國立臺灣大學管理學院財務金融學系

博士論文

Department of Finance

College of Management

National Taiwan University

Doctoral Dissertation

利用股價與選擇權的數據來估計 GARCH 選擇權定價模型

Using Stock and Options Data to Estimate the GARCH Options Pricing Model

鄭宏文

Hung-Wen Cheng

指導教授：傅承德 博士

王耀輝 博士

Advisor: Cheng-Der Fuh, Ph.D.

Yaw-Huei Wang, Ph.D.

中華民國 100 年 6 月

June, 2011

國立臺灣大學博士學位論文
口試委員會審定書

利用股價與選擇權的數據來估計 GARCH 選擇權定價模型

Using Stock and Options Data to Estimate the GARCH Options Pricing Model

本論文係 鄭宏文 君（D94723006）在國立臺灣大學財務金融所完成之博士學位論文，於民國一百年六月二十七日承下列考試委員審查通過及口試及格，特此證明

口試委員：

王耀輝（簽名）

（指導教授）

金慶剛

張森林　王耀輝

廖四郎　葉小蓁

系主任、所長　胡星陽　（簽名）

誌謝

本篇論文能順利完成，獲得許多人的協助跟支持，在此獻上我最誠摯的感謝。

首先，感謝我的指導老師 傅承德教授，在我讀博士班這五年給予的指導及提攜，帶領我進入財務工程領域，從如何找問題、看問題及想問題到如何解決問題，總是能用淺而易懂的方式讓我了解，這對我在研究上是一大助益。除學術上的教導，傅老師也在生活上給予我很多寶貴的人生經驗，著實受益良多。其次，要感謝我在中研院統計所的老闆 銀慶剛教授，帶領我進入計量經濟學領域，銀老師對做研究嚴謹及追求完美的態度，使我在做研究過程中能更紮實且更謹慎的完成。除研究之餘，銀老師也給了我很多生活上的幫助，使我受惠頗多。再者，要感謝上海交大高級金融學院 張純信教授（Charles Chang），帶領我進入財務實證領域，在學業上給予我很多財務知識及國際觀，在研究上要求完美的個性，使我學習及成長很多，而生活上也給予我很多的考驗與挑戰，讓我學習到做任何事都需全力以赴以達到完美境界為目標而邁進，真是亦師亦友的好伙伴。

在論文口試時，感謝口試委員張森林教授、王耀輝教授、葉小蓁教授及廖四郎教授給予寶貴的建議使得論文更加完善。在台灣大學財金所就讀期間，感謝系上黃達業教授、管中閔教授、胡星陽教授及陳業寧教授的指導及幫助。特別要感謝黃達業教授南下高雄於我婚禮上致詞，為婚禮增添了許多風采。也要感謝我的合著者俞淑惠教授及林建志教授，在討論研究的過程中，激發很多火花使得我們論文能更完整的呈現。再來，要感謝我的師兄林士貴教授及王仁和教授，分享很多他們的意

見與經驗，讓我少走一些冤枉路。以及德峰和俊英，一起討論合作論文，各方不同的意見促使我們有不錯的發現。另外，感謝麗君、柏欣和玄啟，在我遇到處理資料上的困難時，幫助我解決我的問題。感謝蔡董事長志鴻和惠菁，在我有煩惱及困擾時，找你們都能給予我相當的幫助。感謝智勇，在論文上的幫忙。還有，胡羚學姐、俊華學長、正皇學長、育展學長、桂欽學長、鴻崑學長、彥行學長、裴紋學姐、宗岡學長、懿宸學長、玫惠學姐、Uncle山林、景德、子綾，財博籃球隊成員們嘉威、崇賢、宏瑋、陽照，同師門的昭玫學姐、汝芳、盛豐、慈陽、竹嵐、士綱，顥嘉、思綺、維哲、志彥、柏蓉、仁維、以達以及中研院統計所的同仁們等，謝謝你們讓我的博士班生涯充滿回憶。

感謝親愛的老婆—茹惠，對我的包容與付出，幫我照顧好我們的兒子兆鈞，讓我能無後顧之憂的順利完成學業。也要感謝兒子兆鈞的出生，使我更有責任且更有動力的往前走。最後，感謝我父、母、岳父及岳母的支持與鼓勵，將此論文獻給你們，以報答對我的養育之恩。

鄭宏文　　　　　謹誌於

台灣大學財務金融研究所

民國一百年六月二十七日

中文摘要

這個研究推導了當只用股票數據（ST）、只用選擇權數據（OT）及用股票和選擇權數據有含（S+O+E）或沒有含誤差項（S+O）時的 GARCH(1,1) 選擇權價格模型估計的漸近特性。在大樣本理論下的漸近變異數說明了只用選擇權數據（OT）會導致潛在地偏誤和無效率性的估計，反之，用股票和選擇權數據有含誤差項（S+O+E）會產生大致上比其他任何一種方法更有效性的無偏估計。這些結果被有限樣本模擬的研究證實了。因此，介於 S+O+E 和 ST 的估計誤差是實質性地導致顯著地不同風險管理結果。這些誤差大大影響了所採用方法的風險管理指標(如選擇權的 deltas 和 gammas 值)高達 80%。由於這 GARCH 選擇權模型是相對地限制及不能捕捉實證現像（參考 Engle 和 Mustafa（1992）），我們引進一個誤差項到這選擇權定價模型，借貸所需的呆滯到這個估計過程，由此產生了最大有效率性的無偏估計。也就是說，數據多是更好的，但是只有當數據是正確的被應用時。

關鍵詞: GARCH 選擇權模型，漸近行為，估計的有效率性及偏誤，風險管理

Abstract

This study derives asymptotic characteristics of GARCH(1,1) options price model estimators when using stock data only (ST), using option data only (OT), and using stock and options data with (S+O+E) or without an error term (S+O). The asymptotic variance in large sample theory shows that the OT method results in potentially biased and inefficient estimators, whereas S+O+E generates unbiased estimators which are substantially more efficient than either ST (S+O) or OT. These results are confirmed by finite sample simulation studies. Hence, the difference in estimation between S+O+E and ST is substantial and results in significantly different risk management consequences. These errors substantially impact risk management metrics as options deltas and gammas vary by as much as 80%, depending on the method used. Since the GARCH option models are relative restrictive and cannot capture the empirical phenomena (cf. Engle and Mustafa (1992)), we introduce an error term to the options pricing model, lending needed slack to the estimation process and resulting in unbiased estimates that are maximally efficient. That is, more data is better, but only if the data set is appropriately applied.

Keywords: GARCH option model, asymptotic behavior, estimator efficiency and bias, risk management

目 錄

List of Figures

List of Tables

Chapter 1

Introduction

A number of papers have sought to develop methods of appropriately and empirically accurately measuring stock price volatility, a factor that is critical in numerous fields of study including options pricing and risk management. It is a value that is necessary for calculation of hedge ratios and key risk metrics such as the options delta and gamma and, as a result, lies at the very core of traditional risk measurement, options pricing, and hedging strategies. Early work in this field focuses on application of stock price data (henceforth the ST data specification) while more recent efforts have sought to apply both stock and options data to the estimation problem under the assumption that the more data is applied, the more accurate are the resulting estimates. Empiricists have sought to do so in both the GARCH and stochastic volatility (SV) settings, generally applying these data without allowing for error in the options pricing model (henceforth the S+O method). Ultimately, GARCH models have been shown to provide better empirical fit and characteristics, making it an important class of models to consider. However, when it comes to estimating these models, application of the dual dataset becomes difficult owing to the restrictive nature of the specification, leading some to believe that the flexibility of SV models makes it a more desirable environment for estimation. Since there are consequences of the likelihood principle and Engle and Mustafa (1992), S+O method is found to be similar to ST method under GARCH models. That is, more data is not always better, and traditional methods leave estimation of this important class of models severely handicapped.

To resolve this issue, we develop additional two data specifications, a third data specification that applies options data only (OT) and a fourth data specification that applies stock and options data but includes an error term (S+O+E) such that options data need not match precisely the options pricing model.[1] We derive a quasi-maximum likelihood estimator (QMLE) for volatility under the GARCH(1,1) specification, both through analytical derivation of asymptotic behavior and numerical simulation, that the OT method generates inefficient, and more importantly, estimate bias that is economically and statistically significant. After relaxing this important modeling constraint, S+O+E generates asymptotically unbiased estimates that are the most efficient of the four data specifications. Most importantly, we then apply S&P500 stock and options daily data from January 2007 to the end of 2007 to generate commonly used risk and hedging metrics, i.e. the options delta and gamma. We find that those calculated using the S+O+E method are substantially different from those arrived at using stock data alone, indicating that the impact of including options data is economically significant and should be taken into account when determining hedging strategies.

Eraker (2004) offers several advantages using both stock and options data under SV model: A primary advantage is that risk premiums relating to volatility and jumps can

[1]More details about price errors see Engle and Mustafa (1992), Jacquier and Jarrow (2000), Eraker (2004) and Johannes, Polson, and Stroud (2009) and others.

be estimated. Secondly, the one-to-one correspondence of options to the conditional returns distribution allows parameters governing the shape of this distribution to potentially be very accurately estimated from option prices. For example, Eraker, Johannes, and Polson (2003) suggest that estimation from stock data alone requires fairly long samples to properly identify all parameters. Hopefully, the use of option prices can lead to very accurate estimates, even in short samples. Moreover, the use of option prices allows, and in fact requires, the estimation of the latent stochastic volatility process. Since volatility determines the time variation in relative option prices, there is also a strong potential for increased accuracy in the estimated volatility process. Finally, joint estimation also raises an interesting and important question: Are estimates of model parameters and volatility consistent across both markets? How to get the same results under GARCH models? This is the essential question to be addressed in this paper.

Other papers apply both stock and options data to estimate volatility, such as Chernov and Ghysels (2000), Jacquier and Jarrow (2000), Pan (2002), Jones (2003), Aït-Sahalia and Kimmel (2007), and Johannes, Polson, and Stroud (2009), each addressing the estimation issue under different assumptions. However, all of these papers do so under an SV or Black-Scholes (BS) rather than GARCH specification. Chernov and Ghysels (2000), Pan (2002), Jones (2003), and Aït-Sahalia and Kimmel (2007) apply an S+O model whereas Jacquier and Jarrow (2000), Eraker (2004) and Johannes, Polson, and Stroud (2009) apply an S+O+E model. In addition, Chernov and Ghysels (2000) and Pan (2002) use generalized method of moments (GMM) estimators, Jacquier and Jarrow (2000), Jones (2003), Eraker (2004), and Johannes, Polson, and Stroud (2009) use Bayesian inference estimators, whereas Aït-Sahalia and Kimmel (2007) uses maximum likelihood estimators

(MLE). Moreover, each of these papers focuses primarily on the computational aspects of their model, leaving open the important issues of the statistical properties and empirical implications of the models. Our paper explores these issues explicitly.

Importantly, Lehar, Scheicher, and Schittenkopf (2002) analyzes GARCH v/s SV models and finds that GARCH models dominate in terms of fit to observed prices. Given this attractive property, it is fruitful to develop a GARCH model that applies both data sets in an efficient, unbiased way: the task we explore here. Unfortunately, GARCH models that apply both stock and options data are scarce. Engle and Mustafa (1992) is among the earliest of these efforts. We follow this early work in that we also consider the role of error in the options pricing model but differ in that their work focuses on the application nonlinear least square (NLS) estimators for minimum of loss function toward the estimation of implied volatility, that is, this does so under an S+O specification. They find that the persistence of volatility shocks implied by options is found to be similar to that estimated from historical data on the index itself, that is, S+O method is similar to ST method. Christoffersen and Jacobs (2004) also uses NLS estimators and the S+O specification with both stock and options data under GARCH model. The focus of that study is on the accuracy of options pricing models and their ability to describe observed options prices. It does not address the estimation quality of the model nor does it seek to differentiate its data specifications from others. In contrast, our study is the first to derive the asymptotic characteristics of estimators then test these results using empirical data under the different data specifications.

Our theoretical construct builds upon the GARCH(1,1) setups of Heston and Nandi (2000), which propose a class of GARCH models that allow for a closed-form solution for the price of a European call option. We apply this model to the application of different data inclusion specifications and address asymptotic behavior using QMLE methods (cf. Lee and Hansen (1994) and Lumsdaine (1996)). For the S+O specification, we revise Aït-Sahalia and Kimmel's (2007) log-likelihood function to follow a GARCH(1,1), which has only one random source, and derive the asymptotic behavior of the QMLE. From this is a consequence of the likelihood principle, we find that the asymptotic behavior of the QMLE for S+O method is equal to that for ST method under GARCH models, that is, S+O method is similar to ST method. Thus, we don't display S+O method in this paper. Specifically, for the OT specification, we apply the log-likelihood function of Duffie, Pedersen, and Singleton (2003) to derive the asymptotic behavior of the QMLE. Our theoretical findings show that the OT method generates biased estimates and further partly results in higher estimation variance and mean squared error than applying stock data alone, the ST specification. Applying data and Monte Carlo simulations, we confirm these findings for all four variables that we seek to estimate.

These findings alone are perhaps not surprising. The GARCH specification is a restrictive one under which the application of the dual dataset can tend to obtain helpless the model. In contrast, the SV class of models introduces an error term into the volatility measure, thereby providing considerable slack in the model and allowing for the application of a more comprehensive dataset. The intuition behind our S+O+E method is the same. By allowing for additional slack, this time in the options pricing formula itself, we hope to provide the slack necessary for the dual dataset to generate unbiased, maximally efficient estimates for the GARCH class of models.

Specifically, the S+O+E specification assumes that $C_t = C_t^{HN} + e_t$. The error term e_t is assumed to be distributed $N(0,\eta^2)$ and is correlated with the error term of stock return.[2] Put plainly, we allow for options price data to err from the theoretical options price. Under this specification, we find that, for all four variables that we seek to estimate, the asymptotic mean squared error is lower than that of ST specification. That is, inclusion of an error term guarantees that the specification will dominate stock data only in large sample theory, where as applying stock and options data without an error term does not. Importantly, S+O+E also generates asymptotically unbiased estimates. These results are confirmed by our simulations in finite sample studies. Indeed, the OT method generates estimate bias, standard deviation, and mean squared error that are several times higher than those of S+O+E. ST, while generally not substantially biased, also generates estimates with standard deviations and mean squared errors that are several times higher than that of S+O+E. We conclude that the use of option prices can lead to very accurate estimates not only in short samples but also in long samples.

To test the implications of these differences in estimate quality, particularly in the risk management setting, we apply 12 months of stock and options data and show empirically that ST and S+O+E generate substantially different critical risk metrics. We calculate options delta and gamma using both the Black-Scholes and GARCH options pricing models. We find that estimates vary considerably depending on the data specification used. Delta estimates differ by as much as 80% which gamma estimates may differ by more than 60%. Although these differences are not systematically related to the

---

[2]In the information point of view, when good or bad news occur in financial market, these maybe affect the stock price and option price simultaneously. Then, these lead to the emergence of the correlation between stock error and price error.

moneyness of the option, they are nonetheless considerable. We conclude that the unbiased, more efficient estimates derived from our S+O+E method have a concrete and economically important impact, a notion managers would do well to keep in mind as they implement risk management practices.

Finally, we apply the Duan's (1995) options pricing formula for a robustness check. The results again show in this study, both through analytical derivation of asymptotic behavior and numerical simulation, that S+O+E method generates more efficient and unbiased estimates. These errors substantially impact risk management metrics as options deltas and gammas vary by as much as 90%. These results are consistent with our general findings.

The remainder of this paper is organized as follows: Chapter 2 introduces price error. Chapter 3 presents the main results. Chapter 4 proposes a robustness check. Chapter 5 concludes. All proofs of the results are relegated to Appendix.

Chapter 2

An Introduction of Pricing Error

Consider T observations of a contingent claim's market price, C_t, for $t \in \{1,\ldots,T\}$.We think of C_t as a limited liability derivative, like a call or put option. Formally, we can assume that there exists an unobservable equilibrium or arbitrage free price c_t for each observation. Then observed price C_t should be equal to the theoretical price c_t. There is a basic model $f(X_t,\theta)$ for the equilibrium price c_t. The model depends on vectors of observables X_t and parameters θ. We assume that the parameters are constant over the sample span. The model is an approximation, even though it was theoretically derived as being exact. There is an unobservable pricing error, e_t. A quote C_t may also sometimes depart from equilibrium. The error then has a second component ξ_t, which can be thought of as a market error. ξ_t and e_t are not identified without further assumptions. In this paper, we merge these two errors into one common pricing error e_t. Formally,

$$C_t = f(X_t,\theta) + e_t\,. \tag{2.1}$$

This implies a multiplicative error structure on the level, which guarantees the positivity of the call price for any error distribution.

The introduction of a non-zero error e_t is justified. First, simplifying assumptions on the structure of trading or the underlying stochastic process made to derive tractable models. They result in errors, possibly biased and non i.i.d. For example, Renault and Touzi (1996) and Heston (1993), show this within the context of stochastic volatility option pricing models. Renault (1997) shows that even a small non-synchroneity error in the recording of underlying and option prices can measurement can cause skewed Black-Scholes implied volatility smiles. Bakshi, Cao, and Chen (1997) show that adding jumps to a basic stochastic volatility process further improves pricing performance. Bossaerts and Hillion (1997) show that the assumption of continuous trading also leads to smiles while Platen and Schweizer (1994)'s hedging model causes time varying skewed smiles in the Black-Scholes model. In all of the above cases, the model errors are related to the inputs of the model. Second, in typical models, the rational agents are unaware of market or model error and know the parameters of the model. Such models could be biased in the 'larger system' consisting of expression (2.1).

Chapter 3

Parameter Estimation under GARCH Option Price models

3.1. Model setup and OT specification

First, we describe the general stock and option pricing models applied in this paper. Then, we derive QMLE and asymptotic results for the ST and OT specifications, noting the bias and estimator inefficiencies of the OT method. Numerical results confirm these characteristics.

3.1.1. GARCH(1,1) stock and option pricing models

We adopt the generalized setup used by Heston and Nandi (2000), which propose a class of GARCH models that allow for a closed-form solution for the price of a European call option, where the data-generating process for the stock price S is:

$$y_t = \ln S_t - \ln S_{t-1} = r + \lambda h_t + h_t^{1/2} z_t, \text{ under } P \text{ measure}, \tag{3.1}$$

$$h_t = \omega + \alpha\left(z_{t-1} - \gamma h_{t-1}^{1/2}\right)^2 + \beta h_{t-1}, \tag{3.2}$$

where r is the risk free rate and λ is the price of risk. The variance equation (3.2) is in fact a nonlinear asymmetric (NAGARCH) configuration (cf. Engle and Ng (1993)). The process remains stationary with finite mean and variance if $\alpha\gamma^2 + \beta < 1$. We may consider process (3.2) as running indefinitely or we may assume initial values y_0 and h_0, with the latter drawn from the stationary distribution applied by Bollerslev (1986), Nelson (1990), Bougerol and Picard (1992), and others. Let Ψ_t be the sigma-field generated by $\{y_t, y_{t-1}, \ldots\}$ and let $\theta_0 = (\omega_0, \alpha_0, \beta_0, \gamma_0)'$ represent the true parameter vector. Assume that $\theta_0 \in \Theta \subseteq \Re^4$ is in the interior of Θ, a compact, convex parameter space. Specifically, for any vector $(\omega, \alpha, \beta, \gamma) \in \Theta$, $0 < \omega_L \le \omega \le \omega_U$, $0 < \alpha_L \le \alpha \le \alpha_U$, $0 < \beta_L \le \beta \le \beta_U$, $\gamma_L \le \gamma \le \gamma_U$, and $\alpha_U(\lambda + \gamma_U + \frac{1}{2})^2 + \beta_U < 1$.[3] Assume also that $\{z_t\}_{t \in Z}$ is i.i.d., drawn from a symmetric, uni-modal density, bounded in a neighborhood of 0, with mean 0, and variance 1. In addition, assume that h_t is independent of $\{z_t, z_{t+1}, \ldots\}$.

The corresponding model under local risk neutralization reads

$$y_t = \ln S_t - \ln S_{t-1} = r - \frac{1}{2} h_t + h_t^{1/2} z_t^Q, \text{ under } Q \text{ measure} \tag{3.3}$$

$$h_t = \omega + \alpha \left(z_{t-1}^Q - \gamma_Q h_{t-1}^{1/2} \right)^2 + \beta h_{t-1}, \tag{3.4}$$

where $\gamma_Q = \gamma + \lambda + \frac{1}{2}$ and $z_t^Q = z_t + \left(\lambda + \frac{1}{2} \right) h_t^{1/2}$. Then, the GARCH option pricing formula is described as:

[3] Since $\alpha_U(\lambda + \gamma_U + 1/2)^2 + \beta_U < 1$ implies $\alpha\gamma_Q^2 + \beta < 1$, the process (3.4) under Q measure also remains stationary with finite mean and variance. These conditions easy to be arrived from our estimative parameters

$$C_t^{HN}(T,K,S_t,h_{t+1};\theta)=e^{-r(T-t)}E_t^Q\left[\max\left(S_T-K,0\right)\right]$$
$$=e^{-r(T-t)}f^*(1)\left(\frac{1}{2}+\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{K^{-i\phi}f^*(i\phi+1)}{i\phi f^*(1)}\right]d\phi\right) \tag{3.5}$$
$$-e^{-r(T-t)}K\left(\frac{1}{2}+\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{K^{-i\phi}f^*(i\phi)}{i\phi}\right]d\phi\right),$$

where $f^*(\varphi)=E_t^Q\left[e^{\varphi x_T}\right]=S_t^\varphi e^{A_t^\varphi+B_t^\varphi h_{t+1}(\theta)}$, $x_t=\ln(S_t)$, $A_T^\varphi=B_T^\varphi=0$, $A_t^\varphi=A_{t+1}^\varphi+\varphi r+B_{t+1}^\varphi\omega-\frac{\ln(1-2\alpha B_{t+1}^\varphi)}{2}$, $B_t^\varphi=\varphi(\gamma_Q-\frac{1}{2})-\frac{\gamma_Q^2}{2}+\beta B_{t+1}^\varphi+\frac{(\varphi-\gamma_Q)^2}{2(1-2\alpha B_{t+1}^\varphi)}$, $\mathrm{Re}(x)$ is the real part of x, T is the maturity date, and K is exercise price. From the derivative of real part doesn't exist, we consider a complex number $\varphi=m+\phi i$, $m=0,1$ to rewrite (3.5) such that the partial differentiation of $C_t^{HN}(\theta)$ exists. Then, $A_t^\varphi\equiv A_{1m,t}(\phi,\theta)+A_{2m,t}(\phi,\theta)i$ and $B_t^\varphi\equiv B_{1m,t}(\phi,\theta)+B_{2m,t}(\phi,\theta)i$, where

$$A_{1m,t}(\phi,\theta)=A_{1m,t+1}(\phi,\theta)+mr+\omega B_{1m,t+1}(\phi,\theta)-\frac{\ln\left([1-2\alpha B_{1m,t+1}(\phi,\theta)]^2+4\alpha^2B_{2m,t+1}^2(\phi,\theta)\right)}{4},$$

$$A_{2m,t}(\phi,\theta)=A_{2m,t+1}(\phi,\theta)+\phi r+\omega B_{2m,t+1}(\phi,\theta)-\frac{1}{2}\tan^{-1}\left(\frac{-2\alpha B_{2m,t+1}(\phi,\theta)}{1-2\alpha B_{1m,t+1}(\phi,\theta)}\right),$$

$$B_{1m,t}(\phi,\theta)=m(\gamma_Q-\frac{1}{2})-\frac{1}{2}\gamma_Q^2+\beta B_{1m,t+1}(\phi,\theta)+\frac{(1-2\alpha B_{1m,t+1}(\phi,\theta))[(m-\gamma_Q)^2-\phi^2]-4\alpha B_{2m,t+1}(\phi,\theta)(m-\gamma_Q)\phi}{2[(1-2\alpha B_{1m,t+1}(\phi,\theta))^2+4\alpha^2B_{2m,t+1}^2(\phi,\theta)]},$$

and $B_{2m,t}(\phi,\theta)=\phi(\gamma_Q-\frac{1}{2})+\beta B_{2m,t+1}(\phi,\theta)+\frac{\alpha B_{2m,t+1}(\phi,\theta)[(m-\gamma_Q)^2-\phi^2]+(m-\gamma_Q)\phi(1-2\alpha B_{1m,t+1}(\phi,\theta))}{(1-2\alpha B_{1m,t+1}(\phi,\theta))^2+4\alpha^2B_{2m,t+1}^2(\phi,\theta)}$. Hence, we rewrite (3.5) as

---

on S&P 500 index data. Hence, the parameter space Θ is enough large.

$$C_t^{HN}(\theta) = S_t\left[\frac{1}{2}+\frac{1}{\pi}\int_0^\infty \frac{e^{X_{11,t}(\phi,\theta)}\sin\left(X_{21,t}(\phi,\theta)\right)}{\phi}d\phi\right]$$
$$-e^{-r(T-t)}K\left[\frac{1}{2}+\frac{1}{\pi}\int_0^\infty \frac{e^{X_{10,t}(\phi,\theta)}\sin\left(X_{20,t}(\phi,\theta)\right)}{\phi}d\phi\right], \tag{3.6}$$

where $X_{1m,t}(\phi,\theta) = -r(T-t)m + A_{1m,t}(\phi,\theta) + B_{1m,t}(\phi,\theta)h_{t+1}(\theta)$ and $X_{2m,t}(\phi,\theta) = \phi\ln(S_t / K) + A_{2m,t}(\phi,\theta) + B_{2m,t}(\phi,\theta)h_{t+1}(\theta)$, $m = 0,1$.

3.1.2. QMLEs for ST and asymptotic results

We now turn our attention to estimating the parameters in the model. The base case ST uses only stock data. Specifically, h_t is the conditional variance of y_t with respect to Ψ_{t-1}. The estimation model utilizes (3.1) and (3.2), applying estimated parameter values $(\omega,\alpha,\beta,\gamma) = (\theta_1,\theta_2,\theta_3,\theta_4)$. The error terms z_t are computed as $z_0 = \frac{y_0 - r - \lambda h_0}{h_0^{1/2}}$, $z_1 = \frac{y_1 - r - \lambda h_1}{h_1^{1/2}}$, …, where $\{y_t,\ t = 0,\ldots,T\}$ are observed data. The process h_t is not observed but is constructed recursively using estimated parameter values, z_0, and an appropriate startup value, h_0, to be discussed in detail later.

