t^2 = \omega + \beta c_{t-1}^2 + \delta |r_{t-1} - \mu|^2 \quad (9)$$

We restrict $\omega > 0$, $\beta \geq 0$ and $\delta \geq 0$. For simplicity, we select the GARCH(1,1) specification above. This has also by far been the most popular parameterization used to describe stock return volatility.³

The top panel of Table 2 (labeled Stable Data) reports results from estimating this model as well as a restricted homoskedastic model where the scales c_t are constant (equal to c). The GARCH parameter β is estimated to be 0.49 and the ARCH term δ is 0.02, indicating that the volatility of returns is quite persistent but only mildly sensitive to the magnitude of the past innovations to returns. The likelihood ratio (LR) test for the null hypothesis of no GARCH (test for $\beta = \delta = 0$) is reported in the last column of Table 2. Homoskedasticity is easily rejected in favor of GARCH(1,1), and the evidence is overwhelming.

Figure 5 plots the estimated scales from the model in Equations (8) and (9). When seen in conjunction with the raw observable data y_t and the behavior of the filter mean and standard deviation plotted in Figures 1 and 2 respectively, the figure clearly demonstrates both time variation in the volatility of implied returns and its sensitivity to large shocks in the observable data.

³ Pagan and Schwert (1990) fit a GARCH(1,2) model for monthly returns from 1834-1925, while French, Schwert and Stambaugh (1987) fit a similar model to monthly returns from 1928-1984. Both these studies find only weak effects of the second MA term. See also Pagan (1996).

4.3. A GARCH-Stable Model of Quasi>Returns

The quasi-returns in our model are unlikely to be Gaussian. Therefore, our model of volatility in Equations (8) and (9) is likely to be misspecified. We therefore modify that model by assuming that the innovations are symmetric stable. This model takes the following form:

$$r_t = \mu + \zeta_t, \quad \zeta_t \sim c_t z_t, \quad z_t \sim \text{iid} S_\alpha(0,1) \quad (10)$$

$$c_t^\alpha = \omega + \beta c_{t-1}^\alpha + \delta |r_{t-1} - \mu|^\alpha. \quad (11)$$

As before, we restrict $\omega > 0$, $\beta \geq 0$ and $\delta \geq 0$. When ζ_t is normal (that is, when $\alpha = 2$), this model reduces to the familiar GARCH-normal process of section 4.2. Once again, for simplicity, we select the GARCH(1,1) specification. A GARCH-stable model similar to the one given in Equation (11) has been estimated for bond returns by McCulloch (1985) and for daily foreign currency returns by Liu and Brorsen (1995).

The top panel of Table 3 reports results from estimating this model as well as a restricted homoskedastic model where the scales c_t are constant (equal to c). The characteristic exponent α is estimated to be 1.55, indicating highly non-normal leptokurtic behavior. The volatility persistence parameter β is estimated to be lower than in the GARCH-normal case at 0.25 but the ARCH term δ is higher at 0.07. The likelihood ratio (LR) test for the null hypothesis of no GARCH (test for $\beta = \delta = 0$) is once again reported in the last column of Table 3. Once again, homoskedasticity is strongly rejected in favor of GARCH(1,1), although the LR test statistic is now substantially smaller than in the GARCH-normal case. Overall, the implied returns exhibit strong volatility clustering features, and this behavior persists even after

accounting for leptokurtosis in implied returns with symmetric stable innovations (see Ghose and Kroner, 1995, and Groenendijk et al, 1995, for an elaboration on this issue).

Figure 6 plots the estimated scales from the model in Equations (10) and (11) using the implied returns data. The figure clearly illustrates the time-varying behavior of volatility in implied returns.

The model in Equations (10) and (11) is estimated with monthly value-weighted CRSP real stock returns (with dividends) over the 1953-1994 period in Bidarkota and McCulloch (2002). From that study, the volatility persistence parameter β is estimated to be 0.80 and δ is estimated to be 0.04. The LR test statistic $\beta = \delta = 0$ is found to be 16.10. Thus, in our framework, the volatility persistence in returns is too low and the ARCH parameter is about right.

The bottom panels of Tables 2 and 3 (labeled Gaussian Data) report results from estimating the GARCH-normal and GARCH-stable models with implied returns obtained by filtering simulated data drawn from Gaussian distributions for both ε_t and η_t , respectively. In this case, the Kalman filter is the optimal estimator (see Harvey, 1992, chapter 3) and Appendix B provides some details on the estimation of the filter density in this case. In summary, the estimates reported in Tables 2 and 3 indicate that implied returns from filtering Gaussian data are Gaussian and homoskedastic. Specifically, these returns display no volatility clusters unlike the implied returns from filtering non-Gaussian symmetric stable data.

5. Why Filtering May Generate Volatility Clusters

In this section, we provide some intuition that helps us to understand the simulation results. In section 5.1, we discuss why Gaussian signal shocks driving the firm fundamentals and Gaussian observational noise in the data will likely not lead to volatility clustering in implied returns. In section 5.2, we elucidate why non-Gaussian signal shocks and observational noise would likely lead to volatility clustering.

5.1. Why Gaussian Filtering Does Not Generate Volatility Clusters

We can reason intuitively why volatility clusters are not likely when the firm fundamentals and observational noise are both Gaussian. In this case, the state space set up of Equations (1) and (2) reduces to a linear Gaussian framework. Appendix II provides details on filtering in a Gaussian linear state space model. Specifically, the celebrated Kalman filter is the optimal estimator of the unobserved fundamental value in this setup. From the properties of the Kalman filter (see, for instance, Harvey (1992), chapter 3), we know that the filter variance responds only to the variances of the signal shock and observational noise, η_t and ε_t respectively. Specifically, the filter variance does not respond to any outliers that may be present in the observations y_t . Given that, some time after startup, the filter variance will stabilize to a constant value as long as the signal and noise variances are assumed to be time-invariant.

