

Estimating the Structural Credit Risk Model When Equity Prices Are Contaminated by Trading Noises

Jin-Chuan Duan and Andras Fulop*

(August 8, 2006)

Abstract

The transformed-data maximum likelihood estimation (MLE) method for structural credit risk models developed by Duan (1994) is extended to account for the fact that observed equity prices may have been contaminated by trading noises. With the presence of trading noises, the likelihood function based on the observed equity prices can only be evaluated via some nonlinear filtering scheme. We devise a particle filtering algorithm that is practical for conducting the MLE estimation of the structural credit risk model of Merton (1974). We implement the method on the Dow Jones 30 firms and on 100 randomly selected firms, and find that ignoring trading noises can lead to significantly over-estimating the firm's asset volatility. The estimated magnitude of trading noise is in line with the direction that a firm's liquidity will predict based on three common liquidity proxies. A simulation study is then conducted to ascertain the performance of the estimation method.

Keywords: Particle filtering, maximum likelihood, option pricing, credit risk, microstructure.

JEL classification code: C22

*Duan and Fulop are with Joseph L. Rotman School of Management, University of Toronto. Duan acknowledges support received as the Manulife Chair in Financial Services and research funding from both the Social Sciences and Humanities Research Council of Canada and the Natural Sciences and Engineering Research Council of Canada. E-mail: jcduan@rotman.utoronto.ca. Fulop is a PhD student. E-mail: Andras.Fulop@rotman.utoronto.ca.

1 Introduction

Structural credit risk models rely on the notion of claim priority and limited liability, which allows a firm's equity and debt be viewed as contingent claims that partition the asset value of the firm. Black and Scholes (1973) were the first to formally consider equity as a call option on the firm's asset value. However, it was the corporate bond pricing model of Merton (1974) that popularized the structural approach to modeling risky corporate debts. Although the structural approach is based on a powerful and compelling interpretation of a firm's credit risk, implementation is complicated by the fact that the firm's asset values cannot be directly observed as argued in, for example, Jarrow and Turnbull (2000). It seems then that the pertinent parameters of a structural credit risk model cannot be estimated. In fact, the implementation difficulty motivates an alternative approach known as reduced-form, which treats corporate default as an event governed by an exogenous shock that is not based on the firm's asset value failing to cover its debt obligation.

Contrary to a common belief, non-observability of a firm's asset values does not actually impede the implementation of structural credit risk models. Duan (1994) has devised a transformed-data maximum likelihood estimation (MLE) method for structural credit risk models that is solely based on a time series of observed equity values. That MLE method hinges on a recognition that the equity value should result from a one-to-one differentiable transformation of the firm's unobserved asset value under the given structural credit risk model even though the transformation depends on some unknown model parameter(s). Therefore, one can explicitly write out the likelihood function only based on the observed equity time series.

The transformed-data MLE method of Duan (1994, 2000) has been applied in credit risk analysis by Ericsson and Reneby (2004*a,b*), Wong and Choi (2004) and Duan et al. (2003). It has also been adopted for banking research by Duan and Yu (1994), Laeven (2002), Duan and Simonato (2002) and Lehar (2004). Interestingly, the KMV method, which is a popular commercial implementation of the structural credit risk model, is equivalent to the transformed-data MLE method in a restrictive sense (see Duan et al. (2004)).

The contribution of this paper is to generalize the transformed-data MLE method of Duan (1994) to allow for trading noises in the observed equity prices. It has been well documented in the market microstructure literature that observed equity prices can diverge from their equilibrium values due to microstructure noises (e.g. illiquidity, asymmetric information, price discreteness). Some of the examples are Hasbrouck (1993), Harris (1990) and Madhavan et al. (1997). The presence of noises can also induce equity returns to exhibit the moving average feature, a phenomenon that has been long recognized in the literature. Recently, Ait-Sahalia et al. (2005a) have, for example, analyzed the effect of trading noise on how frequently one should sample the equity price. In the realized volatility literature, microstructure noises have also been shown to have material effect on volatility estimation; for example, Ait-Sahalia et al. (2005b) and Bandi and Russell (2006). In markets where the trading noise effect is material, it will be ill-advised to ignore its presence. In the specific context of structural credit risk models, ignoring trading noise could non-trivially inflate one’s estimate for the “true” asset volatility. Since the asset volatility plays a key role in the structural credit risk models, one is then likely to produce misleading estimates for credit spreads, default probabilities and other corporate contingent claims.

When trading noises are present in equity prices, the transformed-data MLE method of Duan (1994, 2000) can no longer be applied. The reason is that the previous one-to-one relationship between the unobserved asset value and the observed equity price when there was no trading noise is broken. The equity value is influenced by two sources of uncertainty – the underlying asset value and the trading noise. In short, estimation becomes a nonlinear filtering problem with the unobserved asset value being the “signal”. The “true” equity value, as a nonlinear function of the unobserved asset value, is thus contaminated by a trading noise.

We devise a nonlinear filtering scheme using the auxiliary particle filtering idea of Pitt and Shephard (1999). The nature of our estimation problem allows us to devise a specific sampler that is localized relative to the “true” asset value. Since the likelihood function based on the typical particle filtering algorithm is not continuous due to a required resampling step, smoothness must be built into the algorithm to make it suitable for parameter estimation.

For this purpose, we adopt the smoothed version of the auxiliary particle filter put forward by Pitt (2002). In addition, we must address the potential cases of small trading noise. Small trading noises can cause the likelihood function to spike in the context of particle filtering. We thus devise a new way of computing the likelihood function to circumvent the spiking problem. It turns out that the likelihood function can be easily evaluated with the use of our localized sampler.

We implement the particle filter-based MLE method on two samples. The first sample consists of the Dow Jones 30 companies on the belief that they are less susceptible to trading noises. We also conduct analysis on a sample of 100 randomly selected U.S. listed firms to represent the general population of the U.S. corporate sector. For either sample, we find trading noise to be significant for a percentage of firms that cannot be attributed to chance. The impact of ignoring trading noise is also studied. We find that omission can cause the asset volatility to be greatly overestimated in many cases. In addition, the estimated magnitude of trading noise is found to be in line with the direction that a firm's liquidity will predict based on three liquidity proxy variables - bid-ask spread, firm size and trading volume.

A Monte Carlo study is conducted to determine whether the asymptotic inference can be reasonably applied when one uses a time series sample of daily equity values over one-year time span. Our results indicate that asymptotic inference works well except for the trading noise parameter. When the magnitude of trading noise is small, the sampling distribution of the trading noise parameter appears to deviate somewhat from normality implied by the asymptotic normality, a result due to being close to this parameter's lower bound. When its magnitude is large, the asymptotic normality turns out to be a reasonable description of its sampling distribution. We conduct a likelihood ratio test for the presence of trading noise. Our test explicitly takes into account that the null hypothesis is a boundary value of the parameter set. Our simulation study indicates that this likelihood ratio test is unbiased and has reasonable power.

2 Merton's model and maximum likelihood estimation via particle filtering

2.1 Equity value in Merton's model

Merton (1974) laid the foundation to the literature on the structural approach to credit risk modeling. The value of the firm at time t , V_t , is assumed to follow a geometric Brownian motion with respect to the physical probability law that generates the asset values. The stochastic process is governed by the drift and volatility rate parameters – μ and σ .

$$\frac{dV_t}{V_t} = \mu dt + \sigma dW_t \quad (1)$$

The risk-free rate of interest is assumed to be constant r . Furthermore, the firm has two classes of claims outstanding – an equity and a zero-coupon debt maturing at time T with face value F .