QMLE is obtained by maximizing, conditional on h_0, as follows:

$$L_T^{ST}(y_0,\ldots,y_T,h_0;\theta) = L_T^{ST}(\theta) = \frac{-1}{2T}\sum_{t=1}^{T}\left(\ln\left(h_t(\theta)\right) + \frac{\left(y_t - r - \lambda h_t(\theta)\right)^2}{h_t(\theta)}\right). \tag{3.7}$$

That is, $\hat{\theta}_T^{ST} = \underset{\theta\in\Theta}{\arg\max}\, L_T^{ST}(\theta)$. This estimator is consistent as $\hat{\theta}_T^{ST} \xrightarrow{p} \theta_0$ and is asymptotically Normal as

$$H_{ST0}^{-1/2} F_{ST0} T^{1/2} (\hat{\theta}_T^{ST} - \theta_0) \overset{A}{\sim} N(0, I_4), \tag{3.8}$$

where $F_{ST0} = -E\left(\frac{\partial^2 L_T^{ST}(\theta_0)}{\partial\theta\partial\theta'}\right)$, $H_{ST0} = E\left(T \frac{\partial L_T^{ST}(\theta_0)}{\partial\theta} \frac{\partial L_T^{ST}(\theta_0)}{\partial\theta'}\right)$, and $I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

A full proof appears in Section 3.5 as Theorem 3.1.

In the interest of computational simplicity, assume that z_t is Normal so that $F_{ST0} = H_{ST0}$, though our general intuition remains the same under the more relaxed aforementioned specification for z_t. The asymptotic covariance matrix V_{ST} and asymptotic mean square errors MSE_{ST} are:

$$MSE_{ST}(\theta_0) = V_{ST}(\theta_0) = F_{ST0}^{-1} H_{ST0} F_{ST0}^{-1} = \frac{1}{F_{ST0}}, \tag{3.9}$$

where $F_{ST0} = H_{ST0} = E\left[\left(\frac{1}{2h_t^2(\theta_0)} + \frac{\lambda^2}{h_t(\theta_0)}\right)\frac{\partial h_t(\theta_0)}{\partial\theta}\frac{\partial h_t(\theta_0)}{\partial\theta'}\right]$.

3.1.3. QMLEs for OT and asymptotic results

For OT, we simply have $C_t = C_t^{HN}$, $t = 1, \ldots, T$, since the pricing formula is assumed to match the observed data exactly. Duffie, Pedersen, and Singleton (2003) provide a treatment for the log-likelihood function when only options data is applied in this fashion. Let S_t be an unobservable stock price. Expressing the stock and options price vector as a function of the state variable vector, we have: $C_t = f(S_t; \theta)$ for a differentiable function *f* that is easily computed. At a given parameter vector θ, we may now express the state variable as a function of observed asset prices as follows: $S_t(\theta) = f^{-1}(C_t;\theta)$ assuming invertibility (which is not an issue in our application). Letting $C = (C_1, \ldots, C_T)$ denote the sequence of observed vector of reference option prices, standard change-of-variable arguments lead to the likelihood

$$P(C;\theta) = \prod_{t=1}^{T} P\left(S_t(\theta) \mid S_{t-1}(\theta);\theta\right) \frac{1}{\left|\det Df\left(S_t(\theta);\theta\right)\right|}, \tag{3.10}$$

where

$$\begin{aligned}
\det Df\left(S_t(\theta);\theta\right) &= \tfrac{\partial C_t^{HN}(S_t(\theta);\theta)}{\partial S_t(\theta)} \\
&= \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \frac{e^{X_{11,t}(\phi,\theta)} \sin\left(X_{31,t}(\phi,\theta)\right)}{\phi} d\phi + \frac{1}{\pi}\int_0^{\infty} e^{X_{11,t}(\phi,\theta)} \cos\left(X_{31,t}(\phi,\theta)\right) d\phi \\
&\quad - e^{-r(T-t)} \frac{K}{S_t(\theta)} \frac{1}{\pi}\int_0^{\infty} e^{X_{10,t}(\phi,\theta)} \cos\left(X_{30,t}(\phi,\theta)\right) d\phi
\end{aligned}$$

and

$$X_{3m,t}(\phi,\theta) = \phi \ln\left(S_t(\theta)/K\right) + A_{2m,t}(\phi,\theta) + B_{2m,t}(\phi,\theta) h_{t+1}(\theta),\ m = 0,1.$$

Then, the log-likelihood function for discrete data of the asset price vector C sampled at dates $0 \le t \le T$ has the form

$$L_T^{OT}(\theta) = -\frac{1}{T}\sum_{t=1}^{T}\left[\ln\left(J_t(\theta)\right) + \frac{1}{2}\left(\ln\left(h_t(\theta)\right) + \frac{Y_t^2(\theta)}{h_t(\theta)}\right)\right]. \tag{3.11}$$

where $J_t(\theta) = \left|\det Df(S_t(\theta);\theta)\right|$ and $Y_t(\theta) = \ln S_t(\theta) - \ln S_{t-1}(\theta) - r - \lambda h_t(\theta)$. And, the QMLE for $\hat{\theta}_T^{OT} = \arg\max_{\theta\in\Theta} L_T^{OT}(\theta)$. Note, then, that this estimator is asymptotically biased since: $\hat{\theta}_T^{OT} - \theta_1 \xrightarrow{p} 0$, where $\theta_1 \equiv \theta_0 + \left(\frac{\partial^2 L_T^{OT}(\theta_0)}{\partial\theta\partial\theta'}\right)^{-1} \frac{1}{T}\sum_{t=1}^{T}\frac{1}{J_t(\theta_0)}\frac{\partial J_t(\theta_0)}{\partial\theta}$. Investigating the bias in particular, we have that $Bias_{OT}(\theta_0) = \left(\frac{\partial^2 L_T^{OT}(\theta_0)}{\partial\theta\partial\theta'} \right)^{-1} \frac{1}{T} \sum_{t=1}^{T} \frac{1}{J_t(\theta_0)} \frac{\partial J_t(\theta_0)}{\partial\theta}$. Full proofs can be found in Section 3.5 as Theorem 3.2.

Theorem 3.2 also shows that the estimator is asymptotically Normally distributed as $H_{OT0}^{-1/2} F_{OT0} T^{1/2} (\hat{\theta}_T^{OT} - \theta_1) \overset{A}{\sim} N(0, I_4)$, where $F_{OT0} = -E\left(\frac{\partial^2 L_T^{OT}(\theta_0)}{\partial\theta\partial\theta'} \right)$ and $H_{OT0} = Var\left(T^{1/2} \frac{\partial L_T^{OT}(\theta_0)}{\partial\theta} \right)$. Again, assume that z_t is Normal. The asymptotic covariance matrix V_{OT} and asymptotic mean square errors MSE_{OT} for the OT case are:

$$V_{OT}(\theta_0) = \frac{1}{F_{OT0} H_{OT0}^{-1} F_{OT0}} \tag{3.12}$$

and

$$MSE_{OT}(\theta_0) = V_{OT}(\theta_0) + Bias_{OT}^2(\theta_0), \tag{3.13}$$

where $H_{OT0}(\theta_0) = E\left[\frac{1}{2h_t^2(\theta_0)} \frac{\partial h_t(\theta_0)}{\partial\theta} \frac{\partial h_t(\theta_0)}{\partial\theta'} \right] + E\left[\frac{1}{h_t(\theta_0)} \frac{\partial Y_t(\theta_0)}{\partial\theta} \frac{\partial Y_t(\theta_0)}{\partial\theta'} \right]$ and $F_{OT0}(\theta_0) = E\left[\frac{1}{J_t(\theta_0)} \frac{\partial^2 J_t(\theta_0)}{\partial\theta\partial\theta'} \right] - E\left[\frac{1}{J_t^2(\theta_0)} \frac{\partial J_t(\theta_0)}{\partial\theta} \frac{\partial J_t(\theta_0)}{\partial\theta'} \right] + H_{OT0}(\theta_0)$. Then, the difference of asymptotic mean square errors between OT and ST $dMSE_{OT}$ is:

$$dMSE_{OT}(\theta_0) = MSE_{OT}(\theta_0) - MSE_{ST}(\theta_0). \tag{3.14}$$

Using these results, which follow from Lemmas 3.1, 3.2, 3.5, and 3.6 in Section 3.5, we can compare the magnitude of mean square errors in large sample theory, a lower asymptotic mean square errors indicating better estimation. Namely, if $dMSE_{OT}(\theta_0) < 0$, then $MSE_{OT} < MSE_{ST}$ and using options data only specification is more efficient than using stock only. Since the covariance matrix V is a 4×4 matrix, we estimate each of the four parameters separately, holding the other three constants.

When α_0, β_0, and γ_0 are known and ω is unknown, we have the asymptotic bias is $Bias_{OT}(\omega_0 \mid \alpha_0,\beta_0,\gamma_0) = \left(\frac{\partial^2 L_T^{OT}(\omega_0)}{\partial\omega^2}\right)^{-1} \frac{1}{T}\sum_{t=1}^{T}\frac{1}{J_t(\theta_0)}\frac{\partial J_t(\omega_0)}{\partial\omega}$. Bias, then, is non-zero in magnitude. Investigating estimate mean square errors, we find that:

$$dMSE_{OT}(\omega_0 \mid \alpha_0,\beta_0,\gamma_0) = MSE_{OT}(\omega_0) - MSE_{ST}(\omega_0). \tag{3.15}$$

Note that $dMSE_{OT}$ may be positive or negative, where a positive results means that results are less efficient than using stock data alone. Since the $dMSE_{OT}$ depends on true parameters, we don't compare these values. Thus, we will calculate these values by numerical simulation in Section 3.1.4. As illustrated later, $dMSE_{OT}$ is in fact sometimes positive.

Similarly, when ω_0, β_0, and γ_0 are known and α is unknown, we have $Bias_{OT}(\alpha_0 \mid \omega_0,\beta_0,\gamma_0) = \left(\frac{\partial^2 L_T^{OT}(\alpha_0)}{\partial\alpha^2}\right)^{-1} \frac{1}{T}\sum_{t=1}^{T}\frac{1}{J_t(\theta_0)}\frac{\partial J_t(\alpha_0)}{\partial\alpha}$ which is again non-zero in magnitude and

$$dMSE_{OT}(\alpha_0 \mid \omega_0,\beta_0,\gamma_0) = MSE_{OT}(\alpha_0) - MSE_{ST}(\alpha_0). \tag{3.16}$$

As demonstrated later, $dMSE_{OT}$ is sometime positive definite and the estimator is sometimes less efficient than that which is found using the ST method.

Similarly, when ω_0, α_0, and γ_0 are known and β is unknown, we have $Bias_{OT}(\beta_0 \mid \omega_0,\alpha_0,\gamma_0) = \left(\frac{\partial^2 L_T^{OT}(\beta_0)}{\partial\beta^2}\right)^{-1} \frac{1}{T}\sum_{t=1}^{T}\frac{1}{J_t(\theta_0)}\frac{\partial J_t(\beta_0)}{\partial\beta}$ which is again non-zero in magnitude and

$$dMSE_{OT}(\beta_0 \mid \omega_0, \alpha_0, \gamma_0) = MSE_{OT}(\beta_0) - MSE_{ST}(\beta_0). \tag{3.17}$$

As demonstrated later, we find that $dMSE_{OT}$ is sometimes positive definite.

Finally, when ω_0, α_0, and β_0 are known and γ is unknown, we find similar to the previous case that $Bias_{OT}(\gamma_0 \mid \omega_0, \alpha_0, \beta_0) = \left(\frac{\partial^2 L_T^{OT}(\gamma_0)}{\partial \gamma^2} \right)^{-1} \frac{1}{T} \sum_{t=1}^{T} \frac{1}{J_t(\theta_0)} \frac{\partial J_t(\gamma_0)}{\partial \gamma}$, a non-zero entity and

$$dMSE_{OT}(\gamma_0 \mid \omega_0, \alpha_0, \beta_0) = MSE_{OT}(\gamma_0) - MSE_{ST}(\gamma_0). \tag{3.18}$$

Again, as for the case where β is unknown, we show that $dMSE_{OT}$ is sometimes positive.

3.1.4. Numerical computation for asymptotic bias and mean square errors

We now generate numerical results to test and illustrate these asymptotic findings. We presume that parameter true values are (λ, ω_0, α_0, β_0, γ_0) = (0.1746, 6.792×10^{-9}, 6.546×10^{-8}, 0.9914, 351.945), and the risk-free rate is fixed at 5%. These parameters are estimated using S&P 500 daily index data from January 1996 to the end of 2007. We use these parameters to run our tests.

First, we investigate and calculate analytically the aforementioned estimate bias. Graphs in Figure 1 show the absolute value of bias divided by true value in the area surrounding true parameter values. Since the orders of magnitude for the four parameters are quite different, we graph the bias of ω on the left, that of α, that of β, and that of γ on the right. True parameters are circled in each graph. Then, in each panel, one variable is varied while the other three are treated as known. Specifically, in Panel A, ω is varied, in B α, in C β, and in D γ.

[Insert Figure 1 here]

Note that, in all graphs, bias is decidedly non-zero and non-trivial for all four parameters. Though not shown here, the absolute value of bias is positive for all four parameters for the entire span of possible parameter values.[4] In Panel A, the bias for ω is sometimes large and sometimes small with ω locally in the region around the true parameter values but that for ω in the true parameter values is non-zero in magnitude. There are the same results for those for α, β, and γ. The bias for ω is always higher than that of others. The order of these values is the bias for ω, that for α, that for γ, and that of β. Corresponding graphs Panels B, C, and D are similar to each other in shape, though their x-axes differ. All in all, using OT, bias is non-zero for each variable estimated, regardless of the true parameter values implemented. In contrast, neither ST nor S+O+E generate asymptotic bias in any variable.

Shifting our attention to the efficiency of the estimator, graphs of $dMSE_{OT}$ are shown in Figure 2. Remember that, the more positive this value, the more efficient the estimator.

[Insert Figure 2 here]

Once again, in Panel A, ω is varied, in B α , in C β, and in D γ. Looking at Panel A, $dMSE_{OT}$ for ω is almost negative and sometimes positive for all four parameters in the area surrounding true parameters but that for ω in the true parameter values is always negative. There are the same results for those for α, β, and γ. $dMSE_{OT}$ for α is always lower than that of others. The order of these values is $dMSE_{OT}$ for γ, that for β, that for ω, and that of α.

[4] Each the parameter for ω (α, β, γ) ranges from 0 to 1 such that $\alpha(\lambda+\gamma+0.5)^2+\beta<1$.

Graphs in Panels B, C, and D are again similar in shape. In three panels, $dMSE_{OT}$ is sometimes positive for all four parameters.

In summary, $dMSE_{OT}$ is sometimes positive for all parameters in the area surrounding true parameters. As such OT does not generally produce efficient estimators. As aforementioned, it furthermore generates significant bias. We conclude that OT is not an optimal estimation method given the restrictive nature of GARCH models. As a result, we seek to develop a method that will allow for asymptotic unbias and efficient estimation of this important class of models.

3.2. The S+O+E specification

We now turn our attention to a new specification that takes both stock and options data into account, but which allows for an error term in the options pricing formula. Then, we derive QMLE and asymptotic results for the S+O+E specification. Numerical results confirm these characteristics.

3.2.1. QMLEs and asymptotic results

For this method, we allow that $C_t = C_t^{HN}(\theta) + e_t$ where $e_t = \eta u_t$ and $t = 1, \ldots, T$. Assume that $u_t \overset{i.i.d.}{\sim} N(0,1)$ and $\eta > 0$. For the purpose of calculating the QMLE, let us assume z_t and u_t, with $correlation(z_t, u_t) = \rho$ where $-1 < \rho < 1$, have a bi-Normally distribution, that is, $\begin{pmatrix} z_t \\ u_t \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right)$. [5] Let $G_t = [S_t, C_t]'$ be a vector of observable stock and option prices, respectively. Then, the joint density is as follows:

$$P(G;\theta) = P(S,C;\theta) = P(C \mid S;\theta) P(S;\theta) = \prod_{t=1}^{T} P(C_t \mid S_t;\theta) P(S_t \mid S_{t-1};\theta)$$

$$= \prod_{t=1}^{T} \frac{1}{2\pi\sqrt{1-\rho^2}\sqrt{h_t(\theta)}\eta} \exp\left[-\frac{1}{2(1-\rho^2)} \left(\frac{(y_t - r - \lambda h_t(\theta))^2}{h_t(\theta)} - 2\rho \frac{y_t - r - \lambda h_t(\theta)}{\sqrt{h_t(\theta)}} \frac{C_t - C_t^{HN}(\theta)}{\eta} + \frac{\left(C_t - C_t^{HN}(\theta)\right)^2}{\eta^2} \right) \right]. \quad (3.19)$$

The log-likelihood function for discrete data on the asset price vector G_t sampled at dates $0 \le t \le T$ has the form:

$$L_T^{S+O+E}(\theta) = \ln\left(P(G;\theta)\right)$$

$$= \frac{-1}{2T} \sum_{t=1}^{T} \left[2\ln\left(2\pi\eta\sqrt{1-\rho^2}\right) + \ln(h_t(\theta)) + \frac{1}{(1-\rho^2)} \left(\frac{(y_t - r - \lambda h_t(\theta))^2}{h_t(\theta)} - 2\rho \frac{y_t - r - \lambda h_t(\theta)}{\sqrt{h_t(\theta)}} \frac{C_t - C_t^{HN}(\theta)}{\eta} + \frac{\left(C_t - C_t^{HN}(\theta)\right)^2}{\eta^2} \right) \right].$$

(3.20)

And, the QMLE for $\hat{\theta}_T^{S+O+E} = \arg\max_{\theta \in \Theta} L_T^{S+O+E}(\theta)$.

Unlike the OT case, this estimator is consistent as $\hat{\theta}_T^{S+O+E} \xrightarrow{p} \theta_0$ and is asymptotically Normally distributed as $H_{S+O+E0}^{-1/2} F_{S+O+E0} T^{1/2} (\hat{\theta}_T^{S+O+E} - \theta_0) \overset{A}{\sim} N(0, I_4)$, where $F_{S+O+E0} = -E\left(\frac{\partial^2 L_T^{S+O+E}(\theta_0)}{\partial\theta\partial\theta'} \right)$ and $H_{S+O+E0} = E\left(T \frac{\partial L_T^{S+O+E}(\theta_0)}{\partial\theta} \frac{\partial L_T^{S+O+E}(\theta_0)}{\partial\theta'} \right)$. A full proof appears as Theorem 3.3 in Section 3.5. Again, assume that z_t is Normal so that

[5]This simple idea likes Eraker (2004) to joint stock error and price error but differ in that he assumes that the relation between stock error and price error is zero and there is the relation between price error at time *t* and price error at time *t-1*.

$F_{S+O+E0} = H_{S+O+E0}$. The asymptotic covariance matrix V_{S+O+E} and asymptotic mean square error MSE_{S+O+E} for the S+O+E case are:

$$\begin{aligned} MSE_{S+O+E}(\theta_0) &= V_{S+O+E}(\theta_0) \\ &= F_{S+O+E0}^{-1} H_{S+O+E0} F_{S+O+E0}^{-1} = \frac{1}{F_{S+O+E0}} = \frac{1}{F_{ST0} + M_{S+O+E}(\theta_0)}, \end{aligned} \qquad (3.21)$$

where $M_{S+O+E}(\theta_0) = F_{S+O+E0} - F_{ST0}$ and

$$M_{S+O+E}(\theta_0) = \frac{\rho^2}{1-\rho^2} E\left[\left(\frac{1}{4h_t^2(\theta_0)} + \frac{\lambda^2}{h_t(\theta_0)}\right)\frac{\partial h_t(\theta_0)}{\partial \theta}\frac{\partial h_t(\theta_0)}{\partial \theta'}\right] - \frac{2\lambda\rho}{\eta(1-\rho^2)} E\left[\frac{1}{h_t^{1/2}(\theta_0)}\frac{\partial C_t^{HN}(\theta_0)}{\partial \theta}\frac{\partial h_t(\theta_0)}{\partial \theta'}\right] + \frac{1}{\eta^2(1-\rho^2)} E\left[\frac{\partial C_t^{HN}(\theta_0)}{\partial \theta}\frac{\partial C_t^{HN}(\theta_0)}{\partial \theta'}\right].$$

These results follow from Lemmas 3.1, 3.2, 3.13, and 3.14 in Section 3.5. As before, we can compare the magnitude of asymptotic mean square errors, again a lower asymptotic mean square errors indicating better estimation. Here, if $M_{S+O+E} > 0$, then $MSE_{S+O+E} < MSE_{ST}$. Namely, we now investigate M_{S+O+E} where the more positive, the more efficient the estimator.

When α_0, β_0, and γ_0 are known and ω is unknown:

$$\begin{aligned} M_{S+O+E}(\omega_0 \mid \alpha_0, \beta_0, \gamma_0) = &\frac{\rho^2}{1-\rho^2} E\left[\left(\frac{1}{4h_t^2(\theta_0)} + \frac{\lambda^2}{h_t(\theta_0)}\right)\left(\frac{\partial h_t(\omega_0)}{\partial \omega}\right)^2\right] \\ &- \frac{2\lambda\rho}{\eta(1-\rho^2)} E\left[\frac{1}{h_t^{1/2}(\theta_0)}\frac{\partial C_t^{HN}(\omega_0)}{\partial \omega}\frac{\partial h_t(\omega_0)}{\partial \omega}\right] + \frac{1}{\eta^2(1-\rho^2)} E\left[\left(\frac{\partial C_t^{HN}(\omega_0)}{\partial \omega}\right)^2\right]. \end{aligned} \qquad (3.22)$$

Note that M_{S+O+E} may be positive or negative, where a positive result means that results are more efficient than using stock data alone. In $\rho = 0$ case, we easy to see that M_{S+O+E} is positive definite from (3.22), indicating that S+O+E generates more efficient estimates than ST. This method makes it the most desirable data specification of the two. In $\rho \neq 0$ case, we don't compare these values since the M_{S+O+E} depend on true parameters. Thus, we will calculate these values by numerical simulation in Section 3.2.2. As illustrated later, M_{S+O+E} is in fact generally positive. When ω_0, β_0, and γ_0 are known and α is not:

$$M_{S+O+E}(\alpha_0 \mid \omega_0, \beta_0, \gamma_0) = \frac{\rho^2}{1-\rho^2} E\left[\left(\frac{1}{4h_t^2(\theta_0)} + \frac{\lambda^2}{h_t(\theta_0)}\right)\left(\frac{\partial h_t(\alpha_0)}{\partial \alpha}\right)^2\right] - \frac{2\lambda\rho}{\eta(1-\rho^2)} E\left[\frac{1}{h_t^{1/2}(\theta_0)} \frac{\partial C_t^{HN}(\alpha_0)}{\partial \alpha} \frac{\partial h_t(\alpha_0)}{\partial \alpha}\right] + \frac{1}{\eta^2(1-\rho^2)} E\left[\left(\frac{\partial C_t^{HN}(\alpha_0)}{\partial \alpha}\right)^2\right]. \tag{3.23}$$

In $\rho = 0$ case, M_{S+O+E} is positive definite, and in $\rho \neq 0$ case, as demonstrated later, M_{S+O+E} is positive definite. Similarly, when ω_0, α_0, and γ_0 are known and β is not:

$$M_{S+O+E}(\beta_0 \mid \alpha_0, \omega_0, \gamma_0) = \frac{\rho^2}{1-\rho^2} E\left[\left(\frac{1}{4h_t^2(\theta_0)} + \frac{\lambda^2}{h_t(\theta_0)}\right)\left(\frac{\partial h_t(\beta_0)}{\partial \beta}\right)^2\right] - \frac{2\lambda\rho}{\eta(1-\rho^2)} E\left[\frac{1}{h_t^{1/2}(\theta_0)} \frac{\partial C_t^{HN}(\beta_0)}{\partial \beta} \frac{\partial h_t(\beta_0)}{\partial \beta}\right] + \frac{1}{\eta^2(1-\rho^2)} E\left[\left(\frac{\partial C_t^{HN}(\beta_0)}{\partial \beta}\right)^2\right]. \tag{3.24}$$

Again, as for the case where α is unknown, we show that M_{S+O+E} is always positive. Finally, when ω_0, α_0, and β_0 are known and γ is not:

$$M_{S+O+E}(\gamma_0 \mid \alpha_0, \beta_0, \omega_0) = \frac{\rho^2}{1-\rho^2} E\left[\left(\frac{1}{4h_t^2(\theta_0)} + \frac{\lambda^2}{h_t(\theta_0)}\right)\left(\frac{\partial h_t(\gamma_0)}{\partial \gamma}\right)^2\right] - \frac{2\lambda\rho}{\eta(1-\rho^2)} E\left[\frac{1}{h_t^{1/2}(\theta_0)} \frac{\partial C_t^{HN}(\gamma_0)}{\partial \gamma} \frac{\partial h_t(\gamma_0)}{\partial \gamma}\right] + \frac{1}{\eta^2(1-\rho^2)} E\left[\left(\frac{\partial C_t^{HN}(\gamma_0)}{\partial \gamma}\right)^2\right]. \tag{3.25}$$

Again, as for the case where β is unknown, we show that M_{S+O+E} is always positive.

3.2.2. Numerical computation for asymptotic mean square errors

We now generate numerical results to test and illustrate these asymptotic findings. As before, we use these parameters (λ, ω_0, α_0, β_0, γ_0) = (0.1746, 6.792×10^{-9}, 6.546×10^{-8}, 0.9914, 351.945) and the risk-free rate is fixed at 5% to run our tests.