Let us now consider how the mean of the Kalman filter $E(x_t | Y_t)$, that we have taken to be the (logarithm of the) quasi-stock price, behaves over time. We can express this quantity at any time t as a weighted average of the filter mean at some specific time

in the past $t-j$ and a linear combination (with declining weights on the past observations) of all the intervening observations up to the present time $\{y_{t-j+1}, y_{t-j+2}, \dots, y_t\}$. After the Kalman filter has stabilized and with a large enough value for j , the weights become virtually time-invariant. For a sufficiently large j , the weight on the past filter mean $E(x_{t-j} | Y_{t-j})$ becomes negligibly small. We can then view the filter mean at any time t , $E(x_t | Y_t)$, as a linear combination with constant weights declining into the past, of past observations y_t .

If the observations are generated by a homoskedastic process, then this linear combination will also behave as a homoskedastic process and specifically will not exhibit any clusters of volatility. Of course, if information about firm performance itself arrives in clusters (that is, if the observed data y_t itself exhibits volatility clustering) then the filter mean will also exhibit clusters, although these would be heavily damped because the filter mean responds to new information with a weight less than one.

Thus, when investors observe information about firm performance (such as accounting data) that contains signals about the fundamental value of the firm (per share), and both the firm fundamentals and noise follow Gaussian stochastic processes, the resulting stock returns implied by investor behavior based on signal extraction will not exhibit volatility clustering.

5.2. Why Non-Gaussian Filtering Can Generate Volatility Clusters

With non-Gaussian shocks driving firm fundamentals and noise in observed data, the state space set up of Equations (1) and (2) reduces to a linear non-Gaussian

framework. In this case, the filter density $p\{x_t | Y_t\}$ responds strongly to new observations y_t and never stabilizes even when the signal and noise variances are time-invariant. As clearly demonstrated in Bidarkota and McCulloch (1998) and Bidarkota (in press), when the disturbances ε_t and η_t are both non-Gaussian symmetric stable, the filter density typically spreads out in response to big jumps in the observed data y_t , at times even becoming multi-modal, reflecting an increased uncertainty regarding the fundamental value of the firm x_t . Gradually the filter density reverts back to a bell-shaped curve. Such behavior of the filter density would lead, after an initial jump in the stock price, to large absolute future returns as well.

One implication is that volatility clusters are originated by big shocks in the accounting data. This is a testable auxiliary restriction implied by the notion that non-Gaussian filtering leads to volatility clustering.

6. Conclusions

We set up a framework in which investors observe data that contains information about the fundamental value of a firm contaminated with noise. Investors then solve a filtering problem to probabilistically extract information about the fundamental value of the firm. They then use this information to price stocks of the firm. If the innovations driving the firm fundamentals and/or the noise in the observed data come from thick-tailed non-Gaussian probability distributions, the implied stock returns on firms can exhibit significant volatility clustering. We illustrate with a simulation study.

Our results indicate that the implied returns from non-Gaussian filtering display statistically significant volatility clustering. The evidence is overwhelming even after

accounting for thick tails in the returns data with symmetric stable innovations in an otherwise standard GARCH model. However, the volatility persistence parameter is somewhat low compared to the well-documented estimates for returns data from financial markets.

We conclude by making the observation that our results on volatility clustering are equally applicable to returns on foreign exchange. In this instance, the observed data could include, for example, macroeconomic news such as balance-of-payments data, political factors, and perhaps news reports on speculative attacks by foreign currency traders.

Table 1: Estimates from Filtering Simulated Accounting Data

$$y_t = x_t + \varepsilon_t, \quad \varepsilon_t \sim S_\alpha(0, c) \quad (3)$$

$$x_t = x_{t-1} + \eta_t, \quad \eta_t \sim S_\alpha(0, \rho c). \quad (4)$$

Estimation is done by maximum likelihood. Appendix B provides details on the likelihood function and some estimation details.

Hessian-based standard errors are in parentheses.

| | α | c | ρ | $\log L$ |
|---------------|----------------|-------------|-------------|-----------|
| Stable Data | 1.93 (0.01) | 6.39 (0.08) | 1.50 (0.03) | -21410.77 |
| Gaussian Data | 2 (restricted) | 1.70 (0.41) | 5.74 (1.48) | -20371.10 |

Table 2: GARCH-Normal Model Estimates for Simulated Returns

Conditionally Heteroskedastic Model

$$r_t = \mu + \zeta_t, \quad \zeta_t \sim c_t z_t, \quad z_t \sim \text{iid } N(0, \sqrt{2}) \quad (8)$$

$$c_t^2 = \omega + \beta c_{t-1}^2 + \delta |r_{t-1} - \mu|^2 \quad (9)$$

Homoskedastic Model

$$r_t = \mu + \zeta_t, \quad \zeta_t \sim c z_t, \quad z_t \sim \text{iid } N(0, \sqrt{2})$$

Two sets of parameter estimates are reported for each of the two models above. One set of estimates is for stock returns implied by filtering of simulated stable data and another set is for stock returns implied by filtering of simulated Gaussian data. Estimation is done by maximum likelihood. Hessian-based standard errors are in parentheses. In the last column, $2\Delta \log L$ is the likelihood ratio test statistic. The null model is the homoskedastic model and the alternative model is the conditionally heteroskedastic GARCH(1,1) model. Two restrictions, namely $\beta = \delta = 0$, on the GARCH model yield the null model. Critical values based on the χ_2^2 distribution are reported in parentheses.

| | μ | ω | β | δ | c | logL | $2\Delta\log L$ |
|-------------------------------|---------|-------------|---------|----------|---------|-----------|-----------------|
| Stable Data | | | | | | | |
| GARCH(1,1) Model Estimates | -53.43 | 925544.08 | 0.49 | 0.02 | | -17996.62 | 217.68 |
| | (38.14) | (134285.99) | (0.07) | (0.01) | | | (5.99) |
| Homoskedastic Model Estimates | -90.99 | | | | 1468.13 | -18105.46 | |
| | (44.22) | | | | (23.22) | | |
| Gaussian Data | | | | | | | |
| GARCH(1,1) Model Estimates | 19.87 | 2141.98 | 1.00 | 0.00 | | -17319.00 | 1.84 |
| | (14.85) | (1891.94) | (0.00) | (0.00) | | | (5.99) |
| Homoskedastic Model Estimates | 19.87 | | | | 991.06 | -17319.92 | |
| | (88.44) | | | | (15.67) | | |