Debt has a claim priority over equity and equity holders are protected by limited liability. When debt matures, equity holders repay the debt and keep the balance, if the value of the firm is sufficient to cover the debt obligation. Otherwise, the firm defaults and the debt holders get to keep the remaining firm value. The payout to the debt holders at time T naturally becomes

$$D_T = \min(V_T, F) \quad (2)$$

The equity holders, on the other hand, receive at time T

$$S_T = \max(V_T - F, 0) \quad (3)$$

Merton (1974) derived a pricing formula for the risky debt defined in this framework and focused on the credit risk issue. Here, we address the flip side, i.e., the equity claim, because we need to establish the link to the observed equity price in order to develop the estimation method. In other words, Merton (1974) focused on the theoretical aspect of credit risk whereas we zero in on the econometric issue so as to make the implementation of Merton's model feasible.

The equity claim in equation (3) can be priced at time $t < T$ by the standard Black-Scholes option pricing model to yield the following solution:

$$S_t \equiv S(V_t; \sigma, F, r, T - t) = V_t \Phi(d_t) - F e^{-r(T-t)} \Phi(d_t - \sigma \sqrt{T-t}) \quad (4)$$

where

$$d_t = \frac{\ln(\frac{V_t}{F}) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma \sqrt{T-t}}$$

and $\Phi(\cdot)$ is the standard normal distribution function. Observe that the equity pricing formula is not a function of the drift parameter μ . Also note that the equity pricing function is invertible with respect to the asset value.

2.2 Estimation is a nonlinear filtering problem

For an exchange listed firm, one can obtain a time series of equity prices denoted by $\mathcal{D}_N = \{S_{\tau_i}, i = 0, \dots, N\}$. For simplicity, we assume that equity prices are sampled at a fixed frequency; that is, we let $\tau_i - \tau_{i-1} = h$.

If the equity prices are not contaminated by trading noises, the likelihood function based on \mathcal{D}_N can be written out and estimated using the transformed-data MLE method developed in Duan (1994, 2000). If one has a reason to believe that the observed equity prices contain trading noises, then estimation becomes much more complex. The market microstructure literature indeed strongly suggests that noises should be expected. Therefore, the relationship between the unobserved asset and the observed equity value predicted by the equity pricing formula is masked by trading noise. We assume a multiplicative error structure for the trading noise to express the logarithmic equity value as follows:

$$\ln S_{\tau_i} = \ln S(V_{\tau_i}; \sigma, F, r, T - \tau_i) + \delta \nu_i \quad (5)$$

where $\{\nu_i, i = 0, N\}$ are i.i.d. standard normal random variables and the nonlinear pricing function $S(V_i; \sigma, F, r, T - t)$ has been given earlier.

Since the unobserved asset value process follows the geometric Brownian motion, we can derive its discrete-time form as

$$\ln V_{\tau_{i+1}} = \ln V_{\tau_i} + (\mu - \frac{\sigma^2}{2})h + \sigma \sqrt{h} \varepsilon_{i+1} \quad (6)$$

where $\{\varepsilon_i, i = 1, N\}$ are i.i.d. standard normal random variables. Equations (5) and (6) constitute a state-space model with the first being the measurement equation and the second the transition equation. To estimate this model, we need to develop a practical scheme to deal with this non-linear filtering problem. If the equity pricing function were linear, one would of course be able to estimate the model using the standard Kalman filter.

The state-space model contains three parameters, denoted by $\Theta = \{\sigma, \delta, \mu\}$. The likelihood function of the observed sample of equity prices can in principle be expressed as

$$f(\mathcal{D}_N | \Theta) = f(S_{\tau_N} | \mathcal{D}_{N-1}, \Theta) \cdots f(S_{\tau_1} | \mathcal{D}_0, \Theta). \quad (7)$$

Note that we can view the conditional density in the above expression in two different ways. First,

$$f(S_{\tau_{i+1}} | \mathcal{D}_i, \Theta) = \int_0^\infty f(S_{\tau_{i+1}} | V_{\tau_{i+1}}, \Theta) f(V_{\tau_{i+1}} | \mathcal{D}_i, \Theta) dV_{\tau_{i+1}} \quad (8)$$

where $f(V_{\tau_{i+1}} | \mathcal{D}_i, \Theta)$ is the density of $V_{\tau_{i+1}}$ conditional on all equity values up to τ_i , which is known as the prediction density. Alternatively,

$$f(S_{\tau_{i+1}} | \mathcal{D}_i, \Theta) = \int_0^\infty f(S_{\tau_{i+1}} | V_{\tau_i}, \Theta) f(V_{\tau_i} | \mathcal{D}_i, \Theta) dV_{\tau_i} \quad (9)$$

where $f(V_{\tau_i} | \mathcal{D}_i, \Theta)$ is the density of V_{τ_i} conditional on all equity values up to τ_i , which is known as the filtering density.

Even though $f(S_{\tau_{i+1}} | V_{\tau_{i+1}}, \Theta)$ in (8) has a simple analytical form, it turns out that for a numerical reason, we need to use the expression in (9), which involves $f(S_{\tau_{i+1}} | V_{\tau_i}, \Theta)$. It should be clear that $f(S_{\tau_{i+1}} | V_{\tau_{i+1}}, \Theta)$ spikes when the trading noise is very small, and as a result it is almost impossible to accurately evaluate the conditional likelihood as stated in equation (8). In contrast, the following derivation suggests a way to circumvent the spiking problem.

$$\begin{aligned} f(S_{\tau_{i+1}} | V_{\tau_i}, \Theta) &= \int_{-\infty}^\infty f(S_{\tau_{i+1}} | \nu_{i+1}, V_{\tau_i}, \Theta) f(\nu_{i+1} | V_{\tau_i}, \Theta) d\nu_{i+1} \\ &= \int_{-\infty}^\infty f(S_{\tau_{i+1}} | \nu_{i+1}, V_{\tau_i}, \Theta) f(\nu_{i+1} | \Theta) d\nu_{i+1} \\ &= \int_{-\infty}^\infty \frac{f(V_{\tau_{i+1}} = V_{\tau_{i+1}}^*(S_{\tau_{i+1}}, \nu_{i+1}) | V_{\tau_i}, \Theta)}{\Phi(d_{\tau_{i+1}}^*) e^{\delta \nu_{i+1}}} f(\nu_{i+1} | \Theta) d\nu_{i+1} \end{aligned} \quad (10)$$

where $V_{\tau_{i+1}}^*(S_{\tau_{i+1}}, \nu_{i+1}) = S^{-1}(S_{\tau_{i+1}} e^{-\delta\nu_{i+1}}; \sigma, F, r, T - \tau_{i+1})$, an inversion of the equity pricing equation in (4), and $e^{-\delta\nu_{i+1}}/\Phi(d_{\tau_{i+1}}^*)$ is the Jacobian of the inverse transformation. The second equality in the above is due to the fact that a future trading noise is independent of the current asset value. The third equality follows because once the trading noise is known, the equity value implies a specific asset value.

Putting (9) and (10) together, we have

$$\begin{aligned}
& f(S_{\tau_{i+1}} | \mathcal{D}_i, \Theta) \\
&= \int_0^\infty \int_{-\infty}^\infty \frac{f(V_{\tau_{i+1}} = V_{\tau_{i+1}}^*(S_{\tau_{i+1}}, \nu_{i+1}) | V_{\tau_i}, \Theta)}{\Phi(d_{\tau_{i+1}}^*) e^{\delta\nu_{i+1}}} f(\nu_{i+1} | \Theta) f(V_{\tau_i} | \mathcal{D}_i, \Theta) d\nu_{i+1} dV_{\tau_i} \\
&= E \left[\frac{f(V_{\tau_{i+1}} = V_{\tau_{i+1}}^*(S_{\tau_{i+1}}, \nu_{i+1}) | V_{\tau_i}, \Theta)}{\Phi(d_{\tau_{i+1}}^*) e^{\delta\nu_{i+1}}} \middle| \mathcal{D}_i \right] \tag{11}
\end{aligned}$$

The above expectation can be evaluated easily within the particle filtering system to be discussed in the next subsection. Our approach shares the spirit of the stepping-back idea of Ionides (2004)) in smoothing the likelihood function.