First, we investigate and calculate analytically the efficiency of the estimator, graphs of M_{S+O+E}, in varied ρ. Remember that, the more positive this value, the more efficient the estimator. Graphs in Figure 3 show that the value of M_{S+O+E} in the true parameter values and ρ from -0.9 to 0.9. Then, in each panel, one variable is unknown while the other three are treated as known. Specifically, in Panel A, ω is varied, in B α, in C β, and in D γ.

[Insert Figure 3 here]

Looking at all graphs, M_{S+O+E} is decidedly non-zero and non-trivial for all four parameters. And, M_{S+O+E} is positive for all four parameters. Specifically, M_{S+O+E} is always minimum in $\rho = 0$ and increases as the absolute value of ρ increases. In all cases, S+O+E generates more efficient estimates than ST.

Graphs of M_{S+O+E} are shown in Figure 4, in the area surrounding true parameter values. We only consider $\rho = 0$ case since M_{S+O+E} is minimum in this case. Since the orders of magnitude for the four parameters are quite different, we graph the M_{S+O+E} of ω, α, β, and γ on the left to the right. True parameters are circled in each graph. Then, in each panel, one variable is varied while the other three are treated as known. Specifically, in Panel A, ω is varied, in B α, in C β, and in D γ.

[Insert Figure 4 here]

Note that, in all graphs, M_{S+O+E} is decidedly non-zero and non-trivial for all four parameters. Thought not shown here, M_{S+O+E} is positive for all four parameters for the entire span of possible parameter values.[6] In Panel A, M_{S+O+E} is always positive for all four

[6] Each parameter for ω (α, β, γ) ranges from 0 to 1 such that $\alpha(\lambda+\gamma+0.5)^2+\beta<1$.

parameters in the area surrounding true parameters. While M_{S+O+E} for ω decreases with ω locally in the region around the true parameter values, those for α, β, and γ are increasing. In Panel B, M_{S+O+E} for ω, α, β, and γ increase with α locally in the region around the true parameter values. Corresponding graphs Panels C and D are similar to each other in shape, though their x-axes differ. However, M_{S+O+E} is always positive in each case.

In summary, M_{S+O+E} is always positive for all parameters in the area surrounding true parameters. We conclude that S+O+E generates more efficient estimates than ST in all cases. This method also generates asymptotically unbiased estimates, making it the most desirable data specification of the two. We also conclude that the use of option prices can lead to very accurate estimates, even in long samples.

3.2.3. Numerical findings and direct comparisons in finite sample studies

We generate parameter estimates using a Monte Carlo method, comparing the bias and variance characteristics of the three data specifications in finite sample studies. Specifically, we simulate 30 days of stock and/or option prices and then run 1,000 iterations over each period.[7] We again presume that parameter true values are (λ, ω_0, α_0, β_0, γ_0) = (0.1746, 6.792×10^{-9}, 6.546×10^{-8}, 0.9914, 351.945). For the S+O+E case, we additionally assume that $\eta = 1$ and that $\rho = 0$. As a robustness check, we re-run these tests for a variety of calibrations of ρ and η and find no qualitative differences.

We estimate our four parameters for out-of-the-money ($S_0/K = 0.9$), at-the-money ($S_0/K = 1.0$), and in-the-money ($S_0/K = 1.1$) cases. Results for 30 days are presented in

[7] We have also re-run all tests using 90 and 360 simulated days with qualitatively identical results.

Table 1 where we report the absolute value of estimate bias (estimate less true value), standard deviation of the estimate (SD), and mean squared errors (MSE).

[Insert Table 1 here]

When ω is unknown, we find that S+O+E arrives at estimates within 4.355×10^{-9} of the true value. In contrast, estimates using ST and OT present biases on the order of roughly 1.1 to 4.4 times higher, respectiviely. Note that OT has the stronger bias regardless of the moneyness of the options. When options are in-the-money, standard deviations are lowest for S+O+E, with ST and OT again about 1.3 times and 1.1 times higher, respectively. With regard to MSE, results are even more staggering with S+O+E exhibiting the lowest values and ST and OT generating errors that are about 9 times and 10 times higher. Note that OT seems to perform particularly poorly when options are in-the-money.

When estimating α, we similarly find that estimation bias is lower for S+O+E than for ST and OT, with the latter of these again having by the largest bias. Standard deviations are also lowest for S+O+E, with ST and OT again about 5 times and 3 times higher for most variables, respectively. MSE exhibits the same behavior as before, with S+O+E exhibiting by the lowest values, with magnitudes of difference similar to before. Note that OT presents particularly poor results with options out-of-the-money.

When estimating β, we similarly find that estimation bias is far lower for S+O+E than for ST and OT, with the latter of these again having by far the largest bias. Standard deviation and MSE exhibit the same behavior as before, with S+O+E exhibiting by far the lowest values, with magnitudes of difference similar to before. Once again, OT presents particularly poor results with options out-of-the-money.

Estimation results for γ are consistent with those of the other three parameters. S+O+E exhibits the smallest bias, the lowest standard deviation, and the lowest MSE of the three methods. Again, OT exhibits by far the worst performance along all four metrics and again particularly poor when options are out-of-the-money. The prevalence of biased estimates is striking for OT and is a particular strength for S+O+E.

In summary, we conclude that the use of option prices can lead to very accurate estimates, even in short samples. This result is consistent with that of Eraker (2004).

3.3. Risk management implications

In this section we document that errors and bias in estimation may have substantial repercussions as relates to risk management benchmarks and practices. To illustrate, we obtain daily stock and options data from the Center for Research in Security Prices (CRSP) and the Option Metrics for the period from January 2007 to the end of 2007. For stock prices (S_t), we use the S&P 500 index, and for options data (C_t), we use the price of a short-maturity at-the-money call options where the price is measured as the midpoint of the last reported bid-ask spread. We assume that $h_0 = (\omega + \alpha)/(1 - \alpha\gamma^2 - \beta)$, and for ease of interpretation, let the risk-free rate equal 0%. Applying the 12 months of stock and options data, we find key parameters to be (λ, ω_0, α_0, β_0, γ_0) = (0.1821, 6.847×10^{-9}, 6.669×10^{-8}, 0.9911, 342) for ST and (λ, ω_0, α_0, β_0, γ_0) = (0.1821, 6.029×10^{-9}, 7.166×10^{-8}, 0.9879, 402) for S+O+E. We omit the OT specification as it has been demonstrated that this method sometimes produces inefficient and, more importantly, biased estimates. As demonstrated in the following discussion, while these parameters may not appear to differ greatly, the resulting risk management implications are quite significant.

We then use these parameter estimates to calculate options deltas and gammas, measuring options stock price sensitivity and convexity, respectively. We calculate deltas and gammas for both the Black-Scholes and GARCH options pricing models, so that we have a total of four risk management metrics. For the former, we have that $\Delta^{BS} = \Phi(d_1)$ and $\Gamma^{BS} = \dfrac{\varphi(d_1)}{S_0\sigma\sqrt{T}}$ where $d_1 = \dfrac{\ln(S_0/K)+(r+\sigma^2/2)T}{\sigma\sqrt{T}}$ and $\sigma^2 = \dfrac{\omega_0+\alpha_0}{1-\alpha_0\gamma_0^2-\beta_0}$ (cf. Duan (1995)). For the latter, we find that $\Delta^{GARCH} = \mathrm{e}^{-rT}\, E_0^Q[\dfrac{S_T}{S_0} 1_{\{S_T \geq K\}}] \cong \dfrac{1}{N}\mathrm{e}^{-rT}\sum_{i=1}^{N}\dfrac{S_{T,i}}{S_0} 1_{\{S_{T,i} \geq K\}}$ and $\Gamma^{GARCH} = \dfrac{\partial \Delta^{GARCH}(S_T)}{\partial S_T} \approx \dfrac{\hat{\Delta}^{GARCH}(S_0+k)-\hat{\Delta}^{GARCH}(S_0-k)}{2k}$ (cf. Engle and Rosenberg (1995)). We calculate these metrics for a variety of levels of moneyness, ranging from 0.9 (out-of-the-money) to 1.1 (in-the-money), and times to maturity, ranging from 30 days to 180 days.

Results are provided in Table 2. Columns labeled I present values for ST while those labeled II present values for S+O+E. First, consistent with the findings of Engle and Rosenberg (1995) and Duan (1995), we find that GARCH and Black-Scholes deltas and gammas may differ but not systematically so and not to a large degree. Our focus is on the difference between these measures across the different estimation methods ST and S+O+E, not between the models GARCH vs Black-Scholes. Hence, note that in columns labeled III we present the quotient of each value in I divided by the corresponding value from II, less 1. For example, the upper right-most value in the area labeled III is -0.8412 = 0.0101/0.0634 – 1. That is the GARCH delta using ST is about 84% lower than the GARCH delta calculated using S+O+E. We see that, although there does not appear to be a systematic relation, deltas and gammas may be significantly different depending on the method of estimation used.

[Insert Table 2 here]

First consider delta. Black-Scholes values from range about 10% higher for ST than for S+O+E when options are in the money to as much as 84% lower when they are out of the money. The difference tends to be negative when options are out of the money and positive when they are in the money. The trend is the same for GARCH deltas, although the magnitude ranges from 8.29% higher when options are in the money to 84.12% lower when they are out of the money. As a result, replicating and hedging portfolios will be significantly different based on the estimation method used, regardless of whether the agent applies a Black-Scholes or GARCH options pricing model.

For gammas, there does not appear be a pattern in the difference related to the moneyness of the options. However, the magnitude of the differences ranges from -66.31% to 49.84%, indicating that gammas are more strongly impacted by the method of estimation used than deltas. As a result, the updating dynamics dictated to maintain hedges will be substantially different depending on the data specification employed. The magnitude of these differences suggests that managers and investors would do well to keep this in mind when calculating risk management metrics for options.

3.4. Asymptotic behavior for ST, OT, and S+O+E

Proofs apply Brown (1971) results regarding Central Limit Theorem analogs for martingale differences. All lemma proofs in this section are deferred to the Appendix A.

3.4.1. Asymptotic behavior for ST

We seek to prove the following:

Theorem 3.1. *The estimator of ST is* $\hat{\theta}_T^{ST} = \arg\max_{\theta\in\Theta} L_T^{ST}(\theta)$. *It is consistent such that* $\hat{\theta}_T^{ST} \xrightarrow{p} \theta_0$ *and asymptotically Normal such that* $H_{ST0}^{-1/2} F_{ST0} T^{1/2}(\hat{\theta}_T^{ST} - \theta_0) \xrightarrow{\mathrm{D}} N(0, I_4)$, *where* $F_{ST0} = -E\left(\dfrac{\partial^2 L_T^{ST}(\theta_0)}{\partial\theta\partial\theta'}\right)$ *and* $H_{ST0} = E\left(T\dfrac{\partial L_T^{ST}(\theta_0)}{\partial\theta}\dfrac{\partial L_T^{ST}(\theta_0)}{\partial\theta'}\right)$.

To begin, the log-likelihood function for ST is given by (3.7) and it follows that:

$$\frac{\partial L_T^{ST}(\theta)}{\partial\theta} = -\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t(\theta)}\left(1-\frac{(y_t-\mu_t(\theta))^2}{h_t(\theta)}\right)-\frac{2\lambda}{h_t^{1/2}(\theta)}\left(\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\right)\right]\frac{\partial h_t(\theta)}{\partial\theta}, \tag{3.26}$$

$$\begin{aligned}\frac{\partial^2 L_T^{ST}(\theta)}{\partial\theta\partial\theta'} = &-\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t^2(\theta)}\left(-1+\frac{2(y_t-\mu_t(\theta))^2}{h_t(\theta)}\right)+\frac{2\lambda}{h_t^{3/2}(\theta)}\left(\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\right)+\frac{2\lambda^2}{h_t(\theta)}\right]\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\\ &-\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t(\theta)}\left(1-\frac{(y_t-\mu_t(\theta))^2}{h_t(\theta)}\right)-\frac{2\lambda}{h_t^{1/2}(\theta)}\left(\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\right)\right]\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'},\end{aligned} \tag{3.27}$$

$$\begin{aligned}\frac{\partial^3 L_T^{ST}(\theta)}{\partial\theta\partial\theta'\partial\theta} = &-\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t^3(\theta)}\left(2-\frac{6(y_t-\mu_t(\theta))^2}{h_t(\theta)}\right)-\frac{8\lambda}{h_t^{5/2}(\theta)}\left(\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\right)-\frac{4\lambda^2}{h_t^2(\theta)}\right]\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\frac{\partial h_t(\theta)}{\partial\theta}\\ &-\frac{3}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t^2(\theta)}\left(-1+\frac{2(y_t-\mu_t(\theta))^2}{h_t(\theta)}\right)+\frac{2\lambda}{h_t^{3/2}(\theta)}\left(\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\right)+\frac{2\lambda^2}{h_t(\theta)}\right]\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'}\frac{\partial h_t(\theta)}{\partial\theta}\\ &-\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t(\theta)}\left(1-\frac{(y_t-\mu_t(\theta))^2}{h_t(\theta)}\right)-\frac{2\lambda}{h_t^{1/2}(\theta)}\left(\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\right)\right]\frac{\partial^3 h_t(\theta)}{\partial\theta\partial\theta'\partial\theta},\end{aligned} \tag{3.28}$$

where $\mu_t(\theta) = r + \lambda h_t(\theta)$.

Consider the asymptotic Normality of the first derivative and the limit of the observed information matrix in (3.26) and (3.27), using Lemmas 3.1 and 3.2, respectively:

Lemma 3.1. *The form given by (3.26) evaluated at* $\theta = \theta_0$ *is asymptotically Gaussian,*

$$H_{ST0}^{-1/2} T^{1/2}\frac{\partial L_T^{ST}(\theta_0)}{\partial\theta} \xrightarrow{D} N(0, I_4),$$

where $H_{ST0} = E\left(T\dfrac{\partial L_T^{ST}(\theta_0)}{\partial\theta}\dfrac{\partial L_T^{ST}(\theta_0)}{\partial\theta'}\right) = E\left[\left(\dfrac{1}{2h_t^2(\theta_0)}+\dfrac{\lambda^2}{h_t(\theta_0)}\right)\dfrac{\partial h_t(\theta_0)}{\partial\theta}\dfrac{\partial h_t(\theta_0)}{\partial\theta'}\right]$.

Lemma 3.2. *The observed information matrix given by (3.27) evaluated at* $\theta=\theta_0$ *converges in probability to* $-F_{ST0}^{-1}\dfrac{\partial^2 L_T^{ST}(\theta_0)}{\partial\theta\partial\theta'}\xrightarrow{p} I_4$*, where* $F_{ST0}=-E\left(\dfrac{\partial^2 L_T^{ST}(\theta_0)}{\partial\theta\partial\theta'}\right)$.

Then, evaluating the third derivative of the likelihood function in (3.28), we seek to show that it is uniformly bounded in a neighborhood around the true parameter value θ_0. The neighborhood $N(\theta_0)$ around the true value θ_0 defined as

$$\begin{aligned}N(\theta_0)=\{\theta\,|\,0<\omega_L\le\omega_0\le\omega_U,\ 0<\alpha_L\le\alpha_0\le\alpha_U,\ 0<\beta_L\le\beta_0\le\beta_U,\ \gamma_L\le\gamma_0\le\gamma_U,\\ \alpha_U(\lambda+\gamma_U+1/2)^2+\beta_U<1\}.\end{aligned}\tag{3.29}$$

Lemma 3.3. *There exists* $N(\theta_0)$*, for all* $1\le i,j,k\le 4$*, for which*

$\sup_{\theta\in N(\theta_0)}\left|\dfrac{\partial^3 L_T^{ST}(\theta)}{\partial\theta_i\partial\theta_j\partial\theta_k}\right|\le g(\omega_L,\omega_U,\alpha_L,\alpha_U,\beta_L,\beta_U,\gamma_L,\gamma_U,T)\xrightarrow{a.s.}M<\infty$ *as* $T\to\infty$ *where* M *is constant.*

In order to prove Lemma 3.3, without loss of generality, consider the case $\theta_i=\theta_j=\theta_k=\beta$. The next lemma establishes that the individual terms of the third derivative $\left(\partial^3 L_T^{ST}/\partial\beta^3\right)(\theta)$ in (3.28) are uniformly bounded in the neighborhood $N(\theta_0)$.

Lemma 3.4. *With* $N(\theta_0)$ *defined in (3.29), then for any t,*

$$\omega_L\le\sup_{\theta\in N(\theta_0)}h_t(\theta)\le H_t\tag{3.30}$$

and

$$\sup_{\theta\in N(\theta_0)}h_{it}(\theta)\le H_{it},\ for\ i=1,2,3,\tag{3.31}$$

where $h_{it}(\theta)$*,* H_t*, and* H_{it} *are given by,*

$h_{it}(\theta) \equiv \dfrac{\partial^i h_t(\theta)}{\partial \beta^i}$, $H_t \equiv \dfrac{\max\{\omega_U + 2\alpha_U, h_0\}}{1-\alpha_U(\lambda+\gamma_U)^2 - \beta_U} + \left(\dfrac{\alpha_U}{\omega_L} + 2\alpha_U(\lambda+\gamma_U)^2\right)\sum_{i=1}^{t}\left(\alpha_U(\lambda+\gamma_U)^2+\beta_U\right)^{i-1}(y_{t-i}-r)^2$,

$H_{1t} \equiv \sum_{i=1}^{t-1} H_{t-i}\left(\alpha_U(\lambda+\gamma_U)^2+\beta_U\right)^{i-1}$, $H_{2t} \equiv 2\sum_{i=1}^{t-1}\left(1+\dfrac{\alpha_U(y_{t-i}-r)^2}{\omega_L^3}H_{1t-i}\right)H_{1t-i}\left(\alpha_U(\lambda+\gamma_U)^2+\beta_U\right)^{i-1}$,

and $H_{3t} \equiv 3\sum_{i=1}^{t-1}\left(2\dfrac{\alpha_U(y_{t-i}-r)^2}{\omega_L^4}H_{1t-i}^3 + 2\dfrac{\alpha_U(y_{t-i}-r)^2}{\omega_L^3}H_{2t-i}H_{1t-i} + H_{2t-i}\right)\left(\alpha_U(\lambda+\gamma_U)^2+\beta_U\right)^{i-1}$.

Proof of Theorem 3.1. From Lemma 3.3, we have that

$$\begin{bmatrix}0\\0\\0\\0\end{bmatrix} = \frac{\partial L_T^{ST}(\hat{\theta}_T^{ST})}{\partial\theta} \approx \frac{\partial L_T^{ST}(\theta_0)}{\partial\theta} + \frac{\partial^2 L_T^{ST}(\theta_0)}{\partial\theta\partial\theta'}(\hat{\theta}_T^{ST}-\theta_0)$$

and

$$\hat{\theta}_T^{ST} - \theta_0 \approx -\left(\frac{\partial^2 L_T^{ST}(\theta_0)}{\partial\theta\partial\theta'}\right)^{-1}\frac{\partial L_T^{ST}(\theta_0)}{\partial\theta}.$$

Combining with Lemmas 3.1 and 3.2, we complete the proof of Theorem 3.1.

3.4.2. Asymptotic behavior for OT

We seek to prove the following:

Theorem 3.2. *The estimator of OT is* $\hat{\theta}_T^{OT} = \arg\max_{\theta\in\Theta} L_T^{OT}(\theta)$. *It is asymptotically biased such that* $\hat{\theta}_T^{OT} - \theta_1 \xrightarrow{p} 0$, *where* $\theta_1 \equiv \theta_0 + \left(\dfrac{\partial^2 L_T^{OT}(\theta_0)}{\partial\theta\partial\theta'}\right)^{-1}\dfrac{1}{T}\sum_{t=1}^{T}\dfrac{1}{J_t(\theta_0)}\dfrac{\partial J_t(\theta_0)}{\partial\theta}$, *and asymptotically Normal such that* $H_{OT0}^{-1/2}F_{OT0}T^{1/2}(\hat{\theta}_T^{OT}-\theta_1)\overset{A}{\sim}N(0,I_4)$, *where* $F_{OT0} = -E\left(\dfrac{\partial^2 L_T^{OT}(\theta_0)}{\partial\theta\partial\theta'}\right)$ *and* $H_{OT0} = Var\left(T^{1/2}\dfrac{\partial L_T^{OT}(\theta_0)}{\partial\theta}\right)$.

To begin, the log-likelihood function for OT is given by (3.11) and it follows that:

$$\frac{\partial L_T^{OT}(\theta)}{\partial\theta} = -\frac{1}{T}\sum_{t=1}^{T}\left[\frac{1}{J_t(\theta)}\frac{\partial J_t(\theta)}{\partial\theta} + \frac{1}{2}\left(\frac{1}{h_t(\theta)} - \frac{Y_t^2(\theta)}{h_t^2(\theta)}\right)\frac{\partial h_t(\theta)}{\partial\theta} + \frac{Y_t(\theta)}{h_t(\theta)}\frac{\partial Y_t(\theta)}{\partial\theta}\right], \tag{3.32}$$

$$\frac{\partial^2 L_T^{OT}(\theta)}{\partial\theta\partial\theta'} = -\frac{1}{T}\sum_{t=1}^{T}\left[\begin{array}{l} -\frac{1}{J_t^2(\theta)}\frac{\partial J_t(\theta)}{\partial\theta}\frac{\partial J_t(\theta)}{\partial\theta'} + \frac{1}{J_t(\theta)}\frac{\partial^2 J_t(\theta)}{\partial\theta\partial\theta'} + \frac{1}{2}\left(\frac{-1}{h_t^2(\theta)} + \frac{2Y_t^2(\theta)}{h_t^3(\theta)}\right)\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'} \\ +\frac{1}{2}\left(\frac{1}{h_t(\theta)} - \frac{Y_t^2(\theta)}{h_t^2(\theta)}\right)\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'} - \frac{2Y_t(\theta)}{h_t^2(\theta)}\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial Y_t(\theta)}{\partial\theta'} + \frac{1}{h_t(\theta)}\frac{\partial Y_t(\theta)}{\partial\theta}\frac{\partial Y_t(\theta)}{\partial\theta'} + \frac{Y_t(\theta)}{h_t(\theta)}\frac{\partial^2 Y_t(\theta)}{\partial\theta\partial\theta'} \end{array}\right], \tag{3.33}$$

$$\frac{\partial^3 L_T^{OT}(\theta)}{\partial\theta\partial\theta'\partial\theta} = -\frac{1}{T}\sum_{t=1}^{T}\left[\begin{array}{l} \frac{2}{J_t^3(\theta)}\frac{\partial J_t(\theta)}{\partial\theta}\frac{\partial J_t(\theta)}{\partial\theta'}\frac{\partial J_t(\theta)}{\partial\theta} - \frac{3}{J_t^2(\theta)}\frac{\partial^2 J_t(\theta)}{\partial\theta\partial\theta'}\frac{\partial J_t(\theta)}{\partial\theta} + \frac{1}{J_t(\theta)}\frac{\partial^3 J_t(\theta)}{\partial\theta\partial\theta'\partial\theta} \\ +\left(\frac{1}{h_t^3(\theta)} - \frac{3Y_t^2(\theta)}{h_t^4(\theta)}\right)\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\frac{\partial h_t(\theta)}{\partial\theta} + \frac{3}{2}\left(\frac{-1}{h_t^2(\theta)} + \frac{2Y_t^2(\theta)}{h_t^3(\theta)}\right)\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'}\frac{\partial h_t(\theta)}{\partial\theta} \\ +\frac{1}{2}\left(\frac{1}{h_t(\theta)} - \frac{Y_t^2(\theta)}{h_t^2(\theta)}\right)\frac{\partial^3 h_t(\theta)}{\partial\theta\partial\theta'\partial\theta} + \frac{6Y_t(\theta)}{h_t^3(\theta)}\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\frac{\partial Y_t(\theta)}{\partial\theta} \\ -\frac{3Y_t(\theta)}{h_t^2(\theta)}\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'}\frac{\partial Y_t(\theta)}{\partial\theta} - \frac{3}{h_t^2(\theta)}\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial Y_t(\theta)}{\partial\theta'}\frac{\partial Y_t(\theta)}{\partial\theta} - \frac{3Y_t(\theta)}{h_t^2(\theta)}\frac{\partial^2 Y_t(\theta)}{\partial\theta\partial\theta'}\frac{\partial h_t(\theta)}{\partial\theta} \\ +\frac{3}{h_t(\theta)}\frac{\partial^2 Y_t(\theta)}{\partial\theta\partial\theta'}\frac{\partial Y_t(\theta)}{\partial\theta} + \frac{Y_t(\theta)}{h_t(\theta)}\frac{\partial^3 Y_t(\theta)}{\partial\theta\partial\theta'\partial\theta} \end{array}\right], \tag{3.34}$$