Table 3: GARCH-Stable Model Estimates for Simulated Returns

Conditionally Heteroskedastic Model

$$r_t = \mu + \zeta_t, \quad \zeta_t \sim c_t z_t, \quad z_t \sim \text{iid} S_\alpha(0,1) \quad (10)$$

$$c_t^\alpha = \omega + \beta c_{t-1}^\alpha + \delta |r_{t-1} - \mu|^\alpha. \quad (11)$$

Homoskedastic Model

$$r_t = \mu + \zeta_t, \quad \zeta_t \sim c z_t, \quad z_t \sim \text{iid} S_\alpha(0,1)$$

Two sets of parameter estimates are reported for each of the two models above. One set of estimates is for stock returns implied by filtering of simulated stable data and another set is for stock returns implied by filtering of simulated Gaussian data. Estimation is done by maximum likelihood. Hessian-based standard errors are in parentheses. In the last column, $2\Delta \log L$ is the likelihood ratio test statistic. The null model is the homoskedastic model and the alternative model is the conditionally heteroskedastic GARCH(1,1) model. Two restrictions, namely $\beta = \delta = 0$, on the GARCH model yield the null model. Critical values based on the χ_2^2 distribution are reported in parentheses.

| | α | μ | ω | β | δ | c | logL | $2\Delta \log L$ |
|-------------------------------|----------|---------|------------|---------|----------|---------|-----------|------------------|
| Stable Data | | | | | | | | |
| GARCH(1,1) Model Estimates | 1.55 | -40.42 | 16810.10 | 0.25 | 0.07 | | -17164.56 | 57.68 |
| | (0.04) | (25.66) | (6089.51) | (0.10) | (0.01) | | | (5.99) |
| Homoskedastic Model Estimates | 1.49 | -33.02 | | | | 705.96 | -17193.40 | |
| | (0.11) | (25.62) | | | | (43.77) | | |
| Gaussian Data | | | | | | | | |
| GARCH(1,1) Model Estimates | 2.00 | -19.89 | 936889.03 | 0.00 | 0.01 | | -17319.35 | 1.14 |
| | (0.00) | (22.83) | (46627.13) | (0.01) | (0.01) | | | (5.99) |
| Homoskedastic Model Estimates | 2.00 | 21.53 | | | | 991.04 | -17319.92 | |
| | (0.00) | (29.45) | | | | (15.68) | | |