The result in (11) immediately implies that our nonlinear filtering design will simplify to the transformed-data approach of Duan (1994, 2000) when there are no trading noises, i.e., $\delta = 0$. This is true because without trading noise the item inside the expectation operator has the unit mass on the single point $V_{\tau_{i+1}}^*(S_{\tau_{i+1}}, 0)$.

2.3 Estimation using a smoothed particle filter

The particle filter is a simulation-based technique to generate consecutive prediction and filtering distributions for general nonlinear or non-Gaussian state-space models. The technique relies on different sets of points to represent the distributions of the unobserved state variable(s) at different stages. Bayes' rule is repeatedly used to re-weight the particles in advancing the system. Particle filtering is generally attributed to Gordon et al. (1993). Doucet et al. (2001) offer a wealth of information on the theory and applications of particle filtering.

For our particular filtering problem, $f(V_{\tau_i} | \mathcal{D}_i, \Theta)$ is represented by a set of M particles $\{V_{\tau_i}^{(m)}(\Theta); m = 1, \dots, M\}$ with equal weights. The empirical prediction density can be

expressed as

$$\hat{f}(V_{\tau_{i+1}} | \mathcal{D}_i, \Theta) \propto \frac{1}{M} \sum_{m=1}^M f(V_{\tau_{i+1}} | V_{\tau_i}^{(m)}, \Theta) \quad (12)$$

By Bayes' rule, $f(V_{\tau_{i+1}} | \mathcal{D}_{i+1}, \Theta)$ can be approximated by the following empirical filtering density:

$$\hat{f}(V_{\tau_{i+1}} | \mathcal{D}_{i+1}, \Theta) \propto f(S_{\tau_{i+1}} | V_{\tau_{i+1}}, \Theta) \hat{f}(V_{\tau_{i+1}} | \mathcal{D}_i, \Theta) \quad (13)$$

Equations (12) and (13) provide the basis for advancing the system to the next set of particles of equal weights corresponding to the next time point. The simplest algorithm for advancing the system is the sampling/importance resampling (SIR) method as follows.

- **Step 1:** Draw M points from $\hat{f}(V_{\tau_{i+1}} | \mathcal{D}_i, \Theta)$. Since $\{V_{\tau_i}^{(m)}; m = 1, \dots, M\}$ is an equal-weight sample, one only needs to sample $V_{\tau_{i+1}}^{(m)}$ from $f(V_{\tau_{i+1}} | V_{\tau_i}^{(m)}, \Theta)$, which can be easily done using equation (6).
- **Step 2:** Assign to $V_{\tau_{i+1}}^{(m)}$ a filtering weight of

$$\pi_{i+1}^{(m)} = \frac{w_{i+1}^{(m)}}{\sum_{k=1}^M w_{i+1}^{(k)}} \text{ where } w_{i+1}^{(m)} = f(S_{\tau_{i+1}} | V_{\tau_{i+1}}^{(m)}, \Theta)$$

- **Step 3:** Resample from the weighted sample $\{(V_{\tau_{i+1}}^{(m)}, \pi_{i+1}^{(m)}); m = 1, \dots, M\}$ to obtain a new equal-weight sample of size M .

Note that resampling is a step critical to the implementation of particle filtering. Without it, the variance of importance weights $\pi_{i+1}^{(m)}$ will grow over time (stochastically) and the quality of the particle filter is bound to be poor.

The SIR particle filter is not efficient because in drawing $V_{\tau_{i+1}}^{(m)}$ in step 1, one has not taken into account the knowledge of $S_{\tau_{i+1}}$. Naturally, one can contemplate a different convenient sampler based on some density $g(V_{\tau_{i+1}} | S_{\tau_{i+1}})$ in place of the sampler in step 1. One must modify the weight calculation in step 2, however. As a result of using a different sampler, the importance weight in step 2 should be replaced with

$$w_{i+1}^{(m)} = \frac{f(S_{\tau_{i+1}} | V_{\tau_{i+1}}^{(m)}, \Theta) \sum_{k=1}^M f(V_{\tau_{i+1}}^{(k)} | V_{\tau_i}^{(k)}, \Theta)}{g(V_{\tau_{i+1}}^{(m)} | S_{\tau_{i+1}})}$$

Unfortunately, this adaptive scheme will require of evaluating density functions in the order of M^2 as opposed to M for the standard SIR scheme.

In this paper, we adopt the auxiliary filtering idea of Pitt and Shephard (1999) to design our particle filter for the structural credit risk model. The basic idea is to enlarge the dimension of the state variable(s). In our case, we can think of it as sampling a pair $(V_{\tau_i}^{(m)}, V_{\tau_{i+1}}^{(m)})$ instead of a point $V_{\tau_{i+1}}^{(m)}$. Enlarging the dimension enables an easy calculation of the importance weight so that the number of density evaluations of the SIR scheme will remain in the order of M . After the weights have been determined, we discard the first entry of the pair and the weighted sample for $V_{\tau_{i+1}}^{(m)}$ represents the marginal filtering distribution.

The likelihood function evaluated via a particle filter is discontinuous with respect to the parameter Θ , making it ill-suited for gradient-based optimization and asymptotic statistical inference. For this reason, resampling will be conducted using the smooth bootstrap procedure of Pitt (2002). The generic version of the smoothed auxiliary SIR scheme can be described as follows:

- **Step 1:** Use an auxiliary sampler $g(V_{\tau_t}, V_{\tau_{i+1}} | \mathcal{D}_{i+1})$ to generate $(V_{\tau_i}^{(m)}, V_{\tau_{i+1}}^{(m)})$.
- **Step 2:** Compute the importance weight

$$w_{i+1}^{(m)} = \frac{f(S_{\tau_{i+1}} | V_{\tau_{i+1}}^{(m)}, \Theta) f(V_{\tau_{i+1}}^{(m)} | V_{\tau_i}^{(m)}, \Theta) \hat{f}(V_{\tau_i}^{(m)} | \mathcal{D}_i, \Theta)}{g(V_{\tau_i}^{(m)}, V_{\tau_{i+1}}^{(m)} | \mathcal{D}_{\tau_{i+1}})}. \quad (14)$$

and then assign $\pi_{i+1}^{(m)} = \frac{w_{i+1}^{(m)}}{\sum_{k=1}^M w_{i+1}^{(m)}}$ to the sample point $V_{\tau_{i+1}}^{(m)}$.

- **Step 3:** Resample from the smoothed empirical distribution constructed from the weighted sample $\{(V_{\tau_{i+1}}^{(m)}, \pi_{i+1}^{(m)}); m = 1, \dots, M\}$ to obtain a new equal-weight sample of size M .

Note that $w_{i+1}^{(m)}$ in (14) is the proper importance weight because its numerator is the filtering joint density of $(V_{\tau_i}^{(m)}, V_{\tau_{i+1}}^{(m)})$ up to a proportional constant. Thus, the marginal filtering distribution for $V_{\tau_{i+1}}$ can be simply represented by the weighted sample $\{(V_{\tau_{i+1}}^{(m)}, \pi_{i+1}^{(m)}); m = 1, \dots, M\}$.

The exact description of our sampler and the particle filtering scheme for the structural credit risk model of Merton (1974) is given in Appendix. In a nutshell, we construct a localized sampler that takes advantage of the knowledge on $S_{\tau_{i+1}}$. We localize the sampling of $V_{\tau_{i+1}}$ around the asset value implied by $S_{\tau_{i+1}}$ under no trading noise. We will refer to our specific scheme as the smoothed localized SIR (SL-SIR) particle filter.