where

$$\begin{aligned}\frac{\partial J_t(\theta)}{\partial\theta} = &\frac{1}{\pi}\int_0^\infty e^{X_{11,t}(\phi,\theta)}\left[\left(\frac{\sin\left(X_{31,t}(\phi,\theta)\right)}{\phi} + \cos\left(X_{31,t}(\phi,\theta)\right)\right)\frac{\partial X_{11,t}(\phi,\theta)}{\partial\theta} + \left(\frac{\cos\left(X_{31,t}(\phi,\theta)\right)}{\phi} - \sin\left(X_{31,t}(\phi,\theta)\right)\right)\frac{\partial X_{31,t}(\phi,\theta)}{\partial\theta}\right]d\phi \\ &+ e^{-r(T-t)}\frac{K}{S_t(\theta)}\frac{1}{\pi}\int_0^\infty e^{X_{10,t}(\phi,\theta)}\left[\frac{\cos\left(X_{30,t}(\phi,\theta)\right)}{S_t(\theta)}\frac{\partial S_t(\theta)}{\partial\theta} - \cos\left(X_{30,t}(\phi,\theta)\right)\frac{\partial X_{10,t}(\phi,\theta)}{\partial\theta} + \sin\left(X_{30,t}(\phi,\theta)\right)\frac{\partial X_{30,t}(\phi,\theta)}{\partial\theta}\right]d\phi,\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 J_t(\theta)}{\partial\theta\partial\theta'} = &\frac{1}{\pi}\int_0^\infty e^{X_{11,t}(\phi,\theta)}\left[\begin{array}{l}\left(\frac{\sin\left(X_{31,t}(\phi,\theta)\right)}{\phi} + \cos\left(X_{31,t}(\phi,\theta)\right)\right)\left(\frac{\partial X_{11,t}(\phi,\theta)}{\partial\theta}\frac{\partial X_{11,t}(\phi,\theta)}{\partial\theta'} - \frac{\partial X_{31,t}(\phi,\theta)}{\partial\theta}\frac{\partial X_{31,t}(\phi,\theta)}{\partial\theta'} + \frac{\partial^2 X_{11,t}(\phi,\theta)}{\partial\theta\partial\theta'}\right) \\ +\left(\frac{\cos\left(X_{31,t}(\phi,\theta)\right)}{\phi} - \sin\left(X_{31,t}(\phi,\theta)\right)\right)\left(2\frac{\partial X_{31,t}(\phi,\theta)}{\partial\theta}\frac{\partial X_{11,t}(\phi,\theta)}{\partial\theta'} + \frac{\partial^2 X_{31,t}(\phi,\theta)}{\partial\theta\partial\theta'}\right)\end{array}\right]d\phi \\ &+ e^{-r(T-t)}\frac{K}{S_t(\theta)}\frac{1}{\pi}\int_0^\infty e^{X_{10,t}(\phi,\theta)}\left[\begin{array}{l}\frac{\cos\left(X_{30,t}(\phi,\theta)\right)}{S_t(\theta)}\left(2\frac{\partial X_{10,t}(\phi,\theta)}{\partial\theta}\frac{\partial S_t(\theta)}{\partial\theta'} - 2\frac{\partial S_t(\theta)}{\partial\theta}\frac{\partial S_t(\theta)}{\partial\theta'} + \frac{\partial^2 S_t(\theta)}{\partial\theta\partial\theta'}\right) - \frac{2\sin\left(X_{30,t}(\phi,\theta)\right)}{S_t(\theta)}\frac{\partial S_t(\theta)}{\partial\theta}\frac{\partial X_{30,t}(\phi,\theta)}{\partial\theta'} \\ -\cos\left(X_{30,t}(\phi,\theta)\right)\left(\frac{\partial X_{10,t}(\phi,\theta)}{\partial\theta}\frac{\partial X_{10,t}(\phi,\theta)}{\partial\theta'} + \frac{\partial^2 X_{10,t}(\phi,\theta)}{\partial\theta\partial\theta'} - \frac{\partial X_{30,t}(\phi,\theta)}{\partial\theta}\frac{\partial X_{30,t}(\phi,\theta)}{\partial\theta'}\right) \\ +\sin\left(X_{30,t}(\phi,\theta)\right)\left(2\frac{\partial X_{10,t}(\phi,\theta)}{\partial\theta}\frac{\partial X_{30,t}(\phi,\theta)}{\partial\theta'} + \frac{\partial^2 X_{30,t}(\phi,\theta)}{\partial\theta\partial\theta'}\right)\end{array}\right]d\phi,\end{aligned}$$

$$
\begin{aligned}
\frac{\partial^3 J_t(\theta)}{\partial\theta\partial\theta'\partial\theta} = \frac{1}{\pi}\int_0^\infty e^{X_{11,t}(\phi,\theta)} & \left[\begin{array}{l} \left(\frac{\sin\left(X_{31,t}(\phi,\theta)\right)}{\phi}+\cos\left(X_{31,t}(\phi,\theta)\right)\right)\left(\begin{array}{l} \frac{\partial X_{11,t}(\phi,\theta)}{\partial\theta}\frac{\partial X_{11,t}(\phi,\theta)}{\partial\theta'}\frac{\partial X_{11,t}(\phi,\theta)}{\partial\theta}-3\frac{\partial X_{31,t}(\phi,\theta)}{\partial\theta}\frac{\partial X_{31,t}(\phi,\theta)}{\partial\theta'}\frac{\partial X_{11,t}(\phi,\theta)}{\partial\theta} \\ +3\frac{\partial^2 X_{11,t}(\phi,\theta)}{\partial\theta\partial\theta'}\frac{\partial X_{11,t}(\phi,\theta)}{\partial\theta}-3\frac{\partial^2 X_{31,t}(\phi,\theta)}{\partial\theta\partial\theta'}\frac{\partial X_{31,t}(\phi,\theta)}{\partial\theta}+\frac{\partial^3 X_{11,t}(\phi,\theta)}{\partial\theta\partial\theta'\partial\theta} \end{array} \right) \\ +\left(\frac{\cos\left(X_{31,t}(\phi,\theta)\right)}{\phi}-\sin\left(X_{31,t}(\phi,\theta)\right)\right)\left(\begin{array}{l} \frac{\partial^3 X_{31,t}(\phi,\theta)}{\partial\theta\partial\theta'\partial\theta}+3\frac{\partial^2 X_{31,t}(\phi,\theta)}{\partial\theta\partial\theta'}\frac{\partial X_{11,t}(\phi,\theta)}{\partial\theta}+3\frac{\partial^2 X_{11,t}(\phi,\theta)}{\partial\theta\partial\theta'}\frac{\partial X_{31,t}(\phi,\theta)}{\partial\theta} \\ +3\frac{\partial X_{31,t}(\phi,\theta)}{\partial\theta}\frac{\partial X_{11,t}(\phi,\theta)}{\partial\theta'}\frac{\partial X_{11,t}(\phi,\theta)}{\partial\theta}-\frac{\partial X_{31,t}(\phi,\theta)}{\partial\theta}\frac{\partial X_{31,t}(\phi,\theta)}{\partial\theta'}\frac{\partial X_{31,t}(\phi,\theta)}{\partial\theta} \end{array} \right) \end{array} \right] d\phi \\
+e^{-r(T-t)}\frac{K}{S_t(\theta)}\frac{1}{\pi}\int_0^\infty e^{X_{10,t}(\phi,\theta)} & \left[\begin{array}{l} \frac{\cos\left(X_{30,t}(\phi,\theta)\right)}{S_t(\theta)}\left(\begin{array}{l} \frac{\partial^3 S_t(\theta)}{\partial\theta\partial\theta'\partial\theta}-2\frac{\partial S_t(\theta)}{\partial\theta}\frac{\partial S_t(\theta)}{\partial\theta'}\frac{\partial X_{10,t}(\phi,\theta)}{\partial\theta}-4\frac{\partial^2 S_t(\theta)}{\partial\theta\partial\theta'}\frac{\partial S_t(\theta)}{\partial\theta}+3\frac{\partial^2 X_{10,t}(\phi,\theta)}{\partial\theta\partial\theta'}\frac{\partial S_t(\theta)}{\partial\theta} \\ +3\frac{\partial^2 S_t(\theta)}{\partial\theta\partial\theta'}\frac{\partial X_{10,t}(\phi,\theta)}{\partial\theta}-3\frac{\partial S_t(\theta)}{\partial\theta}\frac{\partial X_{30,t}(\phi,\theta)}{\partial\theta'}\frac{\partial X_{30,t}(\phi,\theta)}{\partial\theta}+3\frac{\partial X_{10,t}(\phi,\theta)}{\partial\theta}\frac{\partial X_{10,t}(\phi,\theta)}{\partial\theta'}\frac{\partial S_t(\theta)}{\partial\theta} \end{array} \right) \\ +\frac{\cos\left(X_{30,t}(\phi,\theta)\right)}{S_t^2(\theta)}\left(-4\frac{\partial X_{10,t}(\phi,\theta)}{\partial\theta}\frac{\partial S_t(\theta)}{\partial\theta'}\frac{\partial S_t(\theta)}{\partial\theta}+4\frac{\partial S_t(\theta)}{\partial\theta}\frac{\partial S_t(\theta)}{\partial\theta'}\frac{\partial S_t(\theta)}{\partial\theta}-2\frac{\partial^2 S_t(\theta)}{\partial\theta\partial\theta'}\frac{\partial S_t(\theta)}{\partial\theta}\right) \\ +\frac{\sin\left(X_{30,t}(\phi,\theta)\right)}{S_t(\theta)}\left(\begin{array}{l} -6\frac{\partial X_{10,t}(\phi,\theta)}{\partial\theta}\frac{\partial S_t(\theta)}{\partial\theta'}\frac{\partial X_{30,t}(\phi,\theta)}{\partial\theta}-3\frac{\partial^2 S_t(\theta)}{\partial\theta\partial\theta'}\frac{\partial X_{30,t}(\phi,\theta)}{\partial\theta}-2\frac{\partial^2 X_{10,t}(\phi,\theta)}{\partial\theta\partial\theta'}\frac{\partial S_t(\theta)}{\partial\theta} \\ -\frac{\partial^2 X_{30,t}(\phi,\theta)}{\partial\theta\partial\theta'}\frac{\partial S_t(\theta)}{\partial\theta}+2\frac{\partial S_t(\theta)}{\partial\theta}\frac{\partial S_t(\theta)}{\partial\theta'}\frac{\partial X_{30,t}(\phi,\theta)}{\partial\theta} \end{array} \right) \\ +\frac{4\sin\left(X_{30,t}(\phi,\theta)\right)}{S_t^2(\theta)}\frac{\partial S_t(\theta)}{\partial\theta}\frac{\partial X_{30,t}(\phi,\theta)}{\partial\theta'}\frac{\partial S_t(\theta)}{\partial\theta} \\ +\sin\left(X_{30,t}(\phi,\theta)\right)\left(\begin{array}{l} -\frac{\partial X_{30,t}(\phi,\theta)}{\partial\theta}\frac{\partial X_{30,t}(\phi,\theta)}{\partial\theta'}\frac{\partial X_{30,t}(\phi,\theta)}{\partial\theta}+\frac{\partial^3 X_{30,t}(\phi,\theta)}{\partial\theta\partial\theta'\partial\theta}+3\frac{\partial^2 X_{10,t}(\phi,\theta)}{\partial\theta\partial\theta'}\frac{\partial X_{30,t}(\phi,\theta)}{\partial\theta} \\ +3\frac{\partial^2 X_{30,t}(\phi,\theta)}{\partial\theta\partial\theta'}\frac{\partial X_{10,t}(\phi,\theta)}{\partial\theta}+3\frac{\partial X_{10,t}(\phi,\theta)}{\partial\theta}\frac{\partial X_{30,t}(\phi,\theta)}{\partial\theta'}\frac{\partial X_{10,t}(\phi,\theta)}{\partial\theta} \end{array} \right) \\ +\cos\left(X_{30,t}(\phi,\theta)\right)\left(\begin{array}{l} 3\frac{\partial X_{10,t}(\phi,\theta)}{\partial\theta}\frac{\partial X_{30,t}(\phi,\theta)}{\partial\theta'}\frac{\partial X_{30,t}(\phi,\theta)}{\partial\theta}-3\frac{\partial^2 X_{10,t}(\phi,\theta)}{\partial\theta\partial\theta'}\frac{\partial X_{10,t}(\phi,\theta)}{\partial\theta} \\ +3\frac{\partial^2 X_{30,t}(\phi,\theta)}{\partial\theta\partial\theta'}\frac{\partial X_{30,t}(\phi,\theta)}{\partial\theta}-\frac{\partial X_{10,t}(\phi,\theta)}{\partial\theta}\frac{\partial X_{10,t}(\phi,\theta)}{\partial\theta'}\frac{\partial X_{10,t}(\phi,\theta)}{\partial\theta}-\frac{\partial^3 X_{10,t}(\phi,\theta)}{\partial\theta\partial\theta'} \end{array} \right) \end{array} \right] d\phi,
\end{aligned}
$$

$$
\frac{\partial Y_t(\theta)}{\partial\theta}=\frac{1}{S_t(\theta)}\frac{\partial S_t(\theta)}{\partial\theta}-\frac{1}{S_{t-1}(\theta)}\frac{\partial S_{t-1}(\theta)}{\partial\theta}-\lambda\frac{\partial h_t(\theta)}{\partial\theta},
$$

$$
\begin{aligned}
\frac{\partial^2 Y_t(\theta)}{\partial\theta\partial\theta'}=&-\frac{1}{S_t^2(\theta)}\frac{\partial S_t(\theta)}{\partial\theta}\frac{\partial S_t(\theta)}{\partial\theta'}+\frac{1}{S_t(\theta)}\frac{\partial^2 S_t(\theta)}{\partial\theta\partial\theta'}+\frac{1}{S_{t-1}^2(\theta)}\frac{\partial S_{t-1}(\theta)}{\partial\theta}\frac{\partial S_{t-1}(\theta)}{\partial\theta'} \\
&-\frac{1}{S_{t-1}(\theta)}\frac{\partial^2 S_{t-1}(\theta)}{\partial\theta\partial\theta'}-\lambda\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^3 Y_t(\theta)}{\partial\theta\partial\theta'\partial\theta}=&\frac{2}{S_t^3(\theta)}\frac{\partial S_t(\theta)}{\partial\theta}\frac{\partial S_t(\theta)}{\partial\theta'}\frac{\partial S_t(\theta)}{\partial\theta}-\frac{3}{S_t^2(\theta)}\frac{\partial^2 S_t(\theta)}{\partial\theta\partial\theta'}\frac{\partial S_t(\theta)}{\partial\theta}+\frac{1}{S_t(\theta)}\frac{\partial^3 S_t(\theta)}{\partial\theta\partial\theta'\partial\theta} \\
&-\frac{2}{S_{t-1}^3(\theta)}\frac{\partial S_{t-1}(\theta)}{\partial\theta}\frac{\partial S_{t-1}(\theta)}{\partial\theta'}\frac{\partial S_{t-1}(\theta)}{\partial\theta}+\frac{3}{S_{t-1}^2(\theta)}\frac{\partial^2 S_{t-1}(\theta)}{\partial\theta\partial\theta'}\frac{\partial S_{t-1}(\theta)}{\partial\theta} \\
&-\frac{1}{S_{t-1}(\theta)}\frac{\partial^3 S_{t-1}(\theta)}{\partial\theta\partial\theta'\partial\theta}-\lambda\frac{\partial^3 h_t(\theta)}{\partial\theta\partial\theta'\partial\theta},
\end{aligned}
$$

$$
\frac{\partial X_{1m,t}(\phi,\theta)}{\partial\theta}=\frac{\partial A_{1m,t}(\phi,\theta)}{\partial\theta}+\frac{\partial B_{1m,t}(\phi,\theta)}{\partial\theta}h_{t+1}(\theta)+B_{1m,t}(\phi,\theta)\frac{\partial h_{t+1}(\theta)}{\partial\theta},
$$

$$\frac{\partial^2 X_{1m,t}(\phi,\theta)}{\partial\theta\partial\theta'}=\frac{\partial^2 A_{1m,t}(\phi,\theta)}{\partial\theta\partial\theta'}+\frac{\partial^2 B_{1m,t}(\phi,\theta)}{\partial\theta\partial\theta'}h_{t+1}(\theta)+2\frac{\partial B_{1m,t}(\phi,\theta)}{\partial\theta}\frac{\partial h_{t+1}(\theta)}{\partial\theta}$$
$$+B_{1m,t}(\phi,\theta)\frac{\partial^2 h_{t+1}(\theta)}{\partial\theta\partial\theta'},$$

$$\frac{\partial^3 X_{1m,t}(\phi,\theta)}{\partial\theta\partial\theta'\partial\theta}=\frac{\partial^3 A_{1m,t}(\phi,\theta)}{\partial\theta\partial\theta'\partial\theta}+\frac{\partial^3 B_{1m,t}(\phi,\theta)}{\partial\theta\partial\theta'\partial\theta}h_{t+1}(\theta)+3\frac{\partial^2 B_{1m,t}(\phi,\theta)}{\partial\theta\partial\theta'}\frac{\partial h_{t+1}(\theta)}{\partial\theta}$$
$$+3\frac{\partial^2 h_{t+1}(\theta)}{\partial\theta\partial\theta'}\frac{\partial B_{1m,t}(\phi,\theta)}{\partial\theta}+B_{1m,t}(\phi,\theta)\frac{\partial^3 h_{t+1}(\theta)}{\partial\theta\partial\theta'\partial\theta},$$

$$\frac{\partial X_{3m,t}(\phi,\theta)}{\partial\theta}=\frac{\phi}{S_t(\theta)}\frac{\partial S_t(\theta)}{\partial\theta}+\frac{\partial A_{2m,t}(\phi,\theta)}{\partial\theta}+h_{t+1}(\theta)\frac{\partial B_{2m,t}(\phi,\theta)}{\partial\theta}+B_{2m,t}(\phi,\theta)\frac{\partial h_{t+1}(\theta)}{\partial\theta},$$

$$\frac{\partial^2 X_{3m,t}(\phi,\theta)}{\partial\theta\partial\theta'}=-\frac{\phi}{S_t^2(\theta)}\frac{\partial S_t(\theta)}{\partial\theta}\frac{\partial S_t(\theta)}{\partial\theta'}+\frac{\phi}{S_t(\theta)}\frac{\partial^2 S_t(\theta)}{\partial\theta\partial\theta'}+\frac{\partial^2 A_{2m,t}(\phi,\theta)}{\partial\theta\partial\theta'}$$
$$+2\frac{\partial B_{2m,t}(\phi,\theta)}{\partial\theta}\frac{\partial h_{t+1}(\theta)}{\partial\theta'}+h_{t+1}(\theta)\frac{\partial^2 B_{2m,t}(\phi,\theta)}{\partial\theta\partial\theta'}+B_{2m,t}(\phi,\theta)\frac{\partial^2 h_{t+1}(\theta)}{\partial\theta\partial\theta'},$$

and

$$\frac{\partial^3 X_{3m,t}(\phi,\theta)}{\partial\theta\partial\theta'}=\frac{2\phi}{S_t^3(\theta)}\frac{\partial S_t(\theta)}{\partial\theta}\frac{\partial S_t(\theta)}{\partial\theta'}\frac{\partial S_t(\theta)}{\partial\theta}-\frac{3\phi}{S_t^2(\theta)}\frac{\partial^2 S_t(\theta)}{\partial\theta\partial\theta'}\frac{\partial S_t(\theta)}{\partial\theta}+\frac{\phi}{S_t(\theta)}\frac{\partial^3 S_t(\theta)}{\partial\theta\partial\theta'\partial\theta}$$
$$+\frac{\partial^3 A_{2m,t}(\phi,\theta)}{\partial\theta\partial\theta'\partial\theta}+3\frac{\partial^2 B_{2m,t}(\phi,\theta)}{\partial\theta\partial\theta'}\frac{\partial h_{t+1}(\theta)}{\partial\theta}+3\frac{\partial^2 h_{t+1}(\theta)}{\partial\theta\partial\theta'}\frac{\partial B_{2m,t}(\phi,\theta)}{\partial\theta}$$
$$+h_{t+1}(\theta)\frac{\partial^3 B_{2m,t}(\phi,\theta)}{\partial\theta\partial\theta'\partial\theta}+B_{2m,t}(\phi,\theta)\frac{\partial^3 h_{t+1}(\theta)}{\partial\theta\partial\theta'\partial\theta}.$$

Consider the asymptotic Normality of the first derivative and the limit of the observed information matrix in (3.32) and (3.33), using Lemmas 3.5 and 3.6, respectively:

Lemma 3.5. *The form given by (3.32) evaluated at* $\theta=\theta_0$ *is asymptotically Gaussian,*

$$H_{OT0}^{-1/2}T^{1/2}\left(\frac{\partial L_T^{OT}(\theta_0)}{\partial\theta}+\frac{1}{T}\sum_{t=1}^{T}\frac{1}{J_t(\theta_0)}\frac{\partial J_t(\theta_0)}{\partial\theta}\right)\xrightarrow{D}N\left(0,I_4\right),$$

where $H_{OT0}=Var\left(T^{1/2}\frac{\partial L_T^{OT}(\theta_0)}{\partial\theta}\right)=E\left[\frac{1}{2h_t^2(\theta)}\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\right]+E\left[\frac{1}{h_t(\theta)}\frac{\partial Y_t(\theta)}{\partial\theta}\frac{\partial Y_t(\theta)}{\partial\theta'}\right].$

Lemma 3.6. *The observed information matrix given by (3.33) evaluated at* $\theta = \theta_0$ *converges in probability to* $-F_{OT0}^{-1}\dfrac{\partial^2 L_T^{OT}(\theta_0)}{\partial\theta\partial\theta'} \xrightarrow{p} I_4$, *where*

$$
\begin{aligned}
F_{OT0} &= -E\left(\frac{\partial^2 L_T^{OT}(\theta_0)}{\partial\theta\partial\theta'}\right) \\
&= E\left[\frac{1}{J_t(\theta_0)}\frac{\partial^2 J_t(\theta_0)}{\partial\theta\partial\theta'}\right] - E\left[\frac{1}{J_t^2(\theta_0)}\frac{\partial J_t(\theta_0)}{\partial\theta}\frac{\partial J_t(\theta_0)}{\partial\theta'}\right] + E\left[\frac{1}{2h_t^2(\theta_0)}\frac{\partial h_t(\theta_0)}{\partial\theta}\frac{\partial h_t(\theta_0)}{\partial\theta'}\right] + E\left[\frac{1}{h_t(\theta_0)}\frac{\partial Y_t(\theta_0)}{\partial\theta}\frac{\partial Y_t(\theta_0)}{\partial\theta'}\right].
\end{aligned}
$$

Then, evaluating the third derivative of the likelihood function in (3.34), we seek to show that it is uniformly bounded in a neighborhood around the true parameter value θ_0.

Lemma 3.7. *There exists* $N(\theta_0)$*, for all* $1 \le i, j, k \le 4$*, for which*

$\sup_{\theta\in N(\theta_0)}\left|\dfrac{\partial^3 L_T^{OT}(\theta)}{\partial\theta_i\partial\theta_j\partial\theta_k}\right| \le g(\omega_L,\omega_U,\alpha_L,\alpha_U,\beta_L,\beta_U,\gamma_L,\gamma_U,T) \xrightarrow{a.s.} M < \infty$ *as* $T\to\infty$ *where* M *is constant.*

In order to prove Lemma 3.7, again, we only consider the case $\theta_i = \theta_j = \theta_k = \beta$. We want to show that the individual terms of the third derivative $\left(\partial^3 L_T^{OT}/\partial\beta^3\right)(\theta)$ in (3.34) are uniformly bounded in the neighborhood $N(\theta_0)$. It has to apply these results of Lemmas 3.8-3.13. First, we prove that these lemmas.

Lemma 3.8. *With* $N(\theta_0)$ *defined in (3.29), then for any t,* $m = 0,1$,

$$B_{1m,t}(\phi,\theta) = -\frac{\phi^2}{2}b_{1m,t}(\phi,\theta) \tag{3.35}$$

and

$$B_{2m,t}(\phi,\theta) = \frac{\phi}{2}b_{2m,t}(\phi,\theta), \tag{3.36}$$

where

$$b_{1m,t}(\phi,\theta) = \beta b_{1m,t+1}(\phi,\theta) + \frac{(1+\alpha\phi^2 b_{1m,t+1}(\phi,\theta))\left[1-\alpha\left(m(2\gamma_Q-1)-\gamma_Q^2\right)b_{1m,t+1}(\phi,\theta)\right]+\alpha^2\left(\gamma_Q^2-m(2\gamma_Q-1)\right)b_{2m,t+1}^2(\phi,\theta)+2\alpha(m-\gamma_Q)b_{2m,t+1}(\phi,\theta)}{(1+\alpha\phi^2 b_{1m,t+1}(\phi,\theta))^2+\alpha^2\phi^2 b_{2m,t+1}^2(\phi,\theta)},$$

$$b_{2m,t}(\phi,\theta) = 2\gamma_Q - 1 + \beta b_{2m,t+1}(\phi,\theta) + \frac{2(m-\gamma_Q)(1+\alpha\phi^2 b_{1m,t+1}(\phi,\theta))+\alpha[(m-\gamma_Q)^2-\phi^2]b_{2m,t+1}(\phi,\theta)}{(1+\alpha\phi^2 b_{1m,t+1}(\phi,\theta))^2+\alpha^2\phi^2 b_{2m,t+1}^2(\phi,\theta)},$$ *and*

$b_{1m,T}(\phi,\theta) = b_{2m,T}(\phi,\theta) = 0.$

Lemma 3.9. *With* $N(\theta_0)$ *defined in (3.29), then for any t,* $m=0,1$*,* $i=1,2$*, and* $k=1,2,3$*,*

$$0 \le \sup_{\theta\in N(\theta_0)} b_{1m,t}(\phi,\theta) \le b_{1mU}, \tag{3.37}$$

$$b_{2mL} \le \sup_{\theta\in N(\theta_0)} b_{2m,t}(\phi,\theta) \le b_{2mU}, \tag{3.38}$$

and

$$\sup_{\theta\in N(\theta_0)} b_{imk,t}(\phi,\theta) \le b_{imkU}, \tag{3.39}$$

where $b_{imk,t}(\phi,\theta) = \dfrac{\partial^k b_{im,t}(\phi,\theta)}{\partial\beta^k}$ *and the constants,* $b_{1mU}, b_{2mL}, b_{2mU}$*,* b_{imkU}*, are functions of* $\omega_L, \omega_U, \alpha_L, \alpha_U, \beta_L, \beta_U, \gamma_L, \gamma_U, \lambda$.