Fig.1a Simulated Noise in Observed Data

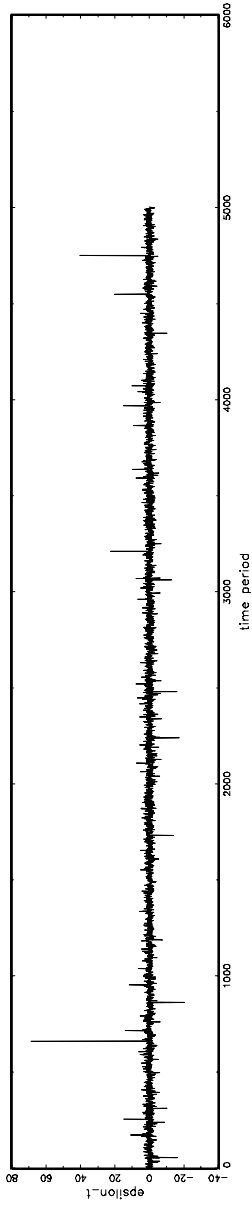


Fig.1b Simulated Signal Shocks Driving the Fundamental Value

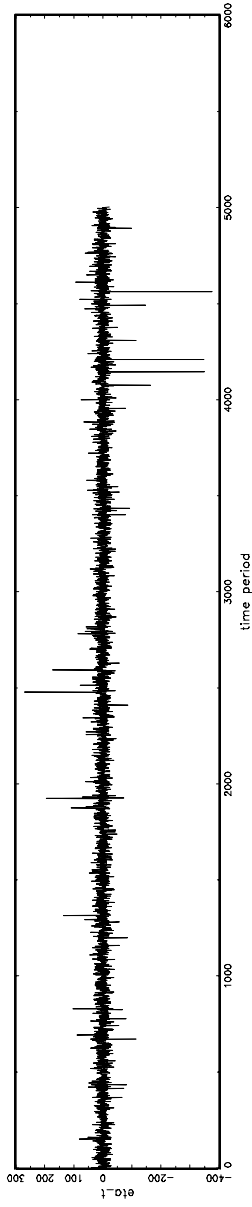


Fig.1c Simulated Observable Data

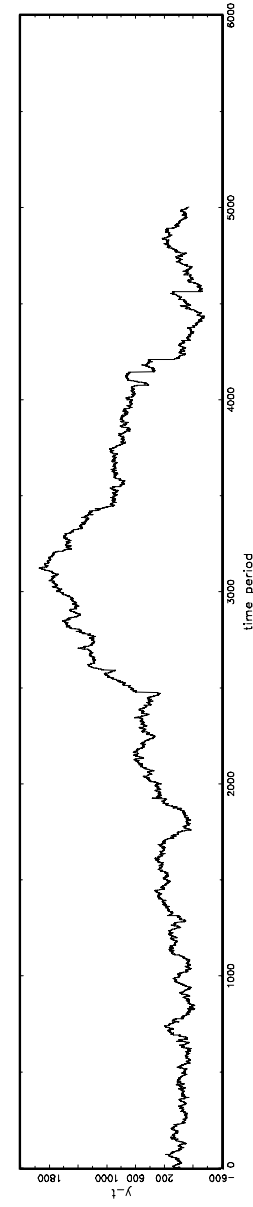


Fig.2.a Stable Filter Mean

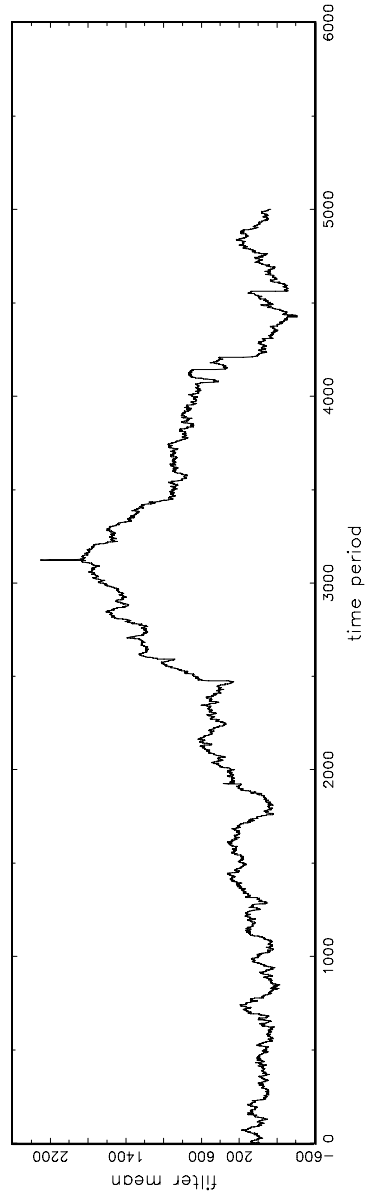


Fig.2.b Stable Filter Standard Deviation

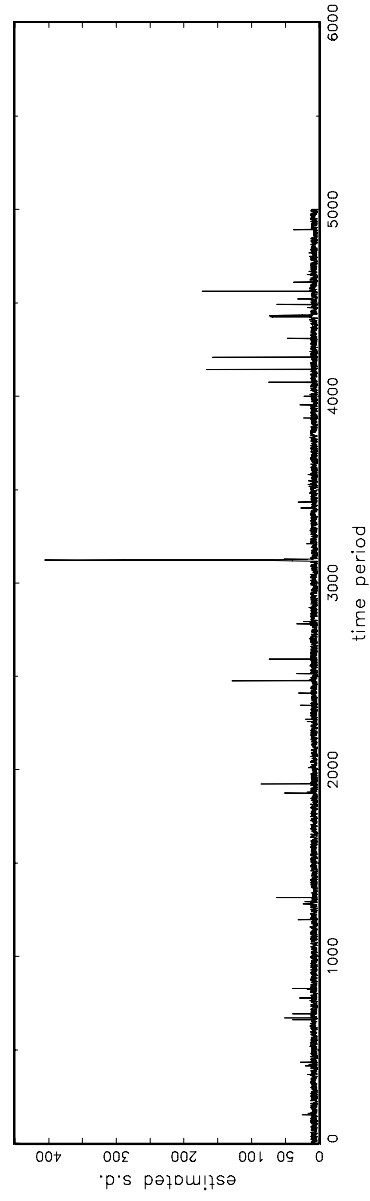