The likelihood function value consists of the components that are in the form of the expression as in (11). The expectation can be computed as follows. Corresponding to $V_{\tau_i}^{(m)}$ in the equal-weight filtering sample, we draw $\nu_{i+1}^{(m)}$ and compute the value for the item inside the expectation operator. Repeat for all m 's and take an average to yield the desired value. This operation consists of the number of density calculations in the order of M . It turns out that there is no need to conduct additional sampling because the SL-SIR particle filter has already generated the item inside the expectation, which happens to coincide with the importance weight of the SL-SIR particle filter. In short, the SL-SIR particle filter is a practical order- M scheme.

2.4 Computing credit spread and default probability

The parameter estimates obtained from the MLE method are meant for credit risk applications. Typically, one is interested in knowing the credit spread of a risky corporate bond over the corresponding Treasury rate and/or the likelihood of a firm going bankrupt. Here we show how their estimates can be computed and statistical inference conducted in the filtering context.

In general, we can generically describe the quantities of interest as a function of the unobserved asset value at the last time point of the sample, V_{τ_N} , and the parameter value, Θ . Denote them by a generic function $H(V_{\tau_N}; \Theta)$. In the case of credit spread, the model of Merton (1974) gives rise to

$$C(V_{\tau_N}; \Theta) = -\frac{1}{T - \tau_N} \ln \left(\frac{V_{\tau_N}}{F} \Phi(-d_{\tau_N}) + e^{-r(T-\tau_N)} \Phi(d_{\tau_N} - \sigma\sqrt{T - \tau_N}) \right) - r \quad (15)$$

where d_t is defined in equation (4). For the default probability, the formula becomes

$$P(V_{\tau_N}; \Theta) = \Phi \left(\frac{\ln(F/V_{\tau_N}) - (\mu - \frac{1}{2}\sigma^2)(T - \tau_N)}{\sigma\sqrt{T - \tau_N}} \right) \quad (16)$$

Generically, the filtering estimate evaluated at the maximum likelihood parameter estimates is

$$h(\hat{\Theta}) \equiv E \left(H(V_{\tau_N}; \hat{\Theta}) \mid \mathcal{D}_N \right), \quad (17)$$

which can be straightforwardly computed with the SL-SIR particle filter. But one must recognize that the parameter values are the maximum likelihood estimates subject to an asymptotic joint normal distribution. Let the asymptotic distribution be denoted by

$$\hat{\Theta} \sim N \left(\Theta_0, A_N^{-1} \right), \quad A_N = - \left. \frac{\partial L(\Theta; \mathcal{D}_N)}{\partial \Theta \partial \Theta'} \right|_{\Theta = \hat{\Theta}} \quad (18)$$

where Θ_0 is the true parameter value vector and $L(\Theta; \mathcal{D}_N)$ is the log-likelihood function. For a differentiable transformation, $h(\hat{\Theta})$, the standard statistical result implies that

$$h(\hat{\Theta}) \sim N \left(h(\Theta_0), \left[\frac{\partial h(\hat{\Theta})}{\partial \Theta} \right]' A_N^{-1} \left[\frac{\partial h(\hat{\Theta})}{\partial \Theta} \right] \right) \quad (19)$$

where $\frac{\partial h(\hat{\Theta})}{\partial \theta}$ is a column vector of the first partial derivatives of the function with respect to the elements of Θ being evaluated at $\hat{\Theta}$.

This standard result was, for example, utilized by Duan (1994) in his the transformed-data MLE analysis. In contrast, the partial derivatives of $h(\hat{\Theta})$ in the current context do not have analytical solutions. The fact that the SL-SIR particle filter is smooth in relation to the parameters makes it possible to approximate the partial derivatives using numerical differences.

3 Empirical analysis

3.1 Dow Jones 30 firms

We implement the MLE method based on the SL-SIR particle filtering on the 30 companies that constitute the Dow Jones Industrial Index. Our data sample consists of daily equity

values of these firms over year 2003. The closing prices of equity and the numbers of shares outstanding are taken from the CRSP database.¹ The balance sheet information is from the Compustat annual files. The Dow Jones 30 companies are big firms with their stocks heavily traded. If trading noises are negligible, one would expect to find supporting evidence in this data set.

The initial maturity of debt is set to 10 years.² We take the book value of liabilities of a company at the year end of 2002 and compound it for 10 years at the risk-free interest rate. The resulting value is our proxy for the face of the debt for the model of Merton (1974). We set $h = 1/250$ to reflect the use of daily equity values. The risk-free interest rate is the 1-year Treasury constant maturity rate obtained from the U.S. Federal Reserve. We run the estimation using the 5000-particle SL-SIR filter.³

Table 1 reports the results for all Dow Jones firms. Names are given in the first column. The second, third and fourth columns contain the maximum likelihood estimates along with their asymptotic standard errors in brackets. The trading noise parameter is multiplied by a factor of 100. The estimated asset volatilities are stated per annum and consistent with the range that is expected of their values. The standard errors are very small, implying these parameters have been fairly accurately estimated. Due to the nature of diffusion models, the drift parameter is expected to have large sampling errors. Our results simply reconfirm this well-known phenomenon.

Although the estimates for trading noise are small in magnitude, they should be understood as a noise around a “true” value, not as a return volatility. Take 3M as an example ($\delta = 0.004044$). The number amounts to 0.4% of the 3M stock price if the trading noise is 1 standard deviation in either direction. In some cases, the trading noise estimate turns

¹The initial equity value is computed as the product of the closing price of equity and number of shares outstanding reported in CRSP. The subsequent equity values are constructed using the daily holding period returns in CRSP so as to account for dividends and stock splits. In CRSP, the price-related quantities are based on the closing transaction prices whenever possible. Otherwise, the average of closing bid and ask is used.

²We have repeated the estimation with an initial maturity of $T = 2$. The results are qualitatively similar to the ones reported in the paper.

³Our experience indicates that 1000 particles are large enough for this particular estimation problem. In the simulation study conducted later, we will use 1000 particles.

out to be zero; for example, American Express. This result may indeed be because American Express faces negligible trading noises. It is also possible that it reflects a variety of factors such as statistical sampling error, the mis-specification of Merton's model and the mis-specified trading noise structure.

We conduct a likelihood ratio (LR) test on the hypothesis of no trading noise; that is, $\delta = 0$. Since the parameter value under the null hypothesis locates right at the lower bound of the parameter set, we must conduct inference in a way that correctly reflects the testing situation. It turns out that the LR test statistic distributes asymptotically as a mixture of 1/2 point mass at 0 and 1/2 chi-square with 1 degree of freedom. This leads to a simple correction to the tail probability calculation. One only needs to cut in half the usual p -value from the chi-square distribution (see, for example, Gouriéroux and Monfort (1995), Chapter 21). The results reported in column five of Table 1 and the summary results in Table 2 suggest that there are 6 out of 30 Dow Jones companies facing significant trading noise at the 5% level.

To examine the effect of trading noise from a practical angle, we ask a what-if question. Suppose we ignore trading noise and proceed with the estimation using the transformed-data MLE method of Duan (1994). What magnitude of bias in volatility will the omission cause? In the sixth column of Table 1, we report the ratio of the estimated asset volatility without trading noise over the one with. Of course, we expect the omission to increase the volatility estimate, because trading noises have been erroneously treated as the genuine asset volatility. In the case of 3M, for example, the omission causes the asset volatility to be overestimated by 15.9%. All in all, the upward bias is in the order of 6.66% on average (see Table 2) with the maximum bias at 23.8% and the minimum at 0%. In summary, the omission can have material impact even for the Dow Jones 30 companies.