Lemma 3.10. *With* $N(\theta_0)$ *defined in (3.29), then for any t,* $m=0,1$*, and* $k=1,2,3$*,*

$$-\frac{\phi^2}{2} b_{1mU} \le \sup_{\theta\in N(\theta_0)} B_{1m,t}(\phi,\theta) \le 0, \tag{3.40}$$

$$\sup_{\theta\in N(\theta_0)} B_{1mk,t}(\phi,\theta) \le \frac{\phi^2}{2} b_{1mkU}, \tag{3.41}$$

$$\frac{\phi}{2} b_{2mL} \le \sup_{\theta\in N(\theta_0)} B_{2m,t}(\phi,\theta) \le \frac{\phi}{2} b_{2mU}, \tag{3.42}$$

and

$$\sup_{\theta\in N(\theta_0)} B_{2mk,t}(\phi,\theta) \le \frac{\phi}{2} b_{2mkU}. \tag{3.43}$$

Lemma 3.11. *With* $N(\theta_0)$ *defined in (3.29), then for any t,* $i=1,2$*, and* $m=0,1$*,*

$$\sup_{\theta\in N(\theta_0)} A_{1m,t}(\phi,\theta)\le mr(T-t)-\frac{\omega_L}{2}\phi^2, \tag{3.44}$$

$$\sup_{\theta\in N(\theta_0)} A_{2m,t}(\phi,\theta)\le\left[\phi\left(r+\frac{\omega_U b_{2mU}}{2}\right)+2c\pi\right](T-t), \tag{3.45}$$

and

$$\sup_{\theta\in N(\theta_0)} A_{im1,t}(\phi,\theta)\le\frac{\phi^{3-i}}{2}A_{im1,t}, \tag{3.46}$$

$$\sup_{\theta\in N(\theta_0)} A_{im2,t}(\phi,\theta)\le\frac{\phi^{3-i}}{2}\left(A_{im21,t}+\phi^2 A_{im22,t}\right), \tag{3.47}$$

$$\sup_{\theta\in N(\theta_0)} A_{im3,t}(\phi,\theta)\le\frac{\phi^{3-i}}{2}\left(A_{im31,t}+\phi^2 A_{im32,t}+\phi^4 A_{im33,t}\right), \tag{3.48}$$

where c *is a positive integer,*

$$A_{1m1,t}=\left(\omega_U b_{1m1U}+\alpha_U b_{1m1U}+\alpha_U^2 b_{2mU}b_{2m1U}\right)(T-t),$$

$$A_{1m21,t}=\left(\omega_U b_{1m2U}+\frac{b_{1m1U}^2}{b_{2mL}^2}+\alpha_U b_{1m2U}+\alpha_U^2 b_{2mU}b_{2m2U}+\alpha_U^2 b_{2m1U}^2\right)(T-t),$$

$$A_{1m22,t}=2\left(\alpha_U b_{1m1U}+\alpha_U^2 b_{2m1U}\right)\left(\alpha_U b_{1m1U}+\alpha_U^2 b_{2mU}b_{2m1U}\right)(T-t),$$

$$A_{1m31,t}=\left(\omega_U b_{1m3U}+\frac{3b_{1m1U}b_{1m2U}}{b_{2mL}^2}+\alpha_U b_{1m3U}+3\alpha_U^2 b_{2m1U}b_{2m2U}+\alpha_U^2 b_{2mU}b_{2m3U}\right)(T-t),$$

$$A_{1m32,t}=\begin{bmatrix}4(\frac{b_{1m1U}^2}{b_{2mL}^2}+\alpha_U b_{1m2U}+\alpha_U^2 b_{2mU}b_{2m2U}+\alpha_U^2 b_{2m1U})(\alpha_U b_{1m1U}+\alpha_U^2 b_{2mU}b_{2m1U})\\ +2(\alpha_U b_{1m1U}+\alpha_U^2 b_{2m1U})(\frac{b_{1m1U}^2}{b_{2mL}^2}+\alpha_U b_{1m2U}+\alpha_U^2 b_{2mU}b_{2m2U}+\alpha_U^2 b_{2m1U})\end{bmatrix}(T-t),$$

$$A_{1m3,t} = 8(\alpha_U b_{1m1U} + \alpha_U^2 b_{2m1U})(\alpha_U b_{1m1U} + \alpha_U^2 b_{2mU} b_{2m1U})^2 (T-t),$$

$$A_{2m1,t} = \left(\omega_U b_{2m1U} + \alpha_U b_{2m1U} + \alpha_U^2 \frac{b_{1m1U}}{b_{2mL}} \right)(T-t),$$

$$A_{2m21,t} = \left(\omega_U b_{2m2U} + \alpha_U b_{2m2U} + \frac{b_{1m2U}}{b_{2mL}} \right)(T-t),$$

$$A_{2m22,t} = 2\alpha_U \left(\alpha_U b_{2m1U} + \frac{b_{1m1U}}{b_{2mL}} \right)\left(b_{1m1U} + \alpha_U b_{2mU} b_{2m1U} \right)(T-t),$$

$$A_{2m31,t} = \left(\omega_U b_{2m3U} + \alpha_U b_{2m3U} + \frac{b_{2m2U} b_{1m1U} + b_{2m1U} b_{1m2U,t+i}}{b_{2mL}^2} + \frac{b_{1m3U}}{b_{2mL}} \right)(T-t),$$

$$A_{2m32,t} = \begin{bmatrix} 4\alpha_U \left(\alpha_U b_{2m2U} + \dfrac{b_{1m2U}}{b_{2mL}} \right)\left(b_{1m1U} + \alpha_U b_{2mU} b_{2m1U} \right) \\ +2\alpha_U \left(\alpha_U b_{2m1U} + \dfrac{b_{1m1U}}{b_{2mL}} \right)\left(b_{1m2U} + \dfrac{b_{1m1U}^2}{\alpha_L b_{2mL}^2} + \alpha_U b_{2m1U}^2 + \alpha_U b_{2mU} b_{2m2U} \right) \end{bmatrix}(T-t), \text{ and}$$

$$A_{2m33,t} = 8\alpha_U^2 \left(\alpha_U b_{2m1U} + \frac{b_{1m1U}}{b_{2mL}} \right)\left(b_{1m1U} + \alpha_U b_{2mU} b_{2m1U} \right)^2 (T-t).$$

Assume that the state variable as a function of observed asset price, $S_t(\theta) = f^{-1}(C_t;\theta)$, satisfies $S_{Lt} \le S_t(\theta) \le S_{Ut}$ and $S_{it}(\theta) \le S_{iUt}$, where $S_{it}(\theta) = \dfrac{\partial^i S_t(\theta)}{\partial \beta^i}$, $i = 1,2,3$, and S_{Lt}, S_{Ut}, and S_{iUt} don't depend on these parameters.

Lemma 3.12. *With $N(\theta_0)$ defined in (3.29), then for any t and $m = 0,1$,*

$$\sup_{\theta \in N(\theta_0)} X_{1m,t}(\phi,\theta) \le -\frac{\omega_L}{2}\phi^2, \qquad (3.49)$$

$$\sup_{\theta \in N(\theta_0)} X_{1m1,t}(\phi,\theta) \le \frac{\phi^2}{2} X_{1m1,t}, \qquad (3.50)$$

$$\sup_{\theta\in N(\theta_0)} X_{1m2,t}(\phi,\theta) \le \frac{\phi^2}{2}\left(X_{1m2,t} + \phi^2 A_{1m22,t}\right), \tag{3.51}$$

$$\sup_{\theta\in N(\theta_0)} X_{1m3,t}(\phi,\theta) \le \frac{\phi^2}{2}\left(X_{1m3,t} + \phi^2 A_{1m32,t} + \phi^4 A_{1m33,t}\right), \tag{3.52}$$

$$\sup_{\theta\in N(\theta_0)} X_{3m,t}(\phi,\theta) \le \phi X_{3m,t} + 2c\pi(T-t), \tag{3.53}$$

$$\sup_{\theta\in N(\theta_0)} X_{3m1,t}(\phi,\theta) \le \frac{\phi}{2} X_{3m1,t}, \tag{3.54}$$

$$\sup_{\theta\in N(\theta_0)} X_{3m2,t}(\phi,\theta) \le \frac{\phi}{2}\left(X_{3m2,t} + \phi^2 A_{2m22,t}\right), \tag{3.55}$$

and

$$\sup_{\theta\in N(\theta_0)} X_{3m3,t}(\phi,\theta) \le \frac{\phi}{2}\left(X_{3m3,t} + \phi^2 A_{2m32,t} + \phi^4 A_{2m33,t}\right), \tag{3.56}$$

where c *is a positive integer,*

$X_{1m1,t} = A_{1m1,t} + b_{1m1U} H_{t+1} + b_{1mU} H_{1t+1}$,

$X_{1m2,t} = A_{1m21,t} + b_{1m2U} H_{t+1} + 2b_{1m1U} H_{1t+1} + b_{1mU} H_{2t+1}$,

$X_{1m3,t} = A_{1m31,t} + b_{1m3U} H_{t+1} + 3b_{1m2U} H_{1t+1} + 3b_{1m1U} H_{2t+1} + b_{1mU} H_{3t+1}$,

$$X_{3m,t} = \left(r + \frac{\omega_U b_{2mU}}{2}\right)(T-t) + \ln(S_{Ut}/K) + \frac{1}{2} b_{2mU} H_{t+1},$$

$$X_{3m1,t} = \frac{2S_{1Ut}}{S_{Lt}} + A_{2m1,t} + b_{2m1U} H_{t+1} + b_{2mU} H_{1t+1},$$

$$X_{3m2,t} = \frac{2S_{1Ut}^2}{S_{Lt}^2} + \frac{2S_{2Ut}}{S_{Ut}} + A_{2m21,t} + b_{2m2U} H_{t+1} + 2b_{2m1U} H_{1t+1} + b_{2mU} H_{2t+1}, \text{ and}$$

$$X_{3m3,t} = \frac{4S_{1Ut}^3}{S_{Lt}^3} + \frac{6S_{2Ut}S_{1Ut}}{S_{Lt}^2} + \frac{2S_{3Ut}}{S_{Lt}} + A_{2m31,t} + b_{2m3U} H_{t+1} + 3b_{2m2U} H_{1t+1} + 3b_{2m1U} H_{2t+1} + b_{2mU} H_{3t+1}.$$

Lemma 3.13. *With* $N(\theta_0)$ *defined in (3.29), then for any t and* $k \geq 0$,

$$\sup_{\theta \in N(\theta_0)} \int_0^\infty \frac{e^{X_{1m,t}(\phi,\theta)}}{\phi} \sin\left(X_{3m,t}(\phi,\theta)\right) d\phi \leq \bar{X}_{3m,t}, \tag{3.57}$$

$$\sup_{\theta \in N(\theta_0)} \int_0^\infty \phi^k e^{X_{1m,t}(\phi,\theta)} \sin\left(X_{3m,t}(\phi,\theta)\right) d\phi \leq c_k, \tag{3.58}$$

and

$$\sup_{\theta \in N(\theta_0)} \int_0^\infty \phi^k e^{X_{1m,t}(\phi,\theta)} \cos\left(X_{3m,t}(\phi,\theta)\right) d\phi \leq c_k, \tag{3.59}$$

where $b = \frac{\omega_L}{2}$, $c_k = \int_0^\infty \phi^k e^{-b\phi^2} d\phi < \infty$, *and* $\bar{X}_{3m,t} = \int_0^1 \frac{\sin\left(\phi X_{3m,t}\right)}{\phi} d\phi + c_0$.

The next lemma shows that the individual terms of the third derivative $\left(\partial^3 L_T^{OT} / \partial \beta^3\right)(\theta)$ in (3.34) are uniformly bounded in the neighborhood $N(\theta_0)$.

Lemma 3.14. *With* $N(\theta_0)$ *defined in (3.29), then for any t and* $i = 1,2,3$,

$$\sup_{\theta \in N(\theta_0)} \frac{1}{J_t(\theta)} \leq \frac{1}{J_t}, \tag{3.60}$$

$$\sup_{\theta \in N(\theta_0)} Y_t(\theta) \leq Y_t, \tag{3.61}$$

$$\sup_{\theta \in N(\theta_0)} J_{it}(\theta) \leq J_{it}, \tag{3.62}$$

and

$$\sup_{\theta \in N(\theta_0)} Y_{it}(\theta) \leq Y_{it}, \tag{3.63}$$

where $J_{it}(\theta) \equiv \dfrac{\partial^i J_t(\theta)}{\partial \beta^i}$, $Y_{it}(\theta) \equiv \dfrac{\partial^i Y_t(\theta)}{\partial \beta^i}$, $Y_t = \ln\left(\dfrac{S_{Ut}}{S_{Lt-1}}\right) + r + \lambda H_t$,

$$Y_{1t} = \frac{S_{1Ut}}{S_{Lt}} + \frac{S_{1Ut-1}}{S_{Lt-1}} + \lambda H_{1t},\ Y_{2t} = \frac{S_{1Ut}^2}{S_{Lt}^2} + \frac{S_{2Ut}}{S_{Lt}} + \frac{S_{1Ut-1}^2}{S_{Lt-1}^2} + \frac{S_{2Ut-1}}{S_{Lt-1}} + \lambda H_{2t},$$

$$Y_{3t} = \frac{2S_{1Ut}^3}{S_{Lt}^3} + \frac{3S_{2Ut}S_{1Ut}}{S_{Lt}^2} + \frac{S_{3Ut}}{S_{Lt}} + \frac{2S_{1Ut-1}^3}{S_{Lt-1}^3} + \frac{3S_{2Ut-1}S_{1Ut-1}}{S_{Lt-1}^2} + \frac{S_{3Ut-1}}{S_{Lt-1}} + \lambda H_{3t},$$

$$J_t = \frac{1}{2} - \frac{1}{\pi}\left[\bar{X}_{31,t} + c_0\left(1 + e^{-r(T-t)}\frac{K}{S_{Lt}}\right)\right],$$

$$J_{1t} = \frac{1}{2\pi}\left[\left(c_1 + c_2\right)X_{111,t} + \left(c_0 + c_1\right)X_{311,t}\right] + e^{-r(T-t)}\frac{K}{S_{Lt}}\frac{1}{2\pi}\left(\frac{2S_{1Ut}}{S_{Lt}}c_0 + c_2 X_{101,t} + c_1 X_{301,t}\right),$$

$$J_{2t} = \frac{1}{\pi}\left[\begin{array}{l} \dfrac{c_3+c_4}{4}X_{111,t}^2 + \dfrac{c_1+c_2}{4}X_{311,t}^2 + \dfrac{c_1+c_2}{2}X_{112,t} + \dfrac{c_3+c_4}{2}A_{1122,t} + \dfrac{c_2+c_3}{2}X_{311,t}X_{111,t} \\ + \dfrac{c_0+c_1}{2}X_{312,t} + \dfrac{c_2+c_3}{2}A_{2122,t} \end{array}\right]$$
$$+ e^{-r(T-t)}\frac{K}{\pi S_{Lt}}\left[\begin{array}{l} c_2\dfrac{S_{1Ut}}{S_{Lt}}X_{101,t} + 2c_0\dfrac{S_{1Ut}^2}{S_{Lt}} + c_0\dfrac{S_{2Ut}}{S_{Lt}} + c_1\dfrac{S_{1Ut}}{S_{Lt}}X_{301,t} + \dfrac{c_4}{4}X_{101,t}^2 + \dfrac{c_2}{2}X_{102,t} \\ + \dfrac{c_4}{2}A_{1022,t} + \dfrac{c_1}{2}X_{301,t} + \dfrac{c_3}{2}X_{101,t}X_{301,t} + \dfrac{c_1}{2}X_{302,t} + \dfrac{c_3}{2}A_{2022,t} \end{array}\right],$$

and

$$
J_{3t}=\frac{1}{\pi}\left[\begin{array}{l}
\frac{c_5+c_6}{8}X_{111,t}^3+\frac{3(c_3+c_4)}{8}X_{311,t}^2X_{111,t}+\frac{3(c_3+c_4)}{4}X_{112,t}X_{111,t}+\frac{3(c_1+c_2)}{4}X_{312,t}X_{211,t}\\
+\frac{c_1+c_2}{2}X_{113,t}+\frac{c_0+c_1}{2}X_{313,t}+\frac{3(c_2+c_3)}{4}X_{312,t}X_{111,t}+\frac{3(c_2+c_3)}{4}X_{112,t}X_{311,t}\\
+\frac{3(c_4+c_5)}{8}X_{311,t}X_{111,t}^2+\frac{c_2+c_3}{8}X_{311,t}^3
\end{array}\right]
$$

$$
+e^{-r(T-t)}\frac{K}{\pi S_{Lt}}\left[\begin{array}{l}
\frac{S_{3Ut}}{S_{Lt}}+c_2\frac{S_{1Ut}^2}{S_{Lt}}X_{101,t}+4\frac{S_{1Ut}S_{2Ut}}{S_{Lt}}+\frac{3c_2}{2}\frac{S_{1Ut}}{S_{Lt}}X_{102,t}+\frac{3c_2}{2}\frac{S_{2Ut}}{S_{Lt}}X_{101,t}+\frac{3c_2}{4}\frac{S_{1Ut}}{S_{Lt}}X_{301,t}^2\\
+\frac{3c_4}{4}\frac{S_{1Ut}}{S_{Lt}}X_{101,t}^2+2c_2\frac{S_{1Ut}^2}{S_{Lt}^2}X_{101,t}+4\frac{S_{1Ut}^3}{S_{Lt}^2}+2\frac{S_{1Ut}S_{2Ut}}{S_{Lt}^2}+\frac{3c_3}{2}\frac{S_{1Ut}}{S_{Lt}}X_{101,t}X_{301,t}\\
+\frac{3c_1}{2}\frac{S_{2Ut}}{S_{Lt}}X_{301,t}+c_2\frac{S_{1Ut}}{S_{Lt}}X_{102,t}+\frac{c_1}{2}\frac{S_{1Ut}}{S_{Lt}}X_{302,t}+c_1\frac{S_{1Ut}^2}{S_{Lt}}X_{301,t}+2c_1\frac{S_{1Ut}^2}{S_{Lt}^2}X_{301,t}\\
+\frac{c_3}{8}X_{301,t}^3+\frac{c_1}{2}X_{303,t}+\frac{3c_3}{4}X_{102,t}X_{301,t}+\frac{3c_3}{4}X_{302,t}X_{101,t}+\frac{3c_5}{8}X_{101,t}^2X_{301,t}\\
+\frac{3c_4}{8}X_{101,t}X_{301,t}^2+\frac{3c_4}{4}X_{102,t}X_{101,t}+\frac{3c_2}{4}X_{302,t}X_{301,t}+\frac{c_6}{8}X_{101,t}^3+\frac{c_2}{2}X_{103,t}
\end{array}\right].
$$

Proof of Theorem 3.2. From Lemma 3.7, we have that

$$
\begin{bmatrix}0\\0\\0\\0\end{bmatrix}=\frac{\partial L_T^{OT}(\hat{\theta}_T^{OT})}{\partial\theta}\approx\frac{\partial L_T^{OT}(\theta_0)}{\partial\theta}+\frac{\partial^2 L_T^{OT}(\theta_0)}{\partial\theta\partial\theta'}(\hat{\theta}_T^{OT}-\theta_0)
$$

and

$$
\hat{\theta}_T^{OT}-\theta_1\approx-\left(\frac{\partial^2 L_T^{OT}(\theta_0)}{\partial\theta\partial\theta'}\right)^{-1}\left(\frac{\partial L_T^{OT}(\theta_0)}{\partial\theta}+\frac{1}{T}\sum_{t=1}^{T}\frac{1}{J_t(\theta_0)}\frac{\partial J_t(\theta_0)}{\partial\theta}\right).
$$

Combining with Lemmas 3.5 and 3.6, we complete the proof of Theorem 3.2.

3.4.3. Asymptotic behavior for S+O+E

We seek to prove the following:

Theorem 3.3. *The estimator of S+O+E is* $\hat{\theta}_T^{S+O+E} = \underset{\theta\in\Theta}{\arg\max}\, L_T^{S+O+E}(\theta)$. *It is consistent such that* $\hat{\theta}_T^{S+O+E} \xrightarrow{p} \theta_0$ *and asymptotically Normal such that* $H_{S+O+E0}^{-1/2} F_{S+O+E0} T^{1/2}(\hat{\theta}_T^{S+O+E} - \theta_0) \overset{A}{\sim} N(0, I_4)$, *where* $F_{S+O+E0} = -E\left(\dfrac{\partial^2 L_T^{S+O+E}(\theta_0)}{\partial\theta\partial\theta'}\right)$ *and* $H_{S+O+E0} = E\left(T \dfrac{\partial L_T^{S+O+E}(\theta_0)}{\partial\theta} \dfrac{\partial L_T^{S+O+E}(\theta_0)}{\partial\theta'}\right)$.

The log-likelihood function for S+O+E is given by (3.20) and it follows that:

$$\begin{aligned}\frac{\partial L_T^{S+O+E}(\theta)}{\partial\theta} =& -\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t(\theta)}\left(1-\frac{1}{1-\rho^2}\frac{(y_t-\mu_t(\theta))^2}{h_t(\theta)}+\frac{\rho}{1-\rho^2}\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\frac{C_t-C_t^{HN}(\theta)}{\eta}\right)+\frac{2\lambda}{(1-\rho^2)h_t^{1/2}(\theta)}\left(\rho\frac{C_t-C_t^{HN}(\theta)}{\eta}-\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\right)\right]\frac{\partial h_t(\theta)}{\partial\theta} \\ &-\frac{1}{T}\sum_{t=1}^{T}\frac{1}{\eta(1-\rho^2)}\left(\rho\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}-\frac{C_t-C_t^{HN}(\theta)}{\eta}\right)\frac{\partial C_t^{HN}(\theta)}{\partial\theta},\end{aligned} \tag{3.64}$$

$$\begin{aligned}\frac{\partial^2 L_T^{S+O+E}(\theta)}{\partial\theta\partial\theta'} =& -\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t^2(\theta)}\left(-1+\frac{2}{1-\rho^2}\frac{(y_t-\mu_t(\theta))^2}{h_t(\theta)}-\frac{3\rho}{2(1-\rho^2)}\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\frac{C_t-C_t^{HN}(\theta)}{\eta}\right)+\frac{2\lambda}{(1-\rho^2)h_t^{3/2}(\theta)}\left(\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}-\rho\frac{C_t-C_t^{HN}(\theta)}{\eta}\right)+\frac{2\lambda^2}{(1-\rho^2)h_t(\theta)}\right]\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'} \\ &-\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t(\theta)}\left(1-\frac{1}{1-\rho^2}\frac{(y_t-\mu_t(\theta))^2}{h_t(\theta)}+\frac{\rho}{1-\rho^2}\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\frac{C_t-C_t^{HN}(\theta)}{\eta}\right)+\frac{2\lambda}{(1-\rho^2)h_t^{1/2}(\theta)}\left(\rho\frac{C_t-C_t^{HN}(\theta)}{\eta}-\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\right)\right]\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'} \\ &+\frac{1}{T}\sum_{t=1}^{T}\frac{\rho}{\eta(1-\rho^2)}\left(\frac{y_t-\mu_t(\theta)}{h_t^{3/2}(\theta)}+\frac{2\lambda}{h_t^{1/2}(\theta)}\right)\frac{\partial C_t^{HN}(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}-\frac{1}{T}\sum_{t=1}^{T}\frac{1}{\eta^2(1-\rho^2)}\frac{\partial C_t^{HN}(\theta)}{\partial\theta}\frac{\partial C_t^{HN}(\theta)}{\partial\theta'}-\frac{1}{T}\sum_{t=1}^{T}\frac{1}{\eta(1-\rho^2)}\left(\rho\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}-\frac{C_t-C_t^{HN}(\theta)}{\eta}\right)\frac{\partial^2 C_t^{HN}(\theta)}{\partial\theta\partial\theta'},\end{aligned} \tag{3.65}$$

$$\begin{aligned}\frac{\partial^3 L_T^{S+O+E}(\theta)}{\partial\theta\partial\theta'\partial\theta} =& -\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t^3(\theta)}\left(2-\frac{6}{1-\rho^2}\frac{(y_t-\mu_t(\theta))^2}{h_t(\theta)}+\frac{15\rho}{4(1-\rho^2)}\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\frac{C_t-C_t^{HN}(\theta)}{\eta}\right)+\frac{1}{(1-\rho^2)h_t^{5/2}(\theta)}\left(\frac{9\rho}{2}\frac{C_t-C_t^{HN}(\theta)}{\eta}-8\lambda\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\right)-\frac{4\lambda^2}{(1-\rho^2)h_t^2(\theta)}\right]\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\frac{\partial h_t(\theta)}{\partial\theta} \\ &-\frac{3}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t^2(\theta)}\left(-1+\frac{2}{1-\rho^2}\frac{(y_t-\mu_t(\theta))^2}{h_t(\theta)}-\frac{3\rho}{2(1-\rho^2)}\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\frac{C_t-C_t^{HN}(\theta)}{\eta}\right)+\frac{2\lambda}{(1-\rho^2)h_t^{3/2}(\theta)}\left(\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}-\rho\frac{C_t-C_t^{HN}(\theta)}{\eta}\right)+\frac{2\lambda^2}{(1-\rho^2)h_t(\theta)}\right]\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'}\frac{\partial h_t(\theta)}{\partial\theta} \\ &-\frac{3}{2T}\sum_{t=1}^{T}\frac{\rho}{\eta(1-\rho^2)}\left(\frac{3}{2}\frac{y_t-\mu_t(\theta)}{h_t^{5/2}(\theta)}+\frac{2\lambda}{h_t^{3/2}(\theta)}\right)\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\frac{\partial C_t^{HN}(\theta)}{\partial\theta}+\frac{3}{2T}\sum_{t=1}^{T}\frac{\rho}{\eta(1-\rho^2)}\left(\frac{y_t-\mu_t(\theta)}{h_t^{3/2}(\theta)}+\frac{2\lambda}{h_t^{1/2}(\theta)}\right)\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'}\frac{\partial C_t^{HN}(\theta)}{\partial\theta} \\ &-\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t(\theta)}\left(1-\frac{1}{1-\rho^2}\frac{(y_t-\mu_t(\theta))^2}{h_t(\theta)}+\frac{\rho}{1-\rho^2}\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\frac{C_t-C_t^{HN}(\theta)}{\eta}\right)+\frac{2\lambda}{(1-\rho^2)h_t^{1/2}(\theta)}\left(\rho\frac{C_t-C_t^{HN}(\theta)}{\eta}-\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\right)\right]\frac{\partial^3 h_t(\theta)}{\partial\theta\partial\theta'\partial\theta} \\ &+\frac{3}{2T}\sum_{t=1}^{T}\frac{\rho}{\eta(1-\rho^2)}\left(\frac{y_t-\mu_t(\theta)}{h_t^{3/2}(\theta)}+\frac{2\lambda}{h_t^{1/2}(\theta)}\right)\frac{\partial^2 C_t^{HN}(\theta)}{\partial\theta\partial\theta'}\frac{\partial h_t(\theta)}{\partial\theta}-\frac{3}{2T}\sum_{t=1}^{T}\frac{1}{\eta^2(1-\rho^2)}\frac{\partial^2 C_t^{HN}(\theta)}{\partial\theta\partial\theta'}\frac{\partial C_t^{HN}(\theta)}{\partial\theta}-\frac{1}{T}\sum_{t=1}^{T}\frac{1}{\eta(1-\rho^2)}\left(\rho\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}-\frac{C_t-C_t^{HN}(\theta)}{\eta}\right)\frac{\partial^3 C_t^{HN}(\theta)}{\partial\theta\partial\theta'\partial\theta}.\end{aligned} \tag{3.66}$$

Once again, consider the asymptotic Normality of the score and observed information matrix in (3.64) and (3.65), using Lemmas 3.15 and 3.16, respectively.