Fig.3a Autocorrelations of Squared Filter Errors

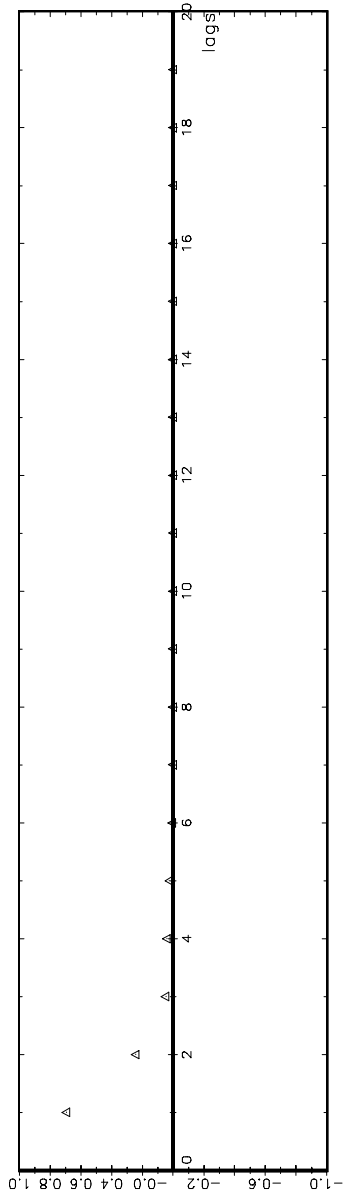


Fig.3b Partial Autocorrelations of Squared Filter Errors

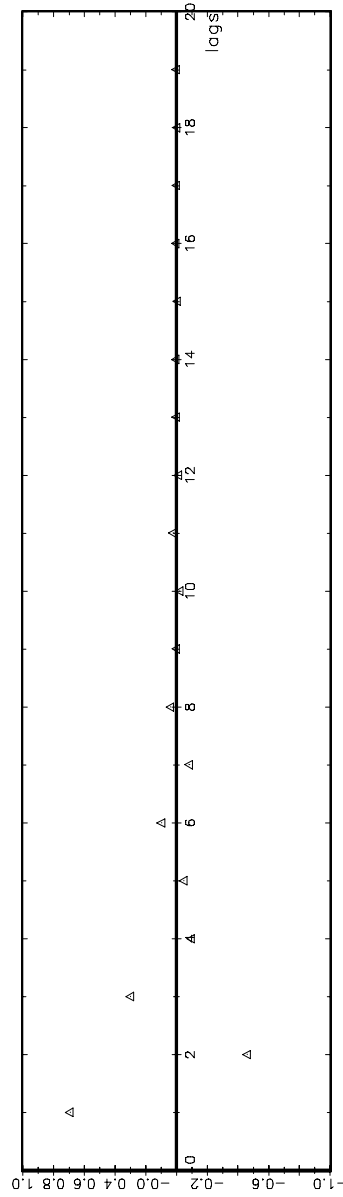


Fig.4 Stable Filter—Implied Stock Returns

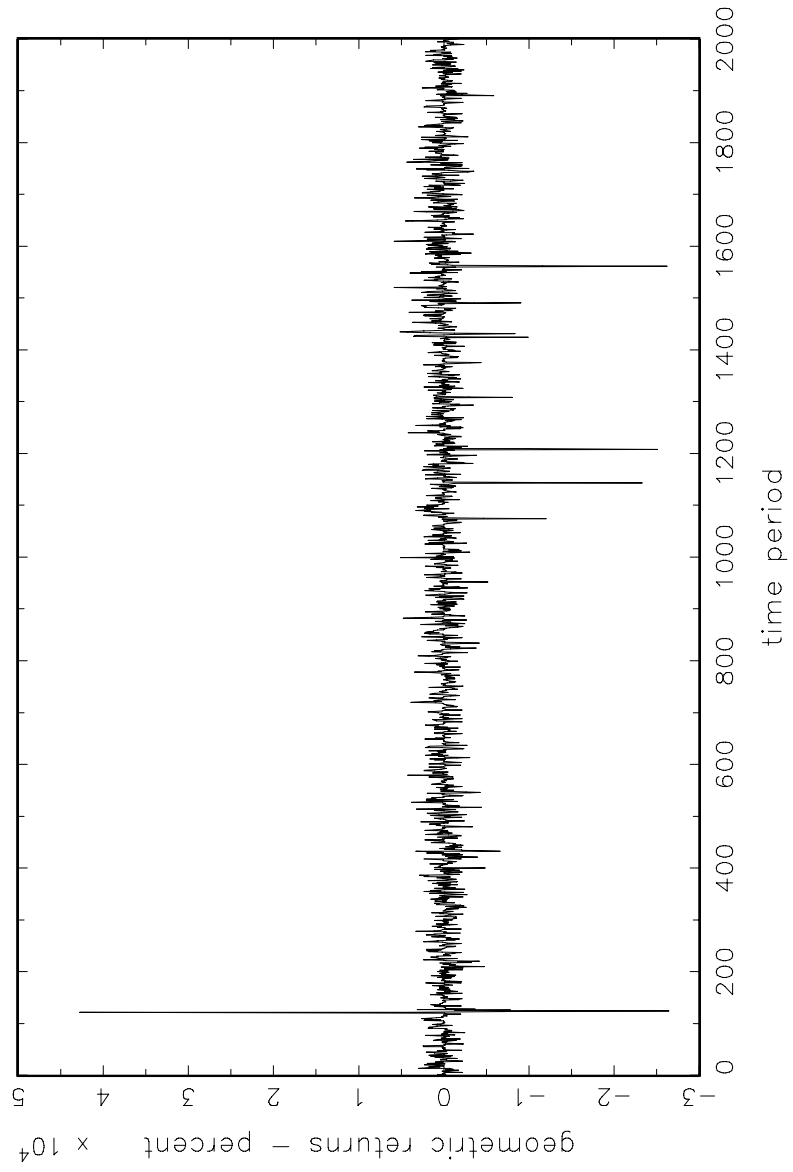


Fig.5 Volatility of Implied Returns from Stable Filtering

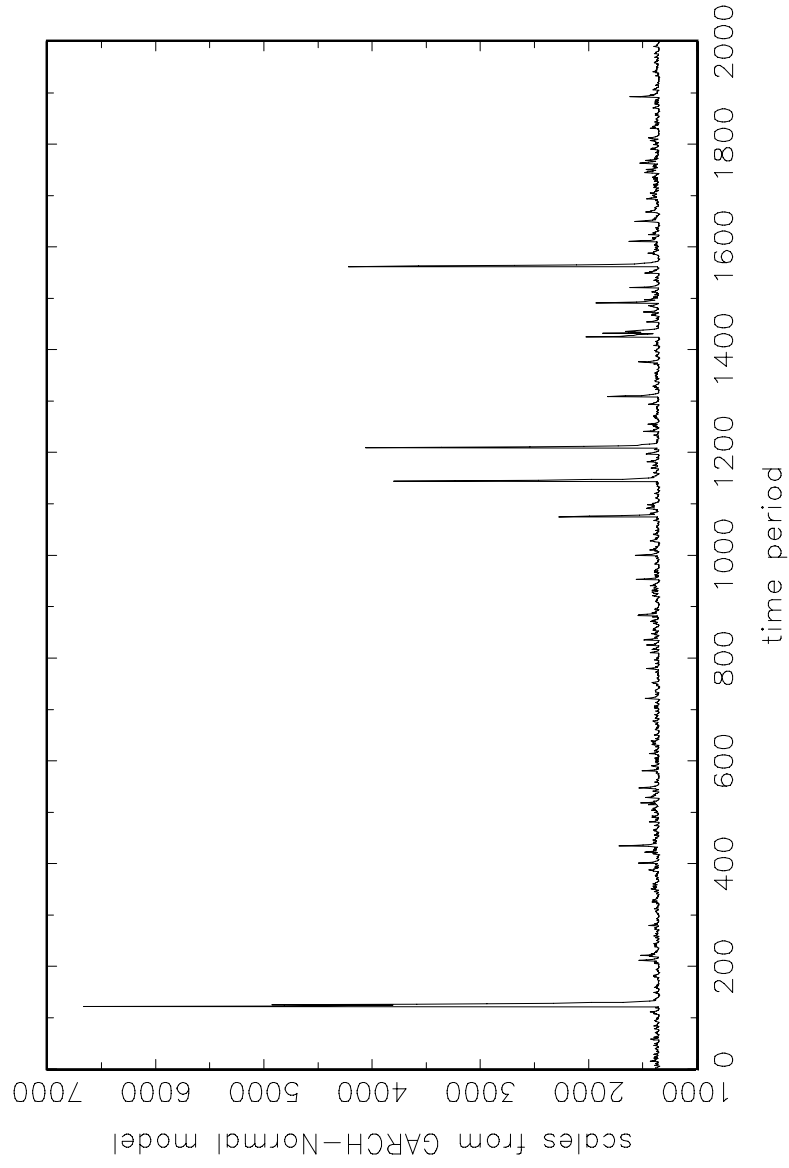
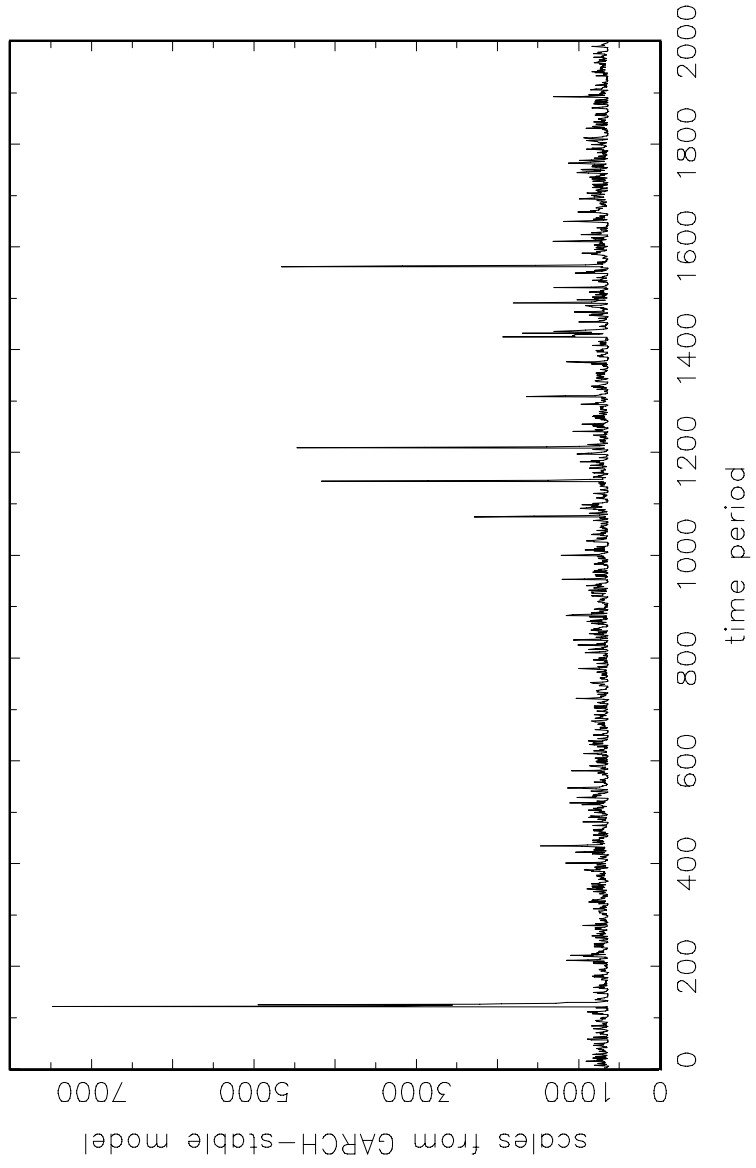