3.2 Randomly selected firms from CRSP

It is reasonable to expect that the Dow Jones 30 companies are subject to smaller trading noises. The results reached in the preceding subsection thus cannot represent the impact of trading noise for a typical U.S. exchange listed firm. We set out to analyze a randomly

chosen sample of 100 firms from the CRSP database. A randomly selected firm is included in the sample only if it has the required CRSP and Compustat data for year 2003 and it must not be a firm already in the Dow Jones sample.

For this randomly selected sample, we implement the MLE in the manner identical to that for the Dow Jones sample. To conserve space, we only report the summary statistics for this sample in Table 3. As expected, the results are stronger for these firms than those for the Dow Jones 30 companies. For 30 out of the 100 firms, the null hypothesis of no trading noise is rejected at the 5% significance level. The upward bias in the volatility estimate due to ignoring trading noises can reach as high as 93.78%.⁴ There are 10% of the firms experiencing a 44.57% or higher upward bias. In short, it is important to recognize trading noise in implementing the structure credit risk model.

3.3 Microstructure noises or mis-specification errors?

Our estimates for the magnitude of trading noise, δ , depend on the use of the Merton (1974) model. Needless to say, a highly stylized model such as Merton's is likely to embed mis-specification errors which may in turn lead to questionable empirical conclusions. To ascertain whether our trading noise estimates are in line with one's prior belief on microstructure noises, we conduct a cross-sectional analysis of the estimated δ 's in relation to the commonly adopted proxies for market liquidity. Intuitively, a more liquid firm, meaningfully measured, should have a smaller δ if Merton's credit risk model is not grossly mis-specified.

Table 4 reports the results of this analysis for the 100 randomly selected firms as described in the preceding subsection. The first proxy is the percentage bid-ask spread, and it is expected to be positively related to microstructure noises. We use the CRSP daily files to compute the difference of the closing ask and bid over the closing bid. For each firm in the sample, we take the average of the daily values over the sample period, i.e., 2003. The cross-sectional average of the percentage bid-ask spreads turns out to be 1.21% while the 100-firm average estimate for δ is roughly 0.6%. So our estimated trading noise seems to be in the same order of magnitude as the expected bid-ask bounce. Table 4 also reports the

⁴The minimum ratio is 0.9987 as opposed to 1, a result due to numerical precision.

regression result that the percentage bid-ask spread is positively and significantly related to the estimated δ , indicating that the trading noise estimates are intuitively plausible. Also reported in the table is the Spearman rank correlation between the estimated δ and the percentage bid-ask spread, which also indicates a strong positive relationship.

The second proxy adopted is the firm size, measured as the market equity capitalization of the firm and computed on the first day of 2003 as the product of the number of shares outstanding and the equity price. This variable is expected to be negatively related to microstructure noises (see for instance Roll (1984)). To have this variable in a percentage sense, we follow Roll (1984) to work with the logarithm of the firm size. The second row in Table 4 confirms a statistically significant negative relationship between δ and the size of the firm. Again, the Spearman rank correlation yields a similar conclusion.

Finally, we investigate how trading volume is related to the estimated magnitude of trading noise. The trading volume is the average of the daily volumes from the CRSP daily file in 2003. Again we expect to see a negative relationship between the estimated δ and this liquidity proxy. The third row of Table 4 indicates the expected direction of the relationship but the estimate is statistically insignificant, measured either by the regression result or the Spearman rank correlation.

The results in Table 4 taken together indicate clearly that our estimated magnitudes of trading noises are intuitively plausible.⁵ A stylized model like Merton (1974) can produce sensible results when its estimation is carried out in a statistically rigorous manner. In summary, our estimation procedure indeed picks up trading noises and can filter them out so that the implementation of the Merton-like models need not be adversely affected by ignoring microstructure noises.

⁵For the Dow Jones 30 sample alone, the estimated δ 's are not correlated with the liquidity proxies. When the two samples are combined, the results are qualitatively the same as those for the 100 randomly selected firms.

4 Simulation analysis

To ascertain the finite-sample performance of our MLE method, we run a simulation experiment. We generate sample paths of noisy equity observations by controlling the end-of-sample pseudo-leverage, i.e., $\frac{E}{V}$. This can be achieved rather easily in the context of Merton's model because the asset's continuously compounded returns are independent. In short, we generate the 250 daily returns and then construct the firm's asset values backward to yield a sample of 251 asset values. Corresponding to the simulated asset value sample, we compute 251 equity values using the measurement equation in (5). For estimation, we act as if we did not know the asset values to mimic the real-life estimation situation. Estimation is performed for each simulated sample and the corresponding asymptotic inference is conducted.

The parameter values are chosen in a way that is consistent with the real data. For the baseline case, we use the median values (rounded) obtained from the 100 randomly chosen firms. The parameter values are $\sigma = 0.3$, $\delta = 0.004$ and $\mu = 0.2$. The ending pseudo-leverage ratio is maintained at 40%. The initial maturity is set to 10 years and gradually declines to 9 years at the end of the simulated sample. We simulate 500 samples in each case. We also vary the two key parameters, σ and δ , to investigate their effect on performance. The parameter values used are $\sigma = 0.7$ and $\delta = 0.016$, which are close to the 90 percentile of the estimates obtained from the 100 randomly chosen firms. We vary one parameter value at a time and run the 1000-particle SL-SIR filter.

Table 5 presents the simulation results for the baseline case. Both median and mean values of all parameter estimates are close to the true values, indicating a pretty good finite-sample behavior of the MLE.

With the exception for δ , the obtained coverage rates suggest that the asymptotic distribution is adequate in describing the sampling property of the estimates.⁶ For the trading

⁶The coverage rate is defined as the percentage of the parameter estimates for which the true parameter value is contained in the α confidence interval implied by the asymptotic distribution. Care is needed in dealing with the cases where the estimated δ is on the lower boundary, i.e., 0. Since δ is not an interior solution, the standard Taylor expansion used to obtain the asymptotic distribution ceases to apply. We thus drop all cases where the δ estimate is zero in computing its coverage rates. For σ and μ , we use the corresponding entries of the Fisher information matrix to get their variances because they are the interior solutions in the parameter set of a reduced dimension.

noise parameter δ , the coverage rates are biased. This may be due to the fact that we have used a relatively low value of δ . Thus, for a relatively small sample of 250, the behavior of this estimator bears resemblance to the situation where the true parameter value is on the boundary of the parameter set. When the parameter value is actually on the boundary, we of course know that the standard asymptotic theory breaks down.

The preceding conjecture about the trading noise parameter is confirmed by the results in Table 6. This table contains the simulation results based on a higher value of δ . We have increased the true parameter value from 0.004 to 0.016 while keeping all other parameters fixed at their previous values. The bias in the coverage rate for δ disappears. The simulation results suggest that the asymptotic inference can be reliably applied to all three parameters using one year worth of daily data when the magnitude of trading noise is large.

Finally, we examine the effect of having a large asset volatility. Table 7 presents the results when the value of σ is increased from 0.3 to 0.7 while keeping other parameters identical to those in the baseline case. The coverage rates suggest the same bias exists for the trading noise parameter δ as in the baseline case, a result that we have argued earlier can be attributed to a low value of δ . Interestingly, the mean and median trading noise parameter estimates become much higher than the true parameter value, in contrast with the earlier results for the baseline case. Overall, we can conclude that a large asset volatility does not fundamentally alter the quality of the estimation procedure except for the trading noise parameter.

Finally, we analyze whether the LR test for the presence of trading noise has a right size. We use the baseline case to generate 500 samples except that δ is set to 0. The results reported in Table 7 suggest that the empirical size for the 5% test is 6.6%, only slightly upward biased. The result for the 10% test is similar with an empirical rejection rate of 11.4%. We then vary the value of δ to examine the power of the LR test. The results indicate a reasonable power; for example, when $\delta = 0.01$, one can expect to reject the hypothesis of no trading noise 66.6% of times using the 5% LR test.