Lemma 3.15. *The score given by (3.64) evaluated at* $\theta = \theta_0$ *is asymptotically Gaussian,* $H_{S+O+E0}^{-1/2} T^{1/2} \dfrac{\partial L_T^{S+O+E}(\theta)}{\partial\theta} \xrightarrow{D} N(0, I_4)$, *where*

$$\begin{aligned}H_{S+O+E0} &= E\left(T\frac{\partial L_T^{S+O+E}(\theta_0)}{\partial\theta}\frac{\partial L_T^{S+O+E}(\theta_0)}{\partial\theta'}\right) \\ &= E\left[\left(\frac{2-\rho^2}{4(1-\rho^2)}\frac{1}{h_t^2(\theta_0)}+\frac{\lambda^2}{(1-\rho^2)h_t(\theta_0)}\right)\frac{\partial h_t(\theta_0)}{\partial\theta}\frac{\partial h_t(\theta_0)}{\partial\theta'}\right]-E\left[\frac{2\lambda\rho}{\eta(1-\rho^2)h_t^{1/2}(\theta_0)}\frac{\partial C_t^{HN}(\theta_0)}{\partial\theta}\frac{\partial h_t(\theta_0)}{\partial\theta'}\right]+\frac{1}{\eta^2(1-\rho^2)}E\left[\frac{\partial C_t^{HN}(\theta_0)}{\partial\theta}\frac{\partial C_t^{HN}(\theta_0)}{\partial\theta'}\right].\end{aligned}$$

Lemma 3.16. *The observed information matrix given by (3.65) evaluated at* $\theta = \theta_0$ *converges in probability to* $-F_{S+O+E0}^{-1} \frac{\partial^2 L_T^{S+O+E}(\theta)}{\partial\theta\partial\theta'} \xrightarrow{p} I_4$, *where* $F_{S+O+E0} = -E\left(\frac{\partial^2 L_T^{S+O+E}(\theta_0)}{\partial\theta\partial\theta'}\right)$.

Then, as before, we have that the third derivative is uniformly bounded as follows:

Lemma 3.17. *There exists* $N(\theta_0)$ *defined in (3.29), for all* $1 \le i, j, k \le 4$, *for which* $\sup_{\theta\in N(\theta_0)} \left|\frac{\partial^3 L_T^{S+O+E}(\theta)}{\partial\theta_i\partial\theta_j\partial\theta_k}\right| \le g(\omega_L, \omega_U, \alpha_L, \alpha_U, \beta_L, \beta_U, \gamma_L, \gamma_U, T) \xrightarrow{a.s.} M < \infty$ *as* $T \to \infty$ *where* M *is constant.*

Without loss of generality, consider the case $\theta_i = \theta_j = \theta_k = \beta$. To show that the individual terms of the third derivative $\left(\partial^3 L_T^{S+O+E} / \partial\beta^3\right)(\theta)$ in (3.66) are uniformly bounded in the neighborhood $N(\theta_0)$. First, it must prove that Lemma 3.18.

Lemma 3.18. *With* $N(\theta_0)$ *defined in (3.29), then for any t and* $m = 0,1$,

$$\sup_{\theta\in N(\theta_0)} X_{2m,t}(\phi,\theta) \le \phi X_{2m,t} + 2c\pi(T-t), \tag{3.67}$$

$$\sup_{\theta\in N(\theta_0)} X_{2m1,t}(\phi,\theta) \le \frac{\phi}{2} X_{2m1,t}, \tag{3.68}$$

$$\sup_{\theta\in N(\theta_0)} X_{2m2,t}(\phi,\theta) \le \frac{\phi}{2}\left(X_{2m2,t} + \phi^2 A_{2m22,t}\right), \tag{3.69}$$

and

$$\sup_{\theta\in N(\theta_0)} X_{2m3,t}(\phi,\theta) \le \frac{\phi}{2}\left(X_{2m3,t} + \phi^2 A_{2m32,t} + \phi^4 A_{2m33,t}\right), \tag{3.70}$$

where c *is a positive integer,*

$$X_{2m,t} = \left(r + \frac{\omega_U b_{2mU}}{2}\right)(T-t) + \ln(S_t / K) + \frac{1}{2} b_{2mU} H_{t+1},$$

$X_{2m1,t} = A_{2m1,t} + b_{2m1U} H_{t+1} + b_{2mU} H_{1t+1}$,

$X_{2m2,t} = A_{2m21,t} + b_{2m2U} H_{t+1} + 2b_{2m1U} H_{1t+1} + b_{2mU} H_{2t+1}$, *and*

$X_{2m3,t} = A_{2m31,t} + b_{2m3U} H_{t+1} + 3b_{2m2U} H_{1t+1} + 3b_{2m1U} H_{2t+1} + b_{2mU} H_{3t+1}$.

The next lemma shows that the individual terms of the third derivative $\left(\partial^3 L_T^{S+O+E} / \partial \beta^3\right)(\theta)$ in (3.66) are uniformly bounded in the neighborhood $N(\theta_0)$.

Lemma 3.19. *With* $N(\theta_0)$ *defined in (3.29), for all t and* $i = 1,2,3$,

$$\sup_{\theta \in N(\theta_0)} C_t^{HN}(\theta) \le S_t \left(\frac{1}{2} + \frac{1}{\pi} \bar{X}_{21t}\right) + e^{-r(T-t)} K \left(\frac{1}{2} + \frac{1}{\pi} \bar{X}_{20t}\right) \qquad (3.71)$$

and

$$\sup_{\theta \in N(\theta_0)} C_{it}^{HN}(\theta) \le S_t C_{i1t}^{HN} + e^{-r(T-t)} K C_{i0t}^{HN}, \qquad (3.72)$$

where $C_{it}^{HN}(\theta) = \dfrac{\partial^i C_t^{HN}(\theta)}{\partial \beta^i}$ *and*

for $m = 0,1$, $\bar{X}_{2m,t} = \int_0^1 \dfrac{\sin\left(X_{2m,t}\phi\right)}{\phi} d\phi + c_0$, $C_{1mt}^{HN} = \dfrac{1}{2}\left(c_1 X_{1m1,t} + c_0 X_{2m1,t}\right)$,

$$C_{2mt}^{HN} = \frac{c_3}{4} X_{1m1,t}^2 + \frac{c_1}{4} X_{2m1,t}^2 + \frac{c_1}{2} X_{1m21,t} + c_3 X_{1m22,t} + \frac{c_2}{2} X_{1m1,t} X_{2m1,t} + \frac{c_0}{2} X_{2m21,t} + c_2 X_{2m22,t},$$

and

$$C_{3mt}^{HN} = \begin{pmatrix} \frac{c_5}{8} X_{1m1,t}^3 + \frac{3c_3}{8} X_{1m1,t} X_{2m1,t}^2 + \frac{3c_3}{4} X_{1m1,t} X_{1m21,t} + \frac{3c_5}{2} X_{1m1,t} X_{1m22,t} \\ + \frac{3c_1}{4} X_{2m1,t} X_{2m21,t} + \frac{3c_3}{2} X_{2m1,t} X_{2m22,t} + \frac{c_1}{2} X_{1m31,t} + c_3 X_{1m32,t} + c_5 X_{1m33,t} \\ \frac{c_2}{8} X_{2m1,t}^3 + \frac{3c_4}{8} X_{2m1,t} X_{1m1,t}^2 + \frac{3c_2}{4} X_{1m1,t} X_{2m21,t} + \frac{3c_4}{2} X_{1m1,t} X_{2m22,t} \\ + \frac{3c_2}{4} X_{2m1,t} X_{1m21,t} + \frac{3c_4}{2} X_{2m1,t} X_{1m22,t} + \frac{c_0}{2} X_{2m31,t} + c_2 X_{2m32,t} + c_4 X_{2m33,t} \end{pmatrix}.$$

Proof of Theorem 3.3. From Lemma 3.17, we have that

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \frac{\partial L_T^{S+O+E}(\hat{\theta}_T^{S+O+E})}{\partial \theta} \approx \frac{\partial L_T^{S+O+E}(\theta_0)}{\partial \theta} + \frac{\partial^2 L_T^{S+O+E}(\theta_0)}{\partial \theta \partial \theta'}(\hat{\theta}_T^{S+O+E} - \theta_0)$$

and

$$\hat{\theta}_T^{S+O+E} - \theta_0 \approx -\left(\frac{\partial^2 L_T^{S+O+E}(\theta_0)}{\partial \theta \partial \theta'} \right)^{-1} \frac{\partial L_T^{S+O+E}(\theta_0)}{\partial \theta}.$$

Combining with Lemmas 3.15 and 3.16, we complete the proof of Theorem 3.3.

Chapter 4

Robustness Checking

In this chapter, we consider an option pricing in Duan (1995) for a robustness check.

4.1. Model setup and ST specification

First, we describe the general stock and options pricing models applied in this paper. Then, we derive QMLE and asymptotic results for the ST specification.

4.1.1. GARCH(1,1) stock and options pricing models

We adopt the generalized setup used by Duan (1995), which propose a class of GARCH models for the price of a European call option, where the data-generating process for the stock price *S* is:

$$y_t = \ln S_t - \ln S_{t-1} = r + \lambda h_t^{1/2} - \frac{1}{2} h_t + \varepsilon_t,\ \varepsilon_t = h_t^{1/2} z_t, \text{ under } P \text{ measure,} \tag{4.1}$$

$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}, \tag{4.2}$$

where ε_t has mean zero and conditional variance h_t under *P* measure, r is the constant one-period risk-free rate of return, and λ the constant unit risk premium. In words, the conditional variance is a linear function of the past squared disturbances and the past conditional variances. The process (4.2) remains stationary if $\alpha+\beta<1$. The GARCH process specified in (4.1) and (4.2) reduces to the standard homoskedastic lognormal process in the Black-Scholes model if $\alpha=0$ and $\beta=0$. This ensures that the Black-Scholes model is a special case. We may consider process (4.2) as running indefinitely or we may assume initial values y_0 and h_0, with the latter drawn from the stationary distribution applied by Bollerslev (1986), Nelson (1990), Bougerol and Picard (1992), and others. Let Ψ_t be the information set (σ-field) generated by $\{y_t, y_{t-1}, \ldots\}$ and let $\theta_0=(\omega_0,\alpha_0,\beta_0)'$ represent the true parameter vector. Assume that $\theta_0\in\Theta\subseteq\Re^3$ is in the interior of Θ, a compact, convex parameter space. Specifically, for any vector $(\omega,\alpha,\beta)\in\Theta$, $0<\omega_L\leq\omega\leq\omega_U$, $0<\alpha_L\leq\alpha\leq\alpha_U$, $0<\beta_L\leq\beta\leq\beta_U$ and $\alpha_U(1+\lambda^2)+\beta_U<1$. Assume also that $\{z_t\}_{t\in Z}$ is i.i.d., drawn from a symmetric, uni-modal density, bounded in a neighborhood of 0, with mean 0, and variance 1. In addition, assume that h_t is independent of $\{z_t, z_{t+1}, \ldots\}$.

The corresponding model under the locally risk-neutral valuation relationship, which is defined by Duan (1995), reads

$$y_t=\ln S_t-\ln S_{t-1}=r-\frac{1}{2}h_t+\varepsilon_t^Q,\ \varepsilon_t^Q=h_t^{1/2}z_t^Q,\ \text{under } Q \text{ measure} \tag{4.3}$$

$$h_t=\omega+\alpha\left(\varepsilon_{t-1}^Q-\lambda h_{t-1}^{1/2}\right)^2+\beta h_{t-1}, \tag{4.4}$$

where $\varepsilon_t^Q=\lambda h_t^{1/2}+\varepsilon_t$ and $z_t^Q=\lambda+z_t$. The conditional variance process under the risk-neutralized pricing measure, is an EGARCH process, was proposed by Nelson (1991).

The process (4.4) under Q measure remains stationary if $\alpha(1+\lambda^2)+\beta<1$. The GARCH(1,1) European call price is described as:

$$C_t^D(T,K,S_t,h_{t+1};\theta)=e^{-r(T-t)}E_t^Q\left[\max\left(S_T-K,0\right)\right], \tag{4.5}$$

$$S_T=S_t\exp\left[(T-t)r-\frac{1}{2}\sum_{s=t+1}^{T}h_s+\sum_{s=t+1}^{T}\varepsilon_s^Q\right], \tag{4.6}$$

where T is the maturity date and K is exercise price.

4.1.2. QMLEs and asymptotic results

We now turn our attention to estimating the parameters in the model. The base case ST uses only stock data. Specifically, h_t is the conditional variance of y_t with respect to Ψ_{t-1}. The estimation model utilizes (4.1) and (4.2), applying estimated parameter values $(\omega,\alpha,\beta)=\left(\theta_1,\theta_2,\theta_3\right)$. The error terms z_t are computed as $z_0=\dfrac{y_0-r-\lambda h_0^{1/2}+0.5h_0}{h_0^{1/2}}$, $z_1=\dfrac{y_1-r-\lambda h_1^{1/2}+0.5h_1}{h_1^{1/2}}$, …, where $\left\{y_t,\ t=0,\ldots,T\right\}$ are observed data. The process h_t is not observed but is constructed recursively using estimated parameter values, z_0, and an appropriate startup value, h_0, to be discussed in detail later.

QMLE is obtained by maximizing, conditional on h_0, as follows:

$$L_T^{ST}(y_0,\ldots,y_T,h_0;\theta)=L_T^{ST}(\theta)=\frac{-1}{2T}\sum_{t=1}^{T}\left(\ln\left(h_t(\theta)\right)+\frac{\left(y_t-\mu_t(\theta)\right)^2}{h_t(\theta)}\right). \tag{4.7}$$

where $\mu_t(\theta) = r + \lambda h_t^{1/2}(\theta) - \frac{1}{2} h_t(\theta)$. That is, $\hat{\theta}_T^{ST} = \arg\max_{\theta \in \Theta} L_T^{ST}(\theta)$. This estimator is consistent as $\hat{\theta}_T^{ST} \xrightarrow{p} \theta_0$. It is also asymptotically Normal as

$$H_{ST0}^{-1/2} F_{ST0} T^{1/2} (\hat{\theta}_T^{ST} - \theta_0) \overset{A}{\sim} N(0, I_3), \tag{4.8}$$

where $F_{ST0} = -E\left(\frac{\partial^2 L_T^{ST}(\theta_0)}{\partial\theta\partial\theta'} \right)$, $H_{ST0} = E\left(T \frac{\partial L_T^{ST}(\theta_0)}{\partial\theta} \frac{\partial L_T^{ST}(\theta_0)}{\partial\theta'} \right)$, and $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. A full proof can be found in the Appendix B as Theorem B.1.

In the interest of computational simplicity, assume that z_t is Normal so that $F_{ST0} = H_{ST0}$, though our general intuition remains the same under the more relaxed aforementioned specification for z_t. The asymptotic covariance matrix V_{ST} and asymptotic mean square errors MSE_{ST} are:

$$MSE_{ST}(\theta_0) = V_{ST}(\theta_0) = F_{ST0}^{-1} H_{ST0} F_{ST0}^{-1} = \frac{1}{F_{ST0}}, \tag{4.9}$$

where $F_{ST0} = H_{ST0} = E\left[\frac{2 + (\lambda - h_t^{1/2}(\theta_0))^2}{4 h_t^2(\theta_0)} \frac{\partial h_t(\theta_0)}{\partial\theta} \frac{\partial h_t(\theta_0)}{\partial\theta'} \right]$.

4.2. The S+O+E specification

We now turn our attention to a new specification that takes both stock and options data into account, but which allows for an error term in the options pricing formula. Then, we derive QMLE and asymptotic results for the S+O+E specification. Numerical results confirm these characteristics.

4.2.1. QMLEs and asymptotic results

For this method, we allow that $C_t = C_t^D + e_t$ where $e_t = \eta u_t$ and $t = 1,\ldots,T$. Assume that $u_t \overset{i.i.d.}{\sim} N(0,1)$ and $\eta > 0$. For the purpose of calculating the QMLE, let us assume z_t and u_t, with $correlation(z_t, u_t) = \rho$ where $-1 < \rho < 1$, have a bi-Normally distribution, that is, $\begin{pmatrix} z_t \\ u_t \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$. Let $G_t = [S_t, C_t]'$ be a vector of observable stock and options prices, respectively. Then, the joint density is as follows:

$$P(G;\theta) = P(S,C;\theta) = P(C \mid S;\theta)P(S;\theta) = \prod_{t=1}^{T} P(C_t \mid S_t;\theta)P(S_t \mid S_{t-1};\theta)$$
$$= \prod_{t=1}^{T} \frac{1}{2\pi\sqrt{1-\rho^2}\sqrt{h_t(\theta)}\eta} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(y_t-\mu_t(\theta))^2}{h_t(\theta)} - 2\rho\frac{y_t-\mu_t(\theta)}{\sqrt{h_t(\theta)}}\frac{C_t-C_t^D(\theta)}{\eta} + \frac{(C_t-C_t^D(\theta))^2}{\eta^2}\right)\right]. \quad (4.10)$$

The log-likelihood function for discrete data on the asset price vector G_t sampled at dates $0 \le t \le T$ has the form:

$$L_T^{S+O+E}(\theta) = \ln\left(P(G;\theta)\right)$$

$$= \frac{-1}{2T}\sum_{t=1}^{T}\left[2\ln\left(2\pi\eta\sqrt{1-\rho^2}\right) + \ln(h_t(\theta)) + \frac{1}{(1-\rho^2)}\left(\frac{(y_t-\mu_t(\theta))^2}{h_t(\theta)} - 2\rho\frac{y_t-\mu_t(\theta)}{\sqrt{h_t(\theta)}}\frac{C_t-C_t^D(\theta)}{\eta} + \frac{(C_t-C_t^D(\theta))^2}{\eta^2}\right)\right]. \quad (4.11)$$

And, the QMLE for $\hat{\theta}_T^{S+O+E} = \arg\max_{\theta\in\Theta} L_T^{S+O+E}(\theta)$.

This estimator is also consistent as $\hat{\theta}_T^{S+O+E} \xrightarrow{p} \theta_0$, and is asymptotically Normally distributed as $H_{S+O+E0}^{-1/2} F_{S+O+E0} T^{1/2}(\hat{\theta}_T^{S+O+E} - \theta_0) \overset{A}{\sim} N(0, I_3)$, where $F_{S+O+E0} = -E\left(\frac{\partial^2 L_T^{S+O+E}(\theta_0)}{\partial\theta\partial\theta'}\right)$ and $H_{S+O+E0} = E\left(T\frac{\partial L_T^{S+O+E}(\theta_0)}{\partial\theta}\frac{\partial L_T^{S+O+E}(\theta_0)}{\partial\theta'}\right)$. A full proof appears as Theorem B.2 in the Appendix B. Again, assume that z_t is Normal so that

$F_{S+O+E0} = H_{S+O+E0}$. The asymptotic covariance matrix V_{S+O+E} and asymptotic mean square error MSE_{S+O+E} for the S+O+E case are:

$$\begin{aligned} MSE_{S+O+E}(\theta_0) &= V_{S+O+E}(\theta_0) \\ &= F^{-1}_{S+O+E0} H_{S+O+E0} F^{-1}_{S+O+E0} = \frac{1}{F_{S+O+E0}} = \frac{1}{F_{ST0} + M_{S+O+E}(\theta_0)}, \end{aligned} \tag{4.12}$$

where

$$\begin{aligned} M_{S+O+E}(\theta_0) &= F_{S+O+E0} - F_{ST0} \\ &= \frac{\rho^2}{4(1-\rho^2)} E\left[\frac{1+(\lambda - h_t^{1/2}(\theta_0))^2}{h_t^2(\theta_0)} \frac{\partial h_t(\theta_0)}{\partial \theta} \frac{\partial h_t(\theta_0)}{\partial \theta'}\right] - \frac{\rho}{\eta(1-\rho^2)} E\left[\frac{\lambda - h_t^{1/2}(\theta_0)}{h_t(\theta_0)} \frac{\partial C_t^D(\theta_0)}{\partial \theta} \frac{\partial h_t(\theta_0)}{\partial \theta'}\right] \\ &\quad + \frac{1}{\eta^2(1-\rho^2)} E\left[\frac{\partial C_t^D(\theta_0)}{\partial \theta} \frac{\partial C_t^D(\theta_0)}{\partial \theta'}\right]. \end{aligned}$$

These results follow from Lemmas B.1, B.2, B.5, and B.6 in the Appendix B. We can now investigate M_{S+O+E} where again the more positive, the more efficient the estimator.

When α_0 and β_0 are known and ω is unknown:

$$\begin{aligned} M_{S+O+E}(\omega_0 \mid \alpha_0, \beta_0) &= \frac{\rho^2}{4(1-\rho^2)} E\left[\frac{1+(\lambda - h_t^{1/2}(\theta_0))^2}{h_t^2(\theta_0)} \left(\frac{\partial h_t(\omega_0)}{\partial \omega}\right)^2\right] + \frac{1}{\eta^2(1-\rho^2)} E\left[\left(\frac{\partial C_t^D(\omega_0)}{\partial \omega}\right)^2\right] \\ &\quad - \frac{\rho}{\eta(1-\rho^2)} E\left[\frac{\lambda - h_t^{1/2}(\theta_0)}{h_t(\theta_0)} \frac{\partial C_t^D(\omega_0)}{\partial \omega} \frac{\partial h_t(\omega_0)}{\partial \omega}\right]. \end{aligned} \tag{4.13}$$

Note that M_{S+O+E} may be positive or negative. In $\rho = 0$ case, we easy to see that M_{S+O+E} is positive definite from (4.13). In $\rho \neq 0$ case, we don't compare these values since the M_{S+O+E} depend on true parameters. Thus, we will calculate these values by numerical simulation in Section 4.2.2. As illustrated later, M_{S+O+E} is in fact generally positive. When ω_0 and β_0 are known and α is not:

$$M_{S+O+E}(\alpha_0 \mid \omega_0, \beta_0) = \frac{\rho^2}{4(1-\rho^2)} E\left[\frac{1+(\lambda - h_t^{1/2}(\theta_0))^2}{h_t^2(\theta_0)} \left(\frac{\partial h_t(\alpha_0)}{\partial \alpha} \right)^2 \right] + \frac{1}{\eta^2(1-\rho^2)} E\left[\left(\frac{\partial C_t^D(\alpha_0)}{\partial \alpha} \right)^2 \right] \\ - \frac{\rho}{\eta(1-\rho^2)} E\left[\frac{\lambda - h_t^{1/2}(\theta_0)}{h_t(\theta_0)} \frac{\partial C_t^D(\alpha_0)}{\partial \alpha} \frac{\partial h_t(\alpha_0)}{\partial \alpha} \right]. \tag{4.14}$$

In $\rho = 0$ case, M_{S+O+E} is positive definite, and in $\rho \neq 0$ case, as demonstrated later, M_{S+O+E} is positive definite. Similarly, when ω_0 and α_0 are known and β is not:

$$M_{S+O+E}(\beta_0 \mid \omega_0, \alpha_0) = \frac{\rho^2}{4(1-\rho^2)} E\left[\frac{1+(\lambda - h_t^{1/2}(\theta_0))^2}{h_t^2(\theta_0)} \left(\frac{\partial h_t(\beta_0)}{\partial \beta} \right)^2 \right] + \frac{1}{\eta^2(1-\rho^2)} E\left[\left(\frac{\partial C_t^D(\beta_0)}{\partial \beta} \right)^2 \right] \\ - \frac{\rho}{\eta(1-\rho^2)} E\left[\frac{\lambda - h_t^{1/2}(\theta_0)}{h_t(\theta_0)} \frac{\partial C_t^D(\beta_0)}{\partial \beta} \frac{\partial h_t(\beta_0)}{\partial \beta} \right]. \tag{4.15}$$

Again, as for the case where α is unknown, we show that M_{S+O+E} is always positive.

4.2.2. Numerical computation for asymptotic mean square errors

We now generate numerical results to test and illustrate these asymptotic findings. We presume that parameter true values are (λ, ω_0, α_0, β_0) = (0.0116, 9.228×10^{-7}, 0.068, 0.9248) and the risk-free rate is fixed at 5%. These parameters are estimated using S&P 500 daily index data from January 1996 to the end of 2007. We use these parameters to run our tests.

First, we investigate and calculate analytically the efficiency of the estimator, graphs of M_{S+O+E}, in varied ρ. Remember that, the more positive this value, the more efficient the estimator. Graphs in Figure 5 show that the value of M_{S+O+E} in the true parameter values and ρ from -0.9 to 0.9. Then, in each panel, one variable is unknown while the other two are treated as known. Specifically, in Panel A, ω is varied, in B α, and in C β.

[Insert Figure 5 here]

Looking at all graphs, M_{S+O+E} is positive for all three parameters. Specifically, M_{S+O+E} is always minimum in $\rho = 0$ and increases as the absolute value of ρ increases.

Graphs of M_{S+O+E} are shown in Figure 6, in the area surrounding true parameter values. We only consider $\rho = 0$ case since M_{S+O+E} is minimum in this case. Since the orders of magnitude for the three parameters are quite different, we graph the M_{S+O+E} of ω on the left and that of α and β on the right. True parameters are circled in each graph. Then, in each panel, one variable is varied while the other two are treated as known. Specifically, in Panel A, ω is varied, in B α, and in C β.