Fig.6 Volatility of Implied Returns from Stable Filtering



Appendix A. Symmetric Stable Distributions

When investors perform signal extraction on the noisy data within a non-Gaussian setting, they assume that the disturbances ε_t and η_t appearing in Equations (1) and (2) are drawn from the symmetric stable family. In this appendix, we briefly describe symmetric stable distributions.

A random variable X is said to have a symmetric stable distribution $S_\alpha(\delta, c)$ if its log-characteristic function can be expressed as:

$$\ln E \exp(iXt) = i\delta t - |ct|^\alpha. \quad (\text{A1})$$

The location parameter $\delta \in (-\infty, \infty)$ shifts the distribution to the left or right, while the scale parameter $c \in (0, \infty)$ expands or contracts it about δ . The parameter $\alpha \in (0, 2]$ is the characteristic exponent governing tail behavior, with a smaller value of α indicating thicker tails. The standard stable distribution function has $c = 1$ and $\delta = 0$.

The normal distribution belongs to the symmetric stable family with $\alpha = 2$, and is the only member with finite variance, equal to $2c^2$. Zolotarev (1986) provides a detailed description of these distributions and McCulloch (1996a) a comprehensive survey on financial applications of these distributions.

Appendix B. Gaussian and Non-Gaussian Filtering

In this appendix, we provide details on how investors make use of the noisily observed data and Equations (1) and (2) to perform filtering or signal extraction and infer $p\{x_t | Y_t\}$, where $Y_t \equiv \{y_t, y_{t-1}, \dots, y_1\}$ is the entire history of noisy data observed to date.

Equations (1) and (2) constitute a linear state space model where Equation (1) is the observation equation and Equation (2) is the state or transition equation. Accordingly, the noise ε_t in Equation (1) is the observation or measurement error and the disturbance η_t appearing in Equation (2) is the signal shock driving the state variable (firm fundamentals) x_t .

We consider two alternative filtering scenarios below. One arises when both the disturbances ε_t and η_t are assumed to be Gaussian. The other arises when both ε_t and η_t are assumed non-Gaussian.

B1. Gaussian Filtering

When both disturbances ε_t and η_t are Gaussian and both the observation and state equations are linear as we have in Equations (1) and (2), we obtain the standard linear Gaussian state space framework (the local level model). Here, the filter density $p\{x_t | Y_t\}$ turns out to be Gaussian as well, and hence is completely specified by its mean and variance. In this case, the celebrated Kalman filter provides recursive formulae for calculating the mean and variance of the filter density. These recursions can be found in any standard textbook, such as Harvey (1992, chapter 3).

B2. Non-Gaussian Filtering

When both disturbances ε_t and η_t are non-Gaussian, we obtain the non-Gaussian state space model. In this case, the filter density $p\{x_t | Y_t\}$ too will turn out to be non-Gaussian as well. Hence, it will not be completely specified by just its mean and variance

alone. In this situation, the linear recursive formulae for updating the mean and variance of the filter density given by the Kalman filter are no longer optimal. The globally optimal filter turns out to be non-linear and is given by the Sorenson-Alspach (1971) filtering algorithm (see also Harvey (1992), p.162-165).

This algorithm provides the following recursive formulae for obtaining one step-ahead prediction $p\{x_t | Y_{t-1}\}$ and filtering $p\{x_t | Y_t\}$ densities for the unobserved state x_t :

$$p(x_t | Y_{t-1}) = \int_{-\infty}^{\infty} p(x_t | x_{t-1}) p(x_{t-1} | Y_{t-1}) dx_{t-1}, \quad (\text{B1})$$

$$p(x_t | Y_t) = p(y_t | x_t) p(x_t | Y_{t-1}) / p(y_t | Y_{t-1}), \quad (\text{B2})$$

$$p(y_t | Y_{t-1}) = \int_{-\infty}^{\infty} p(y_t | x_t) p(x_t | Y_{t-1}) dx_t. \quad (\text{B3})$$

When both disturbance terms ε_t and η_t in Equations (1) and (2) are normally distributed, the Sorenson-Alspach filter collapses to the Kalman filter. In this case, one can evaluate the above integrals analytically. However, in general, these integrals cannot usually be solved in closed form under non-Gaussian distributional assumptions on the error terms.

One approach is to evaluate these integrals numerically, as in Kitagawa (1987), or Hodges and Hale (1993). An alternative that works well with high-dimensional integration is the Monte Carlo integration technique, as in Tanizaki and Mariano (1998) or Durbin and Koopman (2000).