5 Conclusion

We have developed a particle filter-based MLE method for the structural credit risk model of Merton (1974). Our empirical analysis on both the Dow Jones 30 companies and 100 randomly selected firms ascertains the importance of recognizing trading noise. Although our methodological development is presented specifically for Merton's model, the method can be easily adapted to other structural credit risk models. This is in a way similar to the fact that the transformed-data MLE method of Duan (1994, 2000) can be applied to general structural credit risk models under no trading noise. In this paper, the trading noise is assumed to have a log-Gaussian distribution. It is also straightforward to allow for different distributional assumptions. In conclusion, a practical MLE method has been developed in this paper to apply structural credit risk models to a market in which trading noises are likely present.

Appendix: The SL-SIR particle filter for Merton's model

Our localized sampling scheme starts with $V_{\tau_i}^{(m)}$ in the equal-weight filtering sample. Because the trading noise is independent of the unobserved asset value, we can draw $\nu_{i+1}^{(m)}$, which follows the standard normal distribution. We then compute $V_{\tau_{i+1}}^{(m)} = V_{\tau_{i+1}}^*(S_{\tau_{i+1}}, \nu_{i+1}^{(m)})$ where

$$V_{\tau_{i+1}}^*(S_{\tau_{i+1}}, \nu_{i+1}) = S^{-1}(S_{\tau_{i+1}} e^{-\delta \nu_{i+1}}; \sigma, F, r, T - \tau_{i+1})$$

defines the implied asset value by the measurement equation in (5). Since $V_{\tau_{i+1}}^*(S_{\tau_{i+1}}, \nu_{i+1})$ is a function of ν_{i+1} , the standard differential transformation theory can be used to obtain the density function for it as

$$f(V_{\tau_{i+1}}^*(S_{\tau_{i+1}}, \nu_{i+1}^{(m)}) | S_{\tau_{i+1}}) = \frac{\Phi(d_{\tau_{i+1}}^{*(m)}) e^{\delta \nu_{i+1}^{(m)}}}{\delta S_{\tau_{i+1}} \phi(\nu_{i+1}^{(m)})}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the standard normal distribution and density functions, respectively; $d_{\tau_{i+1}}^{*(m)}$ is the function d_{t+1} as defined in (4) being evaluated at $V_{\tau_{i+1}}^*(S_{\tau_{i+1}}, \nu_{i+1}^{(m)})$.

Thus, this localized sampler leads to the following joint density:

$$\begin{aligned} g(V_{\tau_i}^{(m)}, V_{\tau_{i+1}}^{(m)} \mid \mathcal{D}_{\tau_{i+1}}) &= f(V_{\tau_{i+1}}^*(S_{\tau_{i+1}}, \nu_{i+1}^{(m)} \mid S_{\tau_{i+1}}) \hat{f}(V_{\tau_i}^{(m)} \mid \mathcal{D}_i, \Theta) \\ &= \frac{\Phi(d_{\tau_{i+1}}^*(m)) e^{\delta \nu_{i+1}^{(m)}}}{\delta S_{\tau_{i+1}} \phi(\nu_{i+1}^{(m)})} \hat{f}(V_{\tau_i}^{(m)} \mid \mathcal{D}_i, \Theta) \end{aligned}$$

And the corresponding importance weight as defined in (14) becomes

$$\begin{aligned} w_{i+1}^{(m)} &= \frac{f(S_{\tau_{i+1}} \mid V_{\tau_{i+1}}^{(m)}, \Theta) f(V_{\tau_{i+1}}^{(m)} \mid V_{\tau_i}^{(m)}, \Theta) \hat{f}(V_{\tau_i}^{(m)} \mid \mathcal{D}_i, \Theta)}{g(V_{\tau_i}^{(m)}, V_{\tau_{i+1}}^{(m)} \mid \mathcal{D}_{\tau_{i+1}})} \\ &= \frac{f(S_{\tau_{i+1}} \mid V_{\tau_{i+1}}^{(m)}, \Theta) \delta S_{\tau_{i+1}} \phi(\nu_{i+1}^{(m)}) f(V_{\tau_{i+1}}^{(m)} \mid V_{\tau_i}^{(m)}, \Theta)}{\Phi(d_{\tau_{i+1}}^*(m)) e^{\delta \nu_{i+1}^{(m)}}} \\ &= \frac{f(V_{\tau_{i+1}}^{(m)} \mid V_{\tau_i}^{(m)}, \Theta)}{\Phi(d_{\tau_{i+1}}^*(m)) e^{\delta \nu_{i+1}^{(m)}}} \end{aligned}$$

The last equality is true because $f(S_{\tau_{i+1}} \mid V_{\tau_{i+1}}^{(m)}, \Theta) = \frac{1}{\delta S_{\tau_{i+1}} \phi(\nu_{i+1}^{(m)})}$.

We now summarize the three-step SL-SIR scheme for the structural credit risk model of Merton (1974). The SL-SIR particle filter starts at $V_{\tau_0}^{(m)} = V_{\tau_0}^*(S_{\tau_0}, 0)$ for all m 's. The system is advanced via the following three-step procedure:

- **Step 1:** Begin with $V_{\tau_i}^{(m)}$ in the equal-weight filtering sample. Draw a standard normal $\nu_{i+1}^{(m)}$ and compute $V_{\tau_{i+1}}^{(m)} = V_{\tau_{i+1}}^*(S_{\tau_{i+1}}, \nu_{i+1}^{(m)})$ to obtain the pair $(V_{\tau_i}^{(m)}, V_{\tau_{i+1}}^{(m)})$.
- **Step 2:** Compute the importance weight

$$w_{i+1}^{(m)} = \frac{f(V_{\tau_{i+1}}^{(m)} \mid V_{\tau_i}^{(m)}, \Theta)}{\Phi(d_{\tau_{i+1}}^*(m)) e^{\delta \nu_{i+1}^{(m)}}}$$

and then assign $\pi_{i+1}^{(m)} = \frac{w_{i+1}^{(m)}}{\sum_{k=1}^M w_{i+1}^{(m)}}$ to the sample point $V_{\tau_{i+1}}^{(m)}$. Note that by equation (6)

$$f(V_{\tau_{i+1}}^{(m)} \mid V_{\tau_i}^{(m)}, \Theta) = \frac{\phi\left(\frac{\ln(V_{\tau_{i+1}}^{(m)}/V_{\tau_i}^{(m)}) - (\mu - \frac{\sigma^2}{2})h}{\sigma\sqrt{h}}\right)}{V_{\tau_i}^{(m)} \sigma\sqrt{h}}.$$

- **Step 3:** Construct a piecewise linear empirical distribution using the weighted sample $\{(V_{\tau_{i+1}}^{(m)}, \pi_{i+1}^{(m)}); m = 1, \dots, M\}$. Use it to resample a new equal-weight sample of size M .