[Insert Figure 6 here]

Looking at Panel A, M_{S+O+E} is always positive for all three parameters in the area surrounding true parameters. While M_{S+O+E} for ω decreases with ω locally in the region around the true parameter values, those for α and β are increasing, with M_{S+O+E} for β always higher than that of α. In Panel B, M_{S+O+E} for ω, α, and β increase with α locally in the region around the true parameter values. Corresponding graphs Panel C are similar to each other in shape, though their x-axes differ. However, M_{S+O+E} is always positive in each case.

In summary, M_{S+O+E} is always positive for all parameters in the area surrounding true parameters. We conclude that S+O+E generates more efficient estimates than ST in all cases. This method generates asymptotically unbiased estimates, making it the most desirable data specification of the two. We also conclude that the use of option prices can lead to very accurate estimates, even in long samples.

4.2.3. Numerical findings and direct comparisons in finite sample studies

We generate parameter estimates using a Monte Carlo method, comparing the bias and variance characteristics of the two data specifications in finite sample studies. Specifically, we simulate 30 days of stock and/or options prices and then run 1,000 iterations over each period. We again presume that parameter true values are (λ, ω_0, α_0, β_0) = (0.0116, 9.228×10^{-7}, 0.068, 0.9248). For the S+O+E case, we additionally assume that η = 1 and that ρ = 0. As a robustness check, we re-run these tests for a variety of calibrations of ρ and η and find no qualitative differences.

We estimate our four parameters for out-of-the-money (S_0/K = 0.9), at-the-money (S_0/K = 1.0), and in-the-money (S_0/K = 1.1) cases. Results for 30 days are presented in Table 3 where we report the absolute value of estimate bias (estimate less true value), standard deviation of the estimate (SD), and mean squared errors (MSE).

[Insert Table 3 here]

When ω is unknown, we find that S+O+E arrives at estimates within $2.816\text{x}10^{-7}$ of the true value. In contrast, estimates using ST present biases on the order of 1.21 to 2.03 times higher. Note that ST has the stronger bias regardless of the moneyness of the options. Standard deviations for S+O+E are lower than those for ST. With regard to MSE, results are even more staggering with S+O+E exhibiting the lower values and ST generating errors that are about 3 times higher. Note that ST seems to perform particularly poorly when options are out-of-the-money.

When estimating α, we similarly find that estimation bias is lower for S+O+E, with ST again about 6 times higher for most variables. Standard deviations are also lower for S+O+E than for ST. MSE exhibits the same behavior as before, with S+O+E exhibiting by the lower values and ST generating errors that are about 8 times higher. Once again, ST presents particularly poor results with options out-of-the-money.

Estimation results for β are consistent with those of the other two parameters. S+O+E exhibits the smaller bias, the lower standard deviation, and the lower MSE of the two methods. Again, ST exhibits by far the worst performance along all three metrics and again particularly poor when options are out-of-the-money. The prevalence of biased estimates is striking for ST and is a particular strength for S+O+E.

In summary, we conclude that the use of option prices can lead to very accurate estimates, even in short samples. This result is consistent with that of Eraker (2004).

4.3. Risk management implications

In this section we document that errors and bias in estimation may have substantial repercussions as relates to risk management benchmarks and practices. To illustrate, we obtain daily stock and options data for the period from January 2007 to the end of 2007. For stock prices (S_t), we use the S&P 500 index, and for options data (C_t), we use the price of a short-maturity at-the-money call options where the price is measured as the midpoint of the last reported bid-ask spread. We assume that $h_0 = \omega / (1 - \alpha - \beta)$, and for ease of interpretation, let the risk-free rate equal 0%. Applying the 12 months of stock and options data, we find key parameters to be (λ, ω_0, α_0, β_0) = (0.0002, 8.248×10^{-7}, 0.07275, 0.92305) for ST and (λ, ω_0, α_0, β_0) = (0.00012, 8.1633×10^{-7}, 0.07379, 0.92321) for S+O+E. As demonstrated in the following discussion, while these parameters may not appear to differ greatly, the resulting risk management implications are quite significant.

We then use these parameter estimates to calculate options deltas and gammas, measuring options stock price sensitivity and convexity, respectively. We calculate deltas and gammas for both the Black-Scholes and GARCH options pricing models for a variety of levels of moneyness, ranging from 0.8 (out-of-the-money) to 1.2 (in-the-money), and times to maturity, ranging from 30 days to 180 days. Results are provided in Table 4.

[Insert Table 4 here]

First consider delta. Black-Scholes values from range about 4% higher for ST than for S+O+E when options are in the money to as much as 73% lower when they are out of the money. The difference tends to be negative when options are out of the money and positive when they are in the money. The trend is the same for GARCH deltas, although the magnitude ranges from 3.63% higher when options are in the money to 55.82% lower when they are out of the money. As a result, replicating and hedging portfolios will be significantly different based on the estimation method used, regardless of whether the agent applies a Black-Scholes or GARCH options pricing model.

For gammas, there does not appear be a pattern in the difference related to the moneyness of the options. However, the magnitude of the differences ranges from -91.94% to 43.85%, indicating that gammas are more strongly impacted by the method of estimation used than deltas. As a result, the updating dynamics dictated to maintain hedges will be substantially different depending on the data specification employed. The magnitude of these differences suggests that managers and investors would do well to keep this in mind when calculating risk management metrics for options.

Chapter 5

Conclusion

The GARCH class of models has been shown to be empirically superior to other models but is a restrictive model which can be overloaded when applying options data only. We demonstrate the overload of the OT specification both theoretically and numerically, shedding new light on the effectiveness of different methods of estimation and the corresponding asymptotic behavior of estimators. For all reasonable true parameter values, application of options data only generates asymptotically biased and relatively inefficient estimates. However, application of stock and options data without an error term doesn't also generate more efficient estimates since S+O specification is similar to ST specification. As a result, we develop here a method of estimation that applies an error term to the options pricing formula, thereby delivering the additional slack GARCH models require when applying the dual dataset. We show that under this new specification, estimates as unbiased and maximally efficient, i.e., more data is in fact better.

In addition, we demonstrate that different estimation methods will result in significantly different risk metrics of options, regardless of whether a GARCH or Black-Scholes model is used. While these effects do not appear to be systematically related to maturity or moneyness, they can be substantial in magnitude, especially as regards risk management dynamics. This highlights the economic importance of developing an unbiased and efficient estimation method, and financial managers would do well to consider these effects when implementing hedging practices and trades.

The GARCH class of models is conditionally deterministic and, as a result, restrictive. Applying too much data to this specification induces helpless if sufficient slack is not introduced. SV models do not have this quality and have a natural mechanism for slack. Unlike SV model, in order to develop a better method under GARCH models, it seems to be necessary for allowing an error term in the options pricing formula for additional slack.

Appendix A

Proof of Lemma 3.1. Evaluated at $\theta = \theta_0$ the form is given by

$$\frac{\partial L_T^{ST}(\theta)}{\partial \theta} = -\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t(\theta)}\left(1-z_t^2\right) - \frac{2\lambda z_t}{h_t^{1/2}(\theta)}\right]\frac{\partial h_t(\theta)}{\partial \theta} = \frac{1}{T}\sum_{t=1}^{T}\nu_t$$

such that $E\left(\nu_t \mid F_{t-1}\right) = 0$, where $F_t = \sigma\left(z_t, z_{t-1}, \ldots\right)$. Applying the central limit theorem for martingale differences in Brown (1971), consider first

$$\frac{1}{T}\sum_{t=1}^{T}E\left(\nu_t^2 \mid F_{t-1}\right) = \frac{1}{T}\sum_{t=1}^{T}\left[\frac{1}{2h_t^2(\theta)} + \frac{\lambda^2}{h_t(\theta)}\right]\frac{\partial h_t(\theta)}{\partial \theta}\frac{\partial h_t(\theta)}{\partial \theta'} \to E\left[\left(\frac{1}{2h_t^2(\theta)} + \frac{\lambda^2}{h_t(\theta)}\right)\frac{\partial h_t(\theta)}{\partial \theta}\frac{\partial h_t(\theta)}{\partial \theta'}\right]$$

imply that $H_{ST0}^{-1}\frac{1}{T}\sum_{t=1}^{T}E\left(\nu_t^2 \mid F_{t-1}\right) \to I_4$ in probability as $T \to \infty$, using the ergodic theorem. Hence we complete the proof of Lemma 3.1.

Proof of Lemma 3.2. For $\theta = \theta_0$ the observed information is given by

$$\begin{aligned}\frac{\partial^2 L_T^{ST}(\theta)}{\partial \theta \partial \theta'} = &-\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t^2(\theta)}\left(-1+2z_t^2\right) + \frac{2\lambda z_t}{h_t^{3/2}(\theta)} + \frac{2\lambda^2}{h_t(\theta)}\right]\frac{\partial h_t(\theta)}{\partial \theta}\frac{\partial h_t(\theta)}{\partial \theta'} \\ &-\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t(\theta)}\left(1-z_t^2\right) - \frac{2\lambda z_t}{h_t^{1/2}(\theta)}\right]\frac{\partial^2 h_t(\theta)}{\partial \theta \partial \theta'}.\end{aligned}$$

The first term on the right-hand side converges by the ergodic theorem to $-E\left[\left(\frac{1}{2h_t^2(\theta)} + \frac{\lambda^2}{h_t(\theta)}\right)\frac{\partial h_t(\theta)}{\partial \theta}\frac{\partial h_t(\theta)}{\partial \theta'}\right]$; second term on the right-hand side converges in probability to zero. Hence we can complete the proof of Lemma 3.2.

Proof of Lemma 3.4. Since $h_t(\theta) \geq \omega$ and $|x| \leq x^2 + 1$, we have

$$h_t(\theta) \le \omega + 2\alpha + \left(\frac{\alpha}{\omega} + 2\alpha(\lambda+\gamma)^2\right)(y_{t-1}-r)^2 + \left(\alpha(\lambda+\gamma)^2 + \beta\right)h_{t-1}(\theta).$$

By simple recursion,

$$\begin{aligned} h_t(\theta) \le{}& \max\left\{\omega + 2\alpha, h_0\right\}\sum_{i=0}^{t}\left(\alpha(\lambda+\gamma)^2+\beta\right)^i \\ &+\left(\frac{\alpha}{\omega} + 2\alpha(\lambda+\gamma)^2\right)\sum_{i=1}^{t}\left(\alpha(\lambda+\gamma)^2+\beta\right)^{i-1}(y_{t-i}-r)^2 \end{aligned}$$

which implies (3.30).

Here, the first-, second-, and third-order derivatives of $h_t(\theta)$ are

$$h_{1t}(\theta) = h_{t-1}(\theta) + \left(\alpha(\lambda+\gamma)^2 + \beta - \frac{\alpha(y_{t-1}-r)^2}{h_{t-1}^2(\theta)}\right)h_{1t-1}(\theta),$$

$$h_{2t}(\theta) = 2\left(1 + \frac{\alpha(y_{t-1}-r)^2}{h_{t-1}^3(\theta)}h_{1t-1}(\theta)\right)h_{1t-1}(\theta) + \left(\alpha(\lambda+\gamma)^2 + \beta - \frac{\alpha(y_{t-1}-r)^2}{h_{t-1}^2(\theta)}\right)h_{2t-1}(\theta),$$

and

$$\begin{aligned} h_{3t}(\theta) ={}& -6\frac{\alpha(y_{t-1}-r)^2}{h_{t-1}^4(\theta)}h_{1t-1}^3(\theta) + 6\frac{\alpha(y_{t-1}-r)^2}{h_{t-1}^3(\theta)}h_{2t-1}(\theta)h_{1t-1}(\theta) + 3h_{2t-1}(\theta) \\ &+\left(\alpha(\lambda+\gamma)^2 + \beta - \frac{\alpha(y_{t-1}-r)^2}{h_{t-1}^2(\theta)}\right)h_{3t-1}(\theta). \end{aligned}$$

From $\frac{\alpha(y_{t-1}-r)^2}{h_{t-1}^2(\theta)} \ge 0$, h_0 is constant such that $h_{i0}=0$, $i=1,2,3$, and applying simple recursions,

$$h_{1t}(\theta) \le \sum_{i=1}^{t-1} h_{t-i}(\theta)\left(\alpha(\lambda+\gamma)^2+\beta\right)^{i-1},$$

$$h_{2t}(\theta) \le 2\sum_{i=1}^{t-1}\left(1 + \frac{\alpha(y_{t-i}-r)^2}{h_{t-i}^3(\theta)}h_{1t-i}(\theta)\right)h_{1t-i}(\theta)\left(\alpha(\lambda+\gamma)^2+\beta\right)^{i-1},$$

and

$$h_{3t}(\theta) \le 3\sum_{i=1}^{t-1}\left(2\frac{\alpha(y_{t-i}-r)^2}{h_{t-i}^4(\theta)} h_{1t-i}^3(\theta) + 2\frac{\alpha(y_{t-i}-r)^2}{h_{t-i}^3(\theta)} h_{2t-i}(\theta)h_{1t-i}(\theta) + h_{2t-i}(\theta)\right)\left(\alpha(\lambda+\gamma)^2+\beta\right)^{i-1}.$$

Then, (3.31) follows by applying (3.30).

Proof of Lemma 3.3. Without loss of generality, consider the case $\theta_i = \theta_j = \theta_k = \beta$. We show that the individual terms of the third derivative $\left(\partial^3 L_T^{ST} / \partial\beta^3\right)(\theta)$ in (3.28) are uniformly bounded in the neighborhood $N(\theta_0)$. Noting that by definition $y_t = r + \lambda h_t(\theta_0) + h_t^{1/2}(\theta_0) z_t$, the expression for $\left(\partial^3 L_T^{ST} / \partial\beta^3\right)(\theta)$ implies that

$$\left|\frac{\partial^3 L_T^{ST}(\theta)}{\partial\beta^3}\right| \le \frac{1}{T}\sum_{t=1}^{T} w_t(\theta),\text{ where}$$

$$\begin{aligned}
w_t(\theta) &\\
=& \left[\frac{1}{h_t^3(\theta)} + \frac{3\left[\lambda\left(h_t(\theta_0)+h_t(\theta)\right)+h_t^{1/2}(\theta_0)z_t\right]^2}{h_t^4(\theta)} + \frac{4\lambda\left[\lambda\left(h_t(\theta_0)+h_t(\theta)\right)+h_t^{1/2}(\theta_0)z_t\right]}{h_t^3(\theta)} + \frac{2\lambda^2}{h_t^2(\theta)}\right] h_{1t}^3(\theta) \\
&+ \left[\frac{3}{2h_t^2(\theta)} + \frac{3\left[\lambda\left(h_t(\theta_0)+h_t(\theta)\right)+h_t^{1/2}(\theta_0)z_t\right]^2}{h_t^3(\theta)} + \frac{3\lambda\left[\lambda\left(h_t(\theta_0)+h_t(\theta)\right)+h_t^{1/2}(\theta_0)z_t\right]}{h_t^2(\theta)} + \frac{3\lambda^2}{h_t(\theta)}\right] h_{2t}(\theta)h_{1t}(\theta) \\
&+ \left[\frac{1}{2h_t(\theta)} + \frac{1}{2}\frac{\left[\lambda\left(h_t(\theta_0)+h_t(\theta)\right)+h_t^{1/2}(\theta_0)z_t\right]^2}{h_t^2(\theta)} + \frac{\lambda\left[\lambda\left(h_t(\theta_0)+h_t(\theta)\right)+h_t^{1/2}(\theta_0)z_t\right]}{h_t(\theta)}\right] h_{3t}(\theta).
\end{aligned}$$

Lemma 3.4 suffices to show that there exists a neighborhood $N(\theta_0)$ for which

$$\sup_{\theta\in N(\theta_0)}\left|\frac{\partial^3 L_T^{ST}(\theta)}{\partial\beta^3}\right| \le \frac{1}{T}\sum_{t=1}^{T} w_t,$$

where w_t is stationary and has finite moment $Ew_t = M < \infty$ such that $\frac{1}{T}\sum_{t=1}^{T} w_t \xrightarrow{a.s.} M$ by the ergodic theorem, which ends the proof of Lemma 3.3.

Proof of Lemma 3.5. Evaluated at $\theta = \theta_0$ the form is given by

$$\frac{\partial L_T^{OT}(\theta)}{\partial\theta}+\frac{1}{T}\sum_{t=1}^{T}\frac{1}{J_t(\theta)}\frac{\partial J_t(\theta)}{\partial\theta}=-\frac{1}{T}\sum_{t=1}^{T}\left[\frac{1}{2h_t(\theta)}\left(1-z_t^2\right)\frac{\partial h_t(\theta)}{\partial\theta}+\frac{z_t}{h_t^{1/2}(\theta)}\frac{\partial Y_t(\theta)}{\partial\theta}\right]=\frac{1}{T}\sum_{t=1}^{T}\nu_t$$

such that $E\left(\nu_t \mid F_{t-1}\right)=0$, where $F_t=\sigma\left(z_t,z_{t-1},\ldots\right)$. Applying the central limit theorem for martingale differences in Brown (1971), consider first

$$\frac{1}{T}\sum_{t=1}^{T}E\left(\nu_t^2 \mid F_{t-1}\right)=\frac{1}{T}\sum_{t=1}^{T}\left[\frac{1}{2h_t^2(\theta)}\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}+\frac{1}{h_t(\theta)}\frac{\partial Y_t(\theta)}{\partial\theta}\frac{\partial Y_t(\theta)}{\partial\theta'}\right]$$
$$\to E\left[\frac{1}{2h_t^2(\theta)}\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\right]+E\left[\frac{1}{h_t(\theta)}\frac{\partial Y_t(\theta)}{\partial\theta}\frac{\partial Y_t(\theta)}{\partial\theta'}\right]$$

imply that $H_{ST0}^{-1}\frac{1}{T}\sum_{t=1}^{T}E\left(\nu_t^2 \mid F_{t-1}\right)\to I_4$ in probability as $T\to\infty$, using the ergodic theorem. Hence we complete the proof of Lemma 3.5.

Proof of Lemma 3.6. For $\theta=\theta_0$ the observed information is given by

$$\frac{\partial^2 L_T^{OT}(\theta)}{\partial\theta\partial\theta'}=-\frac{1}{T}\sum_{t=1}^{T}\left[-\frac{1}{J_t^2(\theta)}\frac{\partial J_t(\theta)}{\partial\theta}\frac{\partial J_t(\theta)}{\partial\theta'}+\frac{1}{J_t(\theta)}\frac{\partial^2 J_t(\theta)}{\partial\theta\partial\theta'}+\frac{1}{2h_t^2(\theta)}\left(-1+2z_t^2\right)\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\right.$$
$$\left.+\frac{1}{h_t(\theta)}\frac{\partial Y_t(\theta)}{\partial\theta}\frac{\partial Y_t(\theta)}{\partial\theta'}\right]$$
$$-\frac{1}{T}\sum_{t=1}^{T}\left[\frac{1}{2h_t(\theta)}\left(1-z_t^2\right)\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'}-\frac{2z_t}{h_t^{3/2}(\theta)}\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial Y_t(\theta)}{\partial\theta'}+\frac{z_t}{h_t^{1/2}(\theta)}\frac{\partial^2 Y_t(\theta)}{\partial\theta\partial\theta'}\right]$$

The first term on the right-hand side converge by the ergodic theorem to $-E\left[\frac{1}{J_t(\theta)}\frac{\partial^2 J_t(\theta)}{\partial\theta\partial\theta'}-\frac{1}{J_t^2(\theta)}\frac{\partial J_t(\theta)}{\partial\theta}\frac{\partial J_t(\theta)}{\partial\theta'}+\frac{1}{2h_t^2(\theta)}\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}+\frac{1}{h_t(\theta)}\frac{\partial Y_t(\theta)}{\partial\theta}\frac{\partial Y_t(\theta)}{\partial\theta'}\right]$; second term on the right-hand side converges in probability to zero. Hence we can complete the proof of Lemma 3.6.

Proof of Lemma 3.8. If $k=T$, then $B_{1m,T}(\phi,\theta)=0=B_{2m,T}(\phi,\theta)$ holds. If $k=T-1$, then $B_{1m,T-1}(\phi,\theta)=\frac{1}{2}m(m-1)-\frac{\phi^2}{2}$ and $B_{2m,T-1}(\phi,\theta)=\frac{\phi}{2}(2m-1)$ hold since $m=0,1$ such that $m(m-1)=0$. Suppose $k=t+1$ holds. When $k=t$,

$$\begin{aligned}B_{1m,t}(\phi,\theta)&=m(\gamma_Q-\frac{1}{2})-\frac{1}{2}\gamma_Q^2-\frac{\phi^2}{2}\beta b_{1m,t+1}(\phi,\theta)+\frac{(1+\alpha\phi^2 b_{1m,t+1}(\phi,\theta))[(m-\gamma_Q)^2-\phi^2]-2\alpha\phi^2(m-\gamma_Q)b_{2m,t+1}(\phi,\theta)}{2[(1+\alpha\phi^2 b_{1m,t+1}(\phi,\theta))^2+\alpha^2\phi^2 b_{2m,t+1}^2(\phi,\theta)]}\\&=-\frac{\phi^2}{2}b_{1m,t}(\phi,\theta)\end{aligned}$$

and

$$\begin{aligned}B_{2m,t}(\phi,\theta)&=\phi(\gamma_Q-\frac{1}{2})+\frac{\phi}{2}\beta b_{2m,t+1}(\phi,\theta)+\frac{(m-\gamma_Q)\phi(1+\alpha\phi^2 b_{1m,t+1}(\phi,\theta))+\alpha\frac{\phi}{2}b_{2m,t+1}(\phi,\theta)[(m-\gamma_Q)^2-\phi^2]}{(1+\alpha\phi^2 b_{1m,t+1}(\phi,\theta))^2+\alpha^2\phi^2 b_{2m,t+1}^2(\phi,\theta)}\\&=\frac{\phi}{2}b_{2m,t}(\phi,\theta).\end{aligned}$$

By induction, we can complete the proof.

Proof of Lemma 3.9. Since (3.29) and $0\le\frac{\phi^j}{(1+a\phi^2)^2+b^2\phi^2}\le 1$, for all $\phi\ge 0$ and $0\le j<4$, we can complete the proof by induction.

Proof of Lemma 3.10. By Lemmas 3.8 and 3.9, (3.40)-(3.43) follow.