If it is required to estimate the unknown parameters of the model (the hyperparameters), namely the parameters of the distributions for ε_t and η_t , one can

make use of the maximum likelihood estimator. The log-likelihood function, conditional on the hyperparameters of the model, is given by:

$$\log p(y_1, \dots, y_T) = \sum_{t=1}^T \log p(y_t | Y_{t-1}). \quad (\text{B4})$$

B3. Numerical Implementation of Non-Gaussian Filtering

In this paper, filtering in the case when the disturbances ε_t and η_t in the state space model given in Equations (1) and (2) are non-Gaussian is done by evaluating the integrals given in Equations (B1)-(B3) with the numerical integration techniques in Bidarkota and McCulloch (1998). They provide details on the accuracy of their approximation procedure.

The probability density for the symmetric stable distributions required for filtering and maximum likelihood estimation of all the non-Gaussian stable models is computed using the numerical algorithm in McCulloch (1996b).

Appendix C. A Model of Stock Pricing

In this appendix, we develop a simple model of stock pricing that can be used to understand how stock returns are related to the filter distributions. In section C1, we derive within an expected utility framework the certainty-equivalent stock price for taking on the gamble of investment in stocks. Section C2 derives stock returns analytically in a special case where information filtering is perfect, in a sense that is defined there. Section C3 derives stock returns analytically in a special case where information filtering is done in a Gaussian framework.

C1. Certainty-Equivalent Pricing of Stocks

Assume that investors view investment in stocks as a one-shot gamble, repeated every period, with uncertain payoffs with perceived probabilities $p\{x_t | Y_t\}$. For risk-averse investors whose utility $U(\cdot)$ is defined over money holdings, one can calculate the certainty-equivalent stock price Q_t for taking on the gamble of investment in stocks by solving:

$$U(Q_t) = E[U(X_t) | Y_t] \quad (C1)$$

where $E(\cdot)$ is the conditional expectation operator.

For instance, with a constant relative risk aversion (CRRA) utility function:

$$U(X_t) = \frac{X_t^{1-\gamma}}{1-\gamma}, \quad \gamma \geq 0 \quad (C2)$$

where γ is the coefficient of relative risk aversion, the expected utility from investing in stocks:

$$E[U(X_t) | Y_t] = \int_{-\infty}^{+\infty} U(X_t) p(x_t | Y_t) dx_t \quad (C3)$$

can be written as:

$$E[U(X_t) | Y_t] = \int_{-\infty}^{+\infty} \frac{(e^{x_t})^{1-\gamma}}{1-\gamma} p(x_t | Y_t) dx_t. \quad (C4)$$

This can be evaluated for any value of γ once the filter probabilities $p\{x_t | Y_t\}$ are known.

One can then use Equation (C1) to calculate the certainty-equivalent stock price at any time t by:

$$Q_t = \left[(1-\gamma) \int_{-\infty}^{+\infty} \frac{(e^{x_t})^{1-\gamma}}{1-\gamma} p(x_t | Y_t) dx_t \right]^{\frac{1}{1-\gamma}}. \quad (C5)$$

Given the certainty-equivalent stock prices, one can readily evaluate the returns to holding stocks:

$$r_t = \ln(Q_t) - \ln(Q_{t-1}). \quad (C6)$$

C2. Stock Returns Under Perfect Filtering

When investors observe the fundamental value of the firm per share X_t , we can think of this as the perfect filtering situation. In this case, the filter probability density becomes degenerate. That is,

$$\text{prob}\{x_t | Y_t\} = \begin{cases} 1 & \text{when } x_t = X_{t, \text{fundamental value}} \\ 0 & \text{when } x_t \neq X_{t, \text{fundamental value}} \end{cases}. \quad (C7)$$

Using Equation (C3), the expected utility from investing in stocks can then be written as:

$$E[U(X_t) | Y_t] = \sum_{\substack{X_t = X_{t, \text{fundamental value}} \\ X_t \neq X_{t, \text{fundamental value}}}} U(X_t) \text{prob}(x_t | Y_t). \quad (C8)$$

Using Equation (C7) this becomes:

$$E[U(X_t) | Y_t] = U(X_{t, \text{fundamental value}}). \quad (C9)$$

The certainty equivalent stock price Q_t obtained by solving Equation (C1) becomes:

$$Q_t = X_{t, \text{fundamental value}}. \quad (C10)$$

Thus, when investors are able to extract information about the fundamental value of the firm per share perfectly, they price stocks at the fundamental value of the firm per share

$$X_{t, \text{fundamental value}}$$

If the logarithm of the fundamental value of the firm per share x_t behaves as a driftless random walk, that is, if Equation (2) is in fact the true generating process for x_t , then the geometric returns to holding stocks can be evaluated using Equations (2), (C6), and (C10) as:

$$\begin{aligned} r_t &= \ln(Q_t) - \ln(Q_{t-1}) \\ &= \ln(X_{t, \text{fundamental value}}) - \ln(X_{t-1, \text{fundamental value}}) \end{aligned}$$

or
$$r_t = \eta_t. \tag{C11}$$

Therefore, if η_t in Equation (2) is drawn from an iid distribution, then returns will also be iid under perfect filtering. Specifically, returns will not exhibit any volatility clustering.

C3. Stock Returns Under Gaussian Filtering

When investors perform filtering or signal extraction assuming that the observational noise and the disturbance driving the fundamental value of the firm (per share), that is, assuming that ε_t and η_t in Equations (1) and (2) respectively, are both Gaussian, it can be shown that the filter density $p\{x_t | Y_t\}$ is Gaussian as well (see Appendix B1 for further details). In this case, we can solve Equation (C5) analytically

and obtain a closed form solution for the certainty-equivalent stock price Q_t , and hence for returns r_t as well. We now proceed to derive these quantities below.