In line with the arguments of Pitt (2002), the importance weight in Step 2 is exactly the item inside the expectation operator in (11). The conditional likelihood based on the observed equity values from τ_i to τ_{i+1} can thus be computed as the average of the importance weights. Specifically,

$$f(S_{\tau_{i+1}} | \mathcal{D}_i, \Theta) = \frac{1}{M} \sum_{m=1}^M w_{i+1}^{(m)}.$$

References

- Ait-Sahalia, Y., Mykland, P. and Zhang, L. (2005a), ‘How often to sample a continuous-time process in the presence of market microstructure noise’, *Review of Financial Studies* .
- Ait-Sahalia, Y., Mykland, P. and Zhang, L. (2005b), ‘A tale of two time scales: Determining integrated volatility with noisy high-frequency data’ , *Journal of the American Statistical Association* **100**, 1394–1411.
- Bandi, F. and Russell, J. (2006), ‘Separating microstructure noise from volatility’, *Journal of Financial Economics* **79**, 655-692.
- Black, F. and Scholes, M. (1973), ‘The pricing of options and corporate liabilities’, *Journal of Political Economy* **81**, 637–659.
- Doucet, A., de Freitas, N. and Gordon, N., eds (2001), *Sequential Monte Carlo Methods in Practice*, Springer-Verlag.
- Duan, J.-C. (1994), ‘Maximum likelihood estimation using price data of the derivative contract’, *Mathematical Finance* **4**, 155–167.
- Duan, J.-C. (2000), ‘Correction: Maximum likelihood estimation using price data of the derivative contract’, *Mathematical Finance* **10**, 461–462.
- Duan, J.-C., Gauthier, G. and Simonato, J.-G. (2004), ‘On the equivalence of the kmv and maximum likelihood methods for structural credit risk models’, Working Paper, University of Toronto.
- Duan, J.-C., Gauthier, G., Simonato, J.-G. and Zaanoun, S. (2003), ‘Estimating Mertons model by maximum likelihood with survivorship consideration’, Working Paper, University of Toronto.
- Duan, J.-C. and Simonato, J.-G. (2002), ‘Maximum likelihood estimation of deposit insurance value with interest rate risk’, *Journal of Empirical Finance* **9**, 109–132.

- Duan, J.-C. and Yu, M.-T. (1994), ‘Assessing the cost of taiwans deposit insurance’, *Pacific Basin Finance Journal* **2**, 73–90.
- Ericsson, J. and Reneby, J. (2004a), ‘Estimating structural bond pricing models’, *Journal of Business* .
- Ericsson, J. and Reneby, J. (2004c), ‘Valuing corporate liabilities’, Working Paper, McGill University.
- Gordon, N., Salmond, D. and Smith, A. (1993), ‘A novel approach to non-linear and non-gaussian bayesian state estimation’, *IEEE Proceedings F* **140**, 107–113.
- Gourieroux, C. and Monfort, A. (1995), *Statistics and Econometric Models*, Cambridge University Press.
- Harris, L. (1990), ‘Estimation of stock price variances and serial covariances from discrete observations’, *Journal of Financial and Quantitative Analysis* **25**, 291–306.
- Hasbrouck, J. (1993), ‘Assessing the quality of a security market: A new approach to transaction-cost measurement’, *Review of Financial Studies* **6**, 191–212.
- Ionides, E. (2004), ‘Inference and filtering for partially observed diffusion processes via sequential monte carlo’, Working Paper, University of Michigan.
- Jarrow, R. and Turnbull, S. (2000), ‘The intersection of market and credit risk’, *Journal of Banking and Finance* **24**, 271–299.
- Laeven, L. (2002), ‘Bank risk and deposit insurance’, *World Bank Economic Review* **16**, 109–137.
- Lehar, A. (2004), ‘Measuring systemic risk: A risk management approach’, *Journal of Banking and Finance* .
- Madhavan, A., Richardson, M. and Roomans, M. (1997), ‘Why do security prices change? a transaction-level analysis of nyse stocks’, *Review of Financial Studies* **10**, 1035–1064.

- Merton, R. C. (1974), 'On the pricing of corporate debt: The risk structure of interest rates', *Journal of Finance* **29**, 449–70.
- Merton, R. C. (1977), 'An analytic derivation of the cost of deposit insurance and loan guarantees', *Journal of Banking and Finance* **1**, 311.
- Pitt, M. (2002), 'Smooth particle filters likelihood evaluation and maximisation', Working Paper, University of Warwick.
- Roll, R. (1984), 'Simple implicit measure of the effective bid-ask spread in an efficient market', *Journal of Finance* **39**, 1127–1139.
- Pitt, M. and Shephard, N. (1999), 'Filtering via simulation: auxiliary particle filter', *Journal of the American Statistical Association* **94**, 590–599.
- Wong, H. and Choi, T. (2004), 'The impact of default barrier on the market value of firm's asset', Working Paper, Chinese University of Hong Kong.

Table 1: Maximum likelihood estimation for Dow Jones 30 companies, firm-by-firm

Name	Estimates with trading noise (s.e.)			LR test	$\frac{\sigma_{wo}}{\sigma}$
	σ	$\delta \times 100$	μ	p -value	
3M	0.1318 (0.0089)	0.4044 (0.0919)	0.2798 (0.1358)	0.0194	1.1587
Alcoa	0.1589 (0.0181)	0.6820 (0.2082)	0.3130 (0.1640)	0.0645	1.1640
Altria	0.1783 (0.0060)	0.0000 (8.2517)	0.2437 (0.1816)	0.4991	1.0000
American Express	0.0726 (0.0039)	0.0000 (5.7897)	0.0970 (0.0749)	0.4966	1.0000
American Intl	0.0792 (0.0039)	0.0000 (7.1344)	0.0424 (0.0802)	0.4979	1.0001
Boeing	0.1132 (0.0058)	0.0000 (6.7798)	0.1179 (0.1132)	0.4976	1.0001
Caterpillar	0.1215 (0.0044)	0.0000 (6.2252)	0.2847 (0.1235)	0.4990	1.0000
Citigroup	0.0431 (0.0026)	0.0000 (5.7291)	0.0713 (0.0425)	0.4972	1.0001
E.I. du Pont	0.1306 (0.0098)	0.3248 (0.1863)	0.0693 (0.1302)	0.2009	1.0601
Exxon	0.1108 (0.0078)	0.4159 (0.0818)	0.1383 (0.1136)	0.0073	1.1876
General Electric	0.0857 (0.0044)	0.0000 (5.8057)	0.0950 (0.0860)	0.4969	1.0001
General Motors	0.0183 (0.0011)	0.0000 (5.9029)	0.0403 (0.0185)	0.5000	1.0003
Hewlett Packard	0.2511 (0.0281)	0.6391 (0.3831)	0.2079 (0.2887)	0.1416	1.0906
Honeywell	0.1493 (0.0130)	0.4025 (0.2303)	0.2012 (0.1494)	0.2011	1.0651
IBM	0.1412 (0.0110)	0.3696 (0.2020)	0.1119 (0.1415)	0.1282	1.0754
Intel	0.3130 (0.0278)	0.6103 (0.3089)	0.6686 (0.3124)	0.1237	1.0819
J.P. Morgan	0.0275 (0.0017)	0.0000 (6.7264)	0.0520 (0.0271)	0.4976	1.0002
Johnson & Johnson	0.1521 (0.0115)	0.5547 (0.0937)	-0.0394 (0.1532)	0.0035	1.2380
McDonalds	0.2146 (0.0076)	0.0000 (8.0757)	0.3134 (0.2216)	0.4997	1.0000
Merck	0.1986 (0.0072)	0.0041 (5.4404)	-0.0951 (0.2064)	0.4937	1.0000
Microsoft	0.2441 (0.0197)	0.5675 (0.2138)	0.0576 (0.2543)	0.0570	1.1150
Pfizer	0.1950 (0.0121)	0.2417 (0.2239)	0.1372 (0.1950)	0.3582	1.0293
SBC	0.1981 (0.0078)	0.0000 (7.5127)	-0.0027 (0.1984)	0.4978	1.0000
Coca-Cola	0.1739 (0.0112)	0.2502 (0.1830)	0.1457 (0.1752)	0.2538	1.0410
Home Depot	0.2703 (0.0216)	0.4254 (0.3441)	0.3657 (0.3021)	0.2441	1.0478
Procter & Gamble	0.0962 (0.0070)	0.2653 (0.0798)	0.1277 (0.0967)	0.0473	1.1190
United Technologies	0.1341 (0.0105)	0.3625 (0.1665)	0.2804 (0.1338)	0.1729	1.0714
Verizon	0.1277 (0.0111)	0.3659 (0.2843)	-0.0265 (0.1299)	0.2860	1.0542
Wal-Mart	0.1538 (0.0125)	0.5261 (0.1260)	0.0437 (0.1539)	0.0077	1.1866
Walt Disney	0.1601 (0.0128)	0.7463 (0.1177)	0.2041 (0.1638)	0.0206	1.2123

Table 2: Summary of the maximum likelihood estimation for Dow Jones 30 companies

	Estimates with trading noise			$\frac{\sigma_{wo}}{\sigma}$
	σ	$\delta \times 100$	μ	
Mean	0.1482	0.2719	0.1515	1.0666
Median	0.1453	0.2950	0.1228	1.0510
10 Percentile	0.0578	0.0000	-0.0146	1.0000
90 Percentile	0.2476	0.6247	0.3132	1.1871
Min	0.0183	0.0000	-0.0951	1.0000
Max	0.3130	0.7463	0.6686	1.2380

Number of rejections of $H_0 : \delta = 0$ at 5% significance is 6 out of 30

Table 3: Summary of the maximum likelihood estimation for 100 randomly selected companies

	Estimates with trading noise			$\frac{\sigma_{wo}}{\sigma}$
	σ	$\delta \times 100$	μ	
Mean	0.3366	0.6288	0.4263	1.1350
Median	0.2518	0.4266	0.1719	1.0520
10 Percentile	0.0400	0.0000	0.0160	1.0000
90 Percentile	0.7193	1.6252	1.3542	1.4457
Min	0.0038	0.0000	-0.4526	0.9987
Max	1.1759	6.8155	2.5956	1.9378

Number of rejections of $H_0 : \delta = 0$ at 5% significance is 30 out of 100

Table 4: Relationship of the estimated δ 's with alternative measures of trading noises for 100 randomly selected companies

Explanatory variables	$\delta_i = \alpha + \beta x_i$	Spearman rank correlation
Intercept	0.0005 (0.4364)	
Percentage Bid-Ask Spread	0.4540 (0.0000)	0.3398 (0.0005)
Intercept	0.0165 (0.0000)	
log(FirmSize)	-0.0018 (0.0000)	-0.2167 (0.0302)
Intercept	0.0112 (0.0186)	
log(TradingVolume)	-0.0004 (0.2750)	-0.1332 (0.1862)

The percentage bid-ask spread is the average of the daily closing ask minus the closing bid over the closing bid from the CRSP daily file. The firm size is the market equity capitalization on the first day of the year from the CRSP file, computed as the product of the number of shares outstanding and the equity price. The trading volume is the average of the daily volumes from the CRSP daily file. p -values are in parentheses.

Table 5: Simulation results with median parameter values

True Parameters	Estimates with noise			$\frac{\sigma_{w\theta}}{\sigma}$
	σ	$\delta \times 100$	μ	
	0.3	0.4	0.2	
Mean	0.2925	0.4058	0.2121	1.0585
Median	0.2937	0.4368	0.2035	1.0363
St. Dev.	0.0223	0.3343	0.3099	0.0686
10 percentile	0.2651	0.0001	-0.1661	0.9998
90 percentile	0.3213	0.8496	0.6063	1.1537
Min	0.2105	0.0001	-0.7304	0.9972
Max	0.3456	1.2542	1.1250	1.4074
25 % coverage	0.2680	0.4513	0.2220	
50 % coverage	0.5040	0.5769	0.4900	
75 % coverage	0.7640	0.7333	0.7220	
95 % coverage	0.9460	0.8897	0.9340	

Number of δ estimates equal to zero is 110 out of 500

This table presents the results of a Monte-Carlo experiment using 500 independent samples with each consisting of 251 daily observations. The parameters used to simulate the data are set to the median values (rounded) of the 100 randomly chosen firms. The ending pseudo-leverage ratio is maintained at 40%. The 1000-particle SL-SIR filter is used to produce the results. For δ we drop the cases where $\delta = 0$ and compute the coverage rates using the remaining sample. The coverage rates for are based on the entire sample of 500. The standard errors for σ and μ are computed using the corresponding entries of the Fisher information matrix.

Table 6: Simulation results with high δ

True Parameters	Estimates with noise			$\frac{\sigma_{wo}}{\sigma}$
	σ	$\delta \times 100$	μ	
	0.3	1.6	0.2	
Mean	0.2975	1.5992	0.2145	1.4750
Median	0.2969	1.6191	0.2104	1.4605
St. Dev.	0.0330	0.2460	0.3122	0.1772
10 percentile	0.2580	1.2855	-0.1670	1.2439
90 percentile	0.3388	1.8805	0.5981	1.7068
Min	0.2123	0.0001	-0.7454	1.0003
Max	0.4261	2.2198	1.0998	2.2317
25 % coverage	0.2460	0.2485	0.2400	
50 % coverage	0.4960	0.5050	0.4980	
75 % coverage	0.7460	0.7495	0.7360	
95 % coverage	0.9320	0.9419	0.9340	

Number of δ estimates equal to zero is 1 out of 500

This table presents the results of a Monte-Carlo experiment using 500 independent samples with each consisting of 251 daily observations. The parameter values for σ and μ used in simulation are set to the median values (rounded) of the 100 randomly chosen firms. The value for δ is chosen to be close to the 90 percentile of the 100 randomly chosen firms. The ending pseudo-leverage ratio is maintained at 40%. The 1000-particle SL-SIR filter is used to produce the results. For δ we drop the cases where $\delta = 0$ and compute the coverage rates using the remaining sample. The coverage rates for are based on the entire sample of 500. The standard errors for are computed using the corresponding entries of the Fisher information matrix.

Table 7: Simulation results with high σ

True Parameters	Estimates with noise			$\frac{\sigma_{\text{ML}}}{\sigma}$
	σ	$\delta \times 100$	μ	
	0.7	0.4	0.2	
Mean	0.6747	0.6399	0.2181	1.0463
Median	0.6772	0.5459	0.1991	1.0169
St. Dev.	0.0508	0.6102	0.7209	0.0636
10 percentile	0.6140	0.0001	-0.6627	0.9997
90 percentile	0.7403	1.5317	1.1338	1.1358
Min	0.4731	0.0001	-1.9695	0.9948
Max	0.8055	2.3387	2.3599	1.3816
25 % coverage	0.2480	0.3968	0.2260	
50 % coverage	0.5020	0.5228	0.4820	
75 % coverage	0.7420	0.6702	0.7220	
95 % coverage	0.9480	0.8445	0.9320	

Number of δ estimates equal to zero is 127 out of 500

This table presents the results of a Monte-Carlo experiment using 500 independent samples with each consisting of 251 daily observations. The parameter values for δ and μ used in simulation are set to the median values (rounded) of the 100 randomly chosen firms. The value for σ is chosen to be close to the 90 percentile of the 100 randomly chosen firms. The ending pseudo-leverage ratio is maintained at 40%. The 1000-particle SL-SIR filter is used to produce the results. For δ we drop the cases where $\delta = 0$ and compute the coverage rates using the remaining sample. The coverage rates for are based on the entire sample of 500. The standard errors for are computed using the corresponding entries of the Fisher information matrix.

Table 8: Size and power of the LR test

$\delta \times 100$	0	0.2	0.4	1	1.6
Rejection rate at the 5% level	0.066	0.072	0.116	0.666	0.980
Rejection rate at the 10% level	0.114	0.134	0.200	0.768	0.992

This table presents the results of a Monte-Carlo experiment on the size and power of the LR test of $\delta = 0$. The rejection rates (500 samples) for two significance levels are reported. The ending pseudo-leverage ratio is maintained at 40%. $\sigma = 0.3$, $\mu = 0.2$ and the 1000-particle SL-SIR filter is used.