Under Gaussian filtering, since $p\{x_t | Y_t\}$ is Gaussian, it is completely specified by its mean and variance. Let us denote:

$$p\{x_t | Y_t\} \sim N(\mu_{t|t}, \sigma_{t|t}^2).$$

Then, from the properties of Gaussian random variables, we have:

$$p\{(1-\gamma)x_t | Y_t\} \sim N((1-\gamma)\mu_{t|t}, (1-\gamma)^2 \sigma_{t|t}^2).$$

Using the formula for the moment generating function for Gaussian distributions, we can evaluate:

$$E[\exp\{(1-\gamma)x_t\}] = \exp\left\{(1-\gamma)\mu_{t|t} + \frac{(1-\gamma)^2 \sigma_{t|t}^2}{2}\right\}. \quad (C12)$$

We can now solve for the expected utility in this setting by evaluating the integral in Equation (C4) as:

$$E[U(X_t) | Y_t] = \left(\frac{1}{1-\gamma}\right) \exp\left\{(1-\gamma)\mu_{t|t} + \frac{(1-\gamma)^2 \sigma_{t|t}^2}{2}\right\}. \quad (C13)$$

The certainty-equivalent stock price Q_t can be obtained by using the above expression for expected utility in Equation (C5) and simplifying. We get:

$$Q_t = \exp\left\{\mu_{t|t} + \frac{(1-\gamma)\sigma_{t|t}^2}{2}\right\}. \quad (C14)$$

Stock returns can then be evaluated using Equation (C6) as:

$$r_t = (\mu_{t|t} - \mu_{t-1|t-1}) + \left(\frac{1-\gamma}{2}\right)(\sigma_{t|t}^2 - \sigma_{t-1|t-1}^2). \quad (C15)$$

Thus, under Gaussian filtering this equation provides an exact analytical formula for stock returns. It turns out to be a function of the first differences in the mean and variance of the filter density, and the investor's risk aversion coefficient.

Under non-Gaussian filtering, the exact filter density $p\{x_t | Y_t\}$ is available only in the form of integrals, given in Equations (B1)-(B3) in Appendix B. Since these integrals cannot, in general, be evaluated analytically, there is little scope for deriving stock prices and returns analytically in such a setting.

REFERENCES

Bidarkota, P.V. (in press), 'Do fluctuations in U.S. inflation rates reflect infrequent large shocks or frequent small shocks?', The Review of Economics and Statistics.

Bidarkota, P.V. and J.H. McCulloch (1998), 'Optimal univariate inflation forecasting with symmetric stable shocks,' Journal of Applied Econometrics, Vol.13, No.6, 659-670.

Bidarkota, P.V. and J.H. McCulloch (2002), 'Testing for persistence in stock returns with GARCH-stable shocks,' Working Paper, The Ohio State University.

Bookstaber, R.M. and S. Pomerantz (1989), 'An information-based model of market volatility,' Financial Analysts Journal, 37-46.

den Haan, W.J. and S.A. Spear (1998), 'Volatility clustering in real interest rates: Theory and evidence,' Journal of Monetary Economics, 41, 431-453.

Durbin, J. and S.J. Koopman (2000), 'Time series analysis of non-Gaussian observations based on state space models from both classical and Bayesian perspectives,' Journal of The Royal Statistical Society, Series B, 62, Part 1, 3-56.

Engle, R.F. (1982), 'Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation,' Econometrica, 50, 987-1008.

Fama, E. (1991), 'Efficient capital markets: II,' Journal of Finance, Vol.XLVI, No.5, 1575-1617.

French, K.R. and R. Roll (1986), 'Stock return variances: The arrival of information and the reaction of traders,' Journal of Financial Economics, 17, 5-26.

French, K.R., G.W. Schwert, and R.F. Stambaugh (1987), 'Expected stock returns and volatility,' Journal of Financial Economics, 19, 3-29.

Ghose, D. and K.F. Kroner (1995), 'The relationship between GARCH and symmetric stable processes: Finding the source of fat tails in financial data,' Journal of Empirical Finance 2, 225-251.

Groenendijk, P.A., A. Lucas, and C.G. de Vries (1995), 'A note on the relationship between GARCH and symmetric stable processes,' Journal of Empirical Finance 2, 253-264.

Harvey, A.C. (1992), *Forecasting, Structural Time Series Models and the Kalman Filter* (Cambridge University Press, Cambridge, UK).

Hodges, P.E. and D.F. Hale (1993), 'A computational method for estimating densities of non-Gaussian nonstationary univariate time series,' Journal of Time Series Analysis, Vol.14, No.2, 163-178.

Kitagawa, G. (1987), 'Non-Gaussian state space modeling of nonstationary time series,' Journal of the American Statistical Association, Vol.82, No.400, 1032-63.

Liu, S.M. and B.W. Brorsen (1995), 'Maximum likelihood estimation of a GARCH-stable model,' Journal of Applied Econometrics, Vol.10, 273-285.

Mandelbrot, B. (1963), 'The variation of certain speculative prices,' Journal of Business, 36, 394-419.

McCulloch, J.H. (1996a), Financial applications of stable distributions, in: Maddala, G.S., Rao, C.R., eds., Handbook of Statistics, Vol.14 (Elsevier, Amsterdam) 393-425.

_____ (1996b), 'Numerical approximation of the symmetric stable distribution and density,' in R. Adler, R. Feldman, and M.S. Taquq (eds.), A Practical Guide to Heavy Tails: Statistical Techniques for Analyzing Heavy Tailed Distributions, Boston: Birkhauser.

_____ (1985), 'Interest-risk sensitive deposit insurance premia: stable ACH estimates,' Journal of Banking and Finance, 9, 137-156.

Pagan, A. (1996), 'The econometrics of financial markets,' Journal of Empirical Finance, Vol.3, No.1, 15-102.

Pagan, A.R. and G.W. Schwert (1990), 'Alternative models for conditional stock volatility,' Journal of Econometrics, 45, 267-290.

Peng, L. and W. Xiong (2001), 'Time to digest and volatility dynamics,' Working Paper, Princeton University.

Sims, C.A. (2002), 'Implications of rational inattention,' Working Paper, Princeton University.

Sorenson, H.W. and D.L. Alspach (1971), 'Recursive Bayesian estimation using Gaussian sums,' Automatica, 7, 465-479.

Tanizaki, H. and R.S. Mariano (1998), 'Nonlinear and non-Gaussian state-space modeling with Monte Carlo simulations,' Journal of Econometrics, 83, 263-290.

Zolotarev, V.M., 1986, One dimensional stable laws, American Mathematical Society.
(Translation of *Odnomernye Ustoichivye Raspredeleniia* (NAUKA, Moscow, 1983).)