Proof of Lemma 3.11. From simple recursion, we have

$$\begin{aligned}&A_{1m,t}(\phi,\theta)\\&=mr(T-t)-\frac{\phi^2}{2}\omega\sum_{i=1}^{T-t}b_{1m,t+i}(\phi,\theta)-\frac{1}{4}\sum_{i=1}^{T-t}\ln\left[\left(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)\right)^2+\alpha^2\phi^2 b_{2m,t+i}^2(\phi,\theta)\right],\end{aligned}$$

$$\begin{aligned}&A_{1m1,t}(\phi,\theta)\\&=-\frac{\phi^2}{2}\sum_{i=1}^{T-t}\left[\omega b_{1m1,t+i}(\phi,\theta)+\frac{(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta))\alpha b_{1m1,t+i}(\phi,\theta)+\alpha^2 b_{2m,t+i}(\phi,\theta)b_{2m1,t+i}(\phi,\theta)}{\left((1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta))^2+\alpha^2\phi^2 b_{2m,t+i}^2(\phi,\theta)\right)}\right],\end{aligned}$$

$$
\begin{aligned}
&A_{1m2,t}(\phi,\theta)\\
&=-\frac{\phi^2}{2}\sum_{i=1}^{T-t}\left[\begin{array}{l}
\omega b_{1m2,t+i}(\phi,\theta)+\dfrac{\alpha^2\phi^2 b_{1m1,t+i}^2(\phi,\theta)+(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta))\alpha b_{1m2,t+i}(\phi,\theta)+\alpha^2 b_{2m,t+i}(\phi,\theta)b_{2m2,t+i}(\phi,\theta)+\alpha^2 b_{2m1,t+i}^2(\phi,\theta)}{(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta))^2+\alpha^2\phi^2 b_{2m,t+i}^2(\phi,\theta)}\\
-\dfrac{\left((1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta))\alpha b_{1m1,t+i}(\phi,\theta)+\alpha^2 b_{2m1,t+i}(\phi,\theta)\right)\left(2\alpha\phi^2(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta))b_{1m1,t+i}(\phi,\theta)+2\alpha^2\phi^2 b_{2m,t+i}(\phi,\theta)b_{2m1,t+i}(\phi,\theta)\right)}{\left[(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta))^2+\alpha^2\phi^2 b_{2m,t+i}^2(\phi,\theta)\right]^2}
\end{array}\right],
\end{aligned}
$$

$$
\begin{aligned}
&A_{1m3,t}(\phi,\theta)\\
&=-\frac{\phi^2}{2}\sum_{i=1}^{T-t}\left[\begin{array}{l}
\omega b_{1m3,t+i}(\phi,\theta)+\dfrac{3\alpha^2\phi^2 b_{1m1,t+i}(\phi,\theta)b_{1m2,t+i}(\phi,\theta)+(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta))\alpha b_{1m3,t+i}(\phi,\theta)+3\alpha^2 b_{2m1,t+i}(\phi,\theta)b_{2m2,t+i}(\phi,\theta)+\alpha^2 b_{2m,t+i}(\phi,\theta)b_{2m3,t+i}(\phi,\theta)}{(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta))^2+\alpha^2\phi^2 b_{2m,t+i}^2(\phi,\theta)}\\
-4\phi^2\dfrac{\left(\alpha^2\phi^2 b_{1m1,t+i}^2(\phi,\theta)+(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta))\alpha b_{1m2,t+i}(\phi,\theta)+\alpha^2 b_{2m,t+i}(\phi,\theta)b_{2m2,t+i}(\phi,\theta)+\alpha^2 b_{2m1,t+i}^2(\phi,\theta)\right)\left(\alpha(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta))b_{1m1,t+i}(\phi,\theta)+\alpha^2 b_{2m,t+i}(\phi,\theta)b_{2m1,t+i}(\phi,\theta)\right)}{\left[(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta))^2+\alpha^2\phi^2 b_{2m,t+i}^2(\phi,\theta)\right]^2}\\
-2\phi^2\dfrac{\left((1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta))\alpha b_{1m1,t+i}(\phi,\theta)+\alpha^2 b_{2m1,t+i}(\phi,\theta)\right)\left(\alpha^2\phi^2 b_{1m1,t+i}^2(\phi,\theta)+\alpha(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta))b_{1m2,t+i}(\phi,\theta)+\alpha^2 b_{2m1,t+i}^2(\phi,\theta)+\alpha^2 b_{2m,t+i}(\phi,\theta)b_{2m2,t+i}(\phi,\theta)\right)}{\left[(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta))^2+\alpha^2\phi^2 b_{2m,t+i}^2(\phi,\theta)\right]^2}\\
+8\phi^4\dfrac{\left((1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta))\alpha b_{1m1,t+i}(\phi,\theta)+\alpha^2 b_{2m1,t+i}(\phi,\theta)\right)\left(\alpha(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta))b_{1m1,t+i}(\phi,\theta)+\alpha^2 b_{2m,t+i}(\phi,\theta)b_{2m1,t+i}(\phi,\theta)\right)^2}{\left[(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta))^2+\alpha^2\phi^2 b_{2m,t+i}^2(\phi,\theta)\right]^3}
\end{array}\right],
\end{aligned}
$$

$$
A_{2m,t}(\phi,\theta)=\phi r(T-t)+\frac{1}{2}\sum_{i=1}^{T-t}\left[\omega\phi b_{2m,t+i}(\phi,\theta)-\tan^{-1}\left(\frac{-\alpha\phi b_{2m,t+i}(\phi,\theta)}{1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)}\right)\right],
$$

$$
\begin{aligned}
&A_{2m1,t}(\phi,\theta)\\
&=\frac{\phi}{2}\sum_{i=1}^{T-t}\left[\omega b_{2m1,t+i}(\phi,\theta)+\frac{\alpha b_{2m1,t+i}(\phi,\theta)\left(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)\right)-\alpha^2\phi^2 b_{2m,t+i}(\phi,\theta)b_{1m1,t+i}(\phi,\theta)}{\left(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)\right)^2+\alpha^2\phi^2 b_{2m,t+i}^2(\phi,\theta)}\right],
\end{aligned}
$$

$$
\begin{aligned}
&A_{2m2,t}(\phi,\theta)\\
&=\frac{\phi}{2}\sum_{i=1}^{T-t}\left[\begin{array}{l}
\omega b_{2m2,t+i}(\phi,\theta)+\dfrac{\alpha b_{2m2,t+i}(\phi,\theta)\left(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)\right)-\alpha^2\phi^2 b_{2m,t+i}(\phi,\theta)b_{1m2,t+i}(\phi,\theta)}{\left(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)\right)^2+\alpha^2\phi^2 b_{2m,t+i}^2(\phi,\theta)}\\
-2\alpha\phi^2\dfrac{\left(\alpha b_{2m1,t+i}(\phi,\theta)\left(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)\right)-\alpha^2\phi^2 b_{2m,t+i}(\phi,\theta)b_{1m1,t+i}(\phi,\theta)\right)\left(b_{1m1,t+i}(\phi,\theta)\left(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)\right)+\alpha b_{2m,t+i}(\phi,\theta)b_{2m1,t+i}(\phi,\theta)\right)}{\left[\left(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)\right)^2+\alpha^2\phi^2 b_{2m,t+i}^2(\phi,\theta)\right]^2}
\end{array}\right],
\end{aligned}
$$

$$
\begin{aligned}
&A_{2m3,t}(\phi,\theta)\\
&=\frac{\phi}{2}\sum_{i=1}^{T-t}\left[\begin{array}{l}
\omega b_{2m3,t+i}(\phi,\theta)+\dfrac{\alpha b_{2m3,t+i}(\phi,\theta)\left(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)\right)+\alpha^2\phi^2 b_{2m2,t+i}(\phi,\theta)b_{1m1,t+i}(\phi,\theta)-\alpha^2\phi^2 b_{2m1,t+i}(\phi,\theta)b_{1m2,t+i}(\phi,\theta)-\alpha^2\phi^2 b_{2m,t+i}(\phi,\theta)b_{1m3,t+i}(\phi,\theta)}{\left(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)\right)^2+\alpha^2\phi^2 b_{2m,t+i}^2(\phi,\theta)}\\
-4\alpha\phi^2\dfrac{\left(\alpha b_{2m2,t+i}(\phi,\theta)\left(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)\right)-\alpha^2\phi^2 b_{2m,t+i}(\phi,\theta)b_{1m2,t+i}(\phi,\theta)\right)\left(b_{1m1,t+i}(\phi,\theta)\left(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)\right)+\alpha b_{2m,t+i}(\phi,\theta)b_{2m1,t+i}(\phi,\theta)\right)}{\left[\left(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)\right)^2+\alpha^2\phi^2 b_{2m,t+i}^2(\phi,\theta)\right]^2}\\
-2\alpha\phi^2\dfrac{\left(\alpha b_{2m1,t+i}(\phi,\theta)\left(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)\right)-\alpha^2\phi^2 b_{2m,t+i}(\phi,\theta)b_{1m1,t+i}(\phi,\theta)\right)\left(b_{1m2,t+i}(\phi,\theta)\left(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)\right)+\alpha\phi^2 b_{1m1,t+i}^2(\phi,\theta)+\alpha b_{2m1,t+i}^2(\phi,\theta)+\alpha b_{2m,t+i}(\phi,\theta)b_{2m2,t+i}(\phi,\theta)\right)}{\left[\left(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)\right)^2+\alpha^2\phi^2 b_{2m,t+i}^2(\phi,\theta)\right]^2}\\
+8\alpha^2\phi^4\dfrac{\left(\alpha b_{2m1,t+i}(\phi,\theta)\left(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)\right)-\alpha^2\phi^2 b_{2m,t+i}(\phi,\theta)b_{1m1,t+i}(\phi,\theta)\right)\left(b_{1m1,t+i}(\phi,\theta)\left(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)\right)+\alpha b_{2m,t+i}(\phi,\theta)b_{2m1,t+i}(\phi,\theta)\right)^2}{\left[\left(1+\alpha\phi^2 b_{1m,t+i}(\phi,\theta)\right)^2+\alpha^2\phi^2 b_{2m,t+i}^2(\phi,\theta)\right]^3}
\end{array}\right].
\end{aligned}
$$

From $b_{1m,t+i}(\phi,\theta)\ge 0$ for all $1\le i\le T-t$ such that $\sum_{i=1}^{T-t} b_{1m,t+i}(\phi,\theta)\ge b_{1m,T-1}(\phi,\theta)=1$, $\left|\tan^{-1}x\right|\le c\pi$, $\frac{1+a}{(1+a)^2+b^2}\le 1$, $\frac{b^2}{a^2+b^2}\le 1$, $\frac{1}{(1+a)^2+b^2}\le 1$, (3.29), and Lemma 3.9, we can complete these proofs of (3.44)-(3.48).

Proof of Lemma 3.12. By Lemmas 3.4, 3.10, and 3.11, (3.49)-(3.56) follow.

Proof of Lemma 3.13. From $\sin\left(x+2c\pi(T-t)\right)=\sin(x)$, $\sin(x)\le 1$, $\int_0^1 \frac{\sin x}{x}dx<\infty$, and Lemma 3.12, we have

$$\int_0^\infty \frac{e^{X_{1m,t}(\phi,\theta)}}{\phi}\sin\left(X_{3m,t}(\phi,\theta)\right)d\phi\le\int_0^\infty \frac{e^{-b\phi^2}}{\phi}\sin\left(\phi X_{3m,t}\right)d\phi=\int_0^1 \frac{\sin\left(\phi X_{3m,t}\right)}{\phi}d\phi+\int_1^\infty e^{-b\phi^2}d\phi$$

which implies (3.57). Similarly, (3.58)-(3.59) follow.

Proof of Lemma 3.14. Here, the first-, second-, and third-order derivatives of $J_t(\theta)$ and $Y_t(\theta)$ are

$$Y_{1t}(\theta)=\frac{S_{1t}(\theta)}{S_t(\theta)}-\frac{S_{1t-1}(\theta)}{S_{t-1}(\theta)}-\lambda h_{1t}(\theta),$$

$$Y_{2t}(\theta)=-\frac{S_{1t}^2(\theta)}{S_t^2(\theta)}+\frac{S_{2t}(\theta)}{S_t(\theta)}+\frac{S_{1t-1}^2(\theta)}{S_{t-1}^2(\theta)}-\frac{S_{2t-1}(\theta)}{S_{t-1}(\theta)}-\lambda h_{2t}(\theta),$$

$$Y_{3t}(\theta)=\frac{2S_{1t}^3(\theta)}{S_t^3(\theta)}-\frac{3S_{2t}(\theta)S_{1t}(\theta)}{S_t^2(\theta)}+\frac{S_{3t}(\theta)}{S_t(\theta)}-\frac{2S_{1t-1}^3(\theta)}{S_{t-1}^3(\theta)}+\frac{3S_{2t-1}(\theta)S_{1t-1}(\theta)}{S_{t-1}^2(\theta)}-\frac{S_{3t-1}(\theta)}{S_{t-1}(\theta)}-\lambda h_{3t}(\theta),$$

$$\begin{aligned}J_{1t}(\theta)=&\frac{1}{\pi}\int_0^\infty e^{X_{11,t}(\phi,\theta)}\left[\left(\frac{\sin\left(X_{31,t}(\phi,\theta)\right)}{\phi}+\cos\left(X_{31,t}(\phi,\theta)\right)\right)X_{111,t}(\phi,\theta)+\left(\frac{\cos\left(X_{31,t}(\phi,\theta)\right)}{\phi}-\sin\left(X_{31,t}(\phi,\theta)\right)\right)X_{311,t}(\phi,\theta)\right]d\phi\\&+e^{-r(T-t)}\frac{K}{S_t(\theta)}\frac{1}{\pi}\int_0^\infty e^{X_{10,t}(\phi,\theta)}\left[\frac{\cos\left(X_{30,t}(\phi,\theta)\right)}{S_t(\theta)}\frac{\partial S_t(\theta)}{\partial\theta}-\cos\left(X_{30,t}(\phi,\theta)\right)X_{101,t}(\phi,\theta)+\sin\left(X_{30,t}(\phi,\theta)\right)X_{301,t}(\phi,\theta)\right]d\phi,\end{aligned}$$

$$
\begin{aligned}
&J_{2t}(\theta) \\
&= \frac{1}{\pi}\int_0^\infty e^{X_{11,t}(\phi,\theta)} \left[\begin{array}{l} \left(\dfrac{\sin\left(X_{31,t}(\phi,\theta)\right)}{\phi} + \cos\left(X_{31,t}(\phi,\theta)\right) \right)\left(X_{111,t}^2(\phi,\theta) - X_{311,t}^2(\phi,\theta) + X_{112,t}(\phi,\theta) \right) \\ + \left(\dfrac{\cos\left(X_{31,t}(\phi,\theta)\right)}{\phi} - \sin\left(X_{31,t}(\phi,\theta)\right) \right)\left(2X_{311,t}(\phi,\theta)X_{111,t}(\phi,\theta) + X_{312,t}(\phi,\theta) \right) \end{array} \right] d\phi \\
&+ e^{-r(T-t)} \frac{K}{S_t(\theta)} \frac{1}{\pi}\int_0^\infty e^{X_{10,t}(\phi,\theta)} \left[\begin{array}{l} \dfrac{\cos\left(X_{30,t}(\phi,\theta)\right)}{S_t(\theta)}\left(2X_{101,t}(\phi,\theta)S_{1t}(\theta) - 2S_{1t}^2(\theta) + S_{2t}(\theta) \right) - \dfrac{2\sin\left(X_{30,t}(\phi,\theta)\right)}{S_t(\theta)} S_{1t}(\theta)X_{301,t}(\phi,\theta) \\ - \cos\left(X_{30,t}(\phi,\theta)\right)\left(X_{101,t}^2(\phi,\theta) + X_{102,t}(\phi,\theta) - X_{301,t}(\phi,\theta) \right) \\ + \sin\left(X_{30,t}(\phi,\theta)\right)\left(2X_{101,t}(\phi,\theta)X_{301,t}(\phi,\theta) + X_{302,t}(\phi,\theta) \right) \end{array} \right] d\phi,
\end{aligned}
$$

and

$$
\begin{aligned}
&J_{3t}(\theta) \\
&= \frac{1}{\pi}\int_0^\infty e^{X_{11,t}(\phi,\theta)} \left[\begin{array}{l} \left(\dfrac{\sin\left(X_{31,t}(\phi,\theta)\right)}{\phi} + \cos\left(X_{31,t}(\phi,\theta)\right) \right)\left(\begin{array}{l} X_{111,t}^3(\phi,\theta) - 3X_{311,t}^2(\phi,\theta)X_{111,t}(\phi,\theta) \\ +3X_{112,t}(\phi,\theta)X_{111,t}(\phi,\theta) - 3X_{312,t}(\phi,\theta)X_{311,t}(\phi,\theta) + X_{113,t}(\phi,\theta) \end{array} \right) \\ + \left(\dfrac{\cos\left(X_{31,t}(\phi,\theta)\right)}{\phi} - \sin\left(X_{31,t}(\phi,\theta)\right) \right)\left(\begin{array}{l} X_{313,t}(\phi,\theta) + 3X_{312,t}(\phi,\theta)X_{111,t}(\phi,\theta) + 3X_{112,t}(\phi,\theta)X_{311,t}(\phi,\theta) \\ +3X_{311,t}(\phi,\theta)X_{111,t}^2(\phi,\theta) - X_{311,t}^3(\phi,\theta) \end{array} \right) \end{array} \right] d\phi \\
&+ e^{-r(T-t)} \frac{K}{S_t(\theta)} \frac{1}{\pi}\int_0^\infty e^{X_{10,t}(\phi,\theta)} \left[\begin{array}{l} \dfrac{\cos\left(X_{30,t}(\phi,\theta)\right)}{S_t(\theta)}\left(\begin{array}{l} S_{3t}(\theta) - 2S_{1t}^2(\theta)X_{101,t}(\phi,\theta) - 4S_{2t}(\theta)S_{1t}(\theta) + 3X_{102,t}(\phi,\theta)S_{1t}(\theta) + 3S_{2t}(\theta)X_{101,t}(\phi,\theta) \\ -3S_{1t}(\theta)X_{301,t}^2(\phi,\theta) + 3X_{101,t}^2(\phi,\theta)S_{1t}(\theta) \end{array} \right) \\ + \dfrac{\cos\left(X_{30,t}(\phi,\theta)\right)}{S_t^2(\theta)}\left(-4X_{101,t}(\phi,\theta)S_{1t}^2(\theta) + 4S_{1t}^3(\theta) - 2S_{2t}(\theta)S_{1t}(\theta) \right) \\ + \dfrac{\sin\left(X_{30,t}(\phi,\theta)\right)}{S_t(\theta)}\left(\begin{array}{l} -6X_{101,t}(\phi,\theta)S_{1t}(\theta)X_{301,t}(\phi,\theta) - 3S_{2t}(\theta)X_{301,t}(\phi,\theta) - 2X_{102,t}(\phi,\theta)S_{1t}(\theta) \\ -X_{302,t}(\phi,\theta)S_{1t}(\theta) + 2S_{1t}^2(\theta)X_{301,t}(\phi,\theta) \end{array} \right) \\ + \dfrac{4\sin\left(X_{30,t}(\phi,\theta)\right)}{S_t^2(\theta)} X_{301,t}(\phi,\theta)S_{1t}^2(\theta) + \sin\left(X_{30,t}(\phi,\theta)\right)\left(\begin{array}{l} -X_{301,t}^3(\phi,\theta) + X_{303,t}(\phi,\theta) + 3X_{102,t}(\phi,\theta)X_{301,t}(\phi,\theta) \\ +3X_{302,t}(\phi,\theta)X_{101,t}(\phi,\theta) + 3X_{101,t}^2(\phi,\theta)X_{301,t}(\phi,\theta) \end{array} \right) \\ + \cos\left(X_{30,t}(\phi,\theta)\right)\left(\begin{array}{l} 3X_{101,t}(\phi,\theta)X_{301,t}^2(\phi,\theta) - 3X_{102,t}(\phi,\theta)X_{101,t}(\phi,\theta) \\ +3X_{302,t}(\phi,\theta)X_{301,t}(\phi,\theta) - X_{101,t}^3(\phi,\theta) - X_{103,t}(\phi,\theta) \end{array} \right) \end{array} \right] d\phi.
\end{aligned}
$$

By Lemmas 3.4 and 3.8-3.13, (3.60)-(3.63) follow.

Proof of Lemma 3.7. Without loss of generality, consider the case $\theta_i = \theta_j = \theta_k = \beta$. Noting that the expression for $\left(\partial^3 L_T^{OT} / \partial\beta^3\right)(\theta)$ in (3.34) implies that

$\left| \dfrac{\partial^3 L_T^{OT}(\theta)}{\partial\beta^3} \right| \le \dfrac{1}{T}\sum_{t=1}^{T} w_t(\theta)$, where

$w_t(\theta)$

$$
\begin{aligned}
&= \frac{2J_{1t}^3(\theta)}{J_t^3(\theta)} + \frac{3J_{2t}(\theta)J_{1t}(\theta)}{J_t^2(\theta)} + \frac{J_{3t}(\theta)}{J_t(\theta)} + \left(\frac{1}{h_t^3(\theta)} + \frac{3Y_t^2(\theta)}{h_t^4(\theta)}\right)h_{1t}^3(\theta) + \frac{3}{2}\left(\frac{1}{h_t^2(\theta)} + \frac{2Y_t^2(\theta)}{h_t^3(\theta)}\right)h_{2t}(\theta)h_{1t}(\theta) \\
&+ \frac{1}{2}\left(\frac{1}{h_t(\theta)} + \frac{Y_t^2(\theta)}{h_t^2(\theta)}\right)h_{3t}(\theta) + \frac{6Y_t(\theta)}{h_t^3(\theta)}h_{1t}^2(\theta)Y_{1t}(\theta) + \frac{3Y_t(\theta)}{h_t^2(\theta)}h_{2t}(\theta)Y_{1t}(\theta) + \frac{3}{h_t^2(\theta)}h_{1t}(\theta)Y_{1t}^2(\theta) \\
&+ \frac{3Y_t(\theta)}{h_t^2(\theta)}Y_{2t}(\theta)h_{1t}(\theta) + \frac{3}{h_t(\theta)}Y_{2t}(\theta)Y_{1t}(\theta) + \frac{Y_t(\theta)}{h_t(\theta)}Y_{3t}(\theta)
\end{aligned}
$$

Lemmas 3.4 and 3.14 suffice to show that there exists a neighborhood $N(\theta_0)$ for which

$$\sup_{\theta \in N(\theta_0)} \left| \frac{\partial^3 L_T^{OT}(\theta)}{\partial \beta^3} \right| \le \frac{1}{T}\sum_{t=1}^{T} w_t,$$

where w_t is stationary and has finite moment $Ew_t = M < \infty$ such that $\frac{1}{T}\sum_{t=1}^{T} w_t \xrightarrow{a.s.} M$ by the ergodic theorem, which ends the proof of Lemma 3.7.

Proof of Lemma 3.15. Evaluated at $\theta = \theta_0$ the form is given by

$$
\begin{aligned}
\frac{\partial L_T^{S+O+E}(\theta)}{\partial \theta} &= \frac{1}{T}\sum_{t=1}^{T}\left\{\left[\frac{-1}{2h_t(\theta)}\left(1 - \frac{z_t^2}{1-\rho^2} + \frac{\rho z_t u_t}{1-\rho^2}\right) - \frac{\lambda(\rho u_t - z_t)}{(1-\rho^2)h_t^{1/2}(\theta)}\right]\frac{\partial h_t(\theta)}{\partial \theta} - \frac{\rho z_t - u_t}{\eta(1-\rho^2)}\frac{\partial C_t^{HN}(\theta)}{\partial \theta}\right\} \\
&= \frac{1}{T}\sum_{t=1}^{T} v_t
\end{aligned}
$$

such that $E(v_t \mid F_{t-1}) = 0$, where $F_t = \sigma(z_t, u_t, z_{t-1}, u_{t-1}, \ldots)$. Applying the central limit theorem for martingale differences in Brown (1971), consider first

$$\frac{1}{T}\sum_{t=1}^{T} E(v_t^2 \mid F_{t-1}) \to H_{S+O+E0}$$

imply that $H_{S+O+E0}^{-1}\frac{1}{T}\sum_{t=1}^{T} E(v_t^2 \mid F_{t-1}) \to I_4$ in probability as $T \to \infty$, using the ergodic theorem. Hence we complete the proof of Lemma 3.15.

Proof of Lemma 3.16. For $\theta = \theta_0$ the observed information is given by

$$\frac{\partial^2 L_T^{S+O+E}(\theta)}{\partial\theta\partial\theta'}$$

$$=-\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t^2(\theta)}\left(-1+\frac{2z_t^2}{1-\rho^2}-\frac{3\rho z_t u_t}{2(1-\rho^2)}\right)+\frac{2\lambda(z_t-\rho u_t)}{(1-\rho^2)h_t^{3/2}(\theta)}+\frac{2\lambda^2}{(1-\rho^2)h_t(\theta)}\right]\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}$$

$$+\frac{1}{T}\sum_{t=1}^{T}\frac{\rho}{\eta(1-\rho^2)}\left(z_t+\frac{2\lambda}{h_t^{1/2}(\theta)}\right)\frac{\partial C_t^{HN}(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}-\frac{1}{T}\sum_{t=1}^{T}\frac{1}{\eta^2(1-\rho^2)}\frac{\partial C_t^{HN}(\theta)}{\partial\theta}\frac{\partial C_t^{HN}(\theta)}{\partial\theta'}$$

$$-\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t(\theta)}\left(1-\frac{z_t^2}{1-\rho^2}+\frac{\rho z_t u_t}{1-\rho^2}\right)-\frac{2\lambda(z_t-\rho u_t)}{(1-\rho^2)h_t^{1/2}(\theta)}\right]\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'}-\frac{1}{T}\sum_{t=1}^{T}\frac{\rho z_t-u_t}{\eta(1-\rho^2)}\frac{\partial^2 C_t^{HN}(\theta)}{\partial\theta\partial\theta'}.$$

The first term on the right-hand side converges by the ergodic theorem to $-E\left[\left(\frac{2-\rho^2}{4(1-\rho^2)}\frac{1}{h_t^2(\theta)}+\frac{\lambda^2}{(1-\rho^2)h_t(\theta)}\right)\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\right]$; second term on the right-hand side converges to $E\left[\frac{2\lambda\rho}{\eta(1-\rho^2)h_t^{1/2}(\theta)}\frac{\partial C_t^{HN}(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\right]$; third term on the right-hand side converges to $-E\left[\frac{1}{\eta^2(1-\rho^2)}\frac{\partial C_t^{HN}(\theta)}{\partial\theta}\frac{\partial C_t^{HN}(\theta)}{\partial\theta'}\right]$; last two terms on the right-hand side converges in probability to zero. Hence we can complete the proof of Lemma 3.16.

Proof of Lemma 3.18. Similar to the proof of Lemma 3.10, (3.67)-(3.70) follow.

Proof of Lemma 3.19. Here, the first-, second-, and third-order derivatives of $C_t^{HN}(\theta)$ are

$$C_{1t}^{HN}(\theta)$$

$$=\frac{1}{\pi}S_t\int_0^\infty\frac{e^{X_{11,t}(\phi,\theta)}}{\phi}\left[\sin\left(X_{21,t}(\phi,\theta)\right)X_{111,t}(\phi,\theta)+\cos\left(X_{21,t}(\phi,\theta)\right)X_{211,t}(\phi,\theta)\right]d\phi$$

$$-\frac{1}{\pi}e^{-r(T-t)}K\int_0^\infty\frac{e^{X_{10,t}(\phi,\theta)}}{\phi}\left[\sin\left(X_{20,t}(\phi,\theta)\right)X_{101,t}(\phi,\theta)+\cos\left(X_{20,t}(\phi,\theta)\right)X_{201,t}(\phi,\theta)\right]d\phi,$$

$$C_{2t}^{HN}(\theta)$$

$$=\frac{1}{\pi}S_t\int_0^\infty\frac{e^{X_{11,t}(\phi,\theta)}}{\phi}\begin{Bmatrix}\sin\left(X_{21,t}(\phi,\theta)\right)\left[X_{111,t}^2(\phi,\theta)-X_{211,t}^2(\phi,\theta)+X_{112,t}(\phi,\theta)\right]\\+\cos\left(X_{21,t}(\phi,\theta)\right)\left[2X_{111,t}(\phi,\theta)X_{211,t}(\phi,\theta)+X_{212,t}(\phi,\theta)\right]\end{Bmatrix}d\phi$$

$$-\frac{1}{\pi}e^{-r(T-t)}K\int_0^\infty\frac{e^{X_{10,t}(\phi,\theta)}}{\phi}\begin{Bmatrix}\sin\left(X_{20,t}(\phi,\theta)\right)\left[X_{101,t}^2(\phi,\theta)-X_{201,t}^2(\phi,\theta)+X_{102,t}(\phi,\theta)\right]\\+\cos\left(X_{20,t}(\phi,\theta)\right)\left[2X_{101,t}(\phi,\theta)X_{201,t}(\phi,\theta)+X_{202,t}(\phi,\theta)\right]\end{Bmatrix}d\phi,$$

and

$$C_{3t}^{HN}(\theta)$$

$$=\frac{1}{\pi}S_t\int_0^\infty \frac{e^{X_{11,t}(\phi,\theta)}}{\phi}\left\{\begin{array}{l}\sin\left(X_{21,t}(\phi,\theta)\right)\left[\begin{array}{l}X_{111,t}^3(\phi,\theta)-3X_{111,t}(\phi,\theta)X_{211,t}^2(\phi,\theta)\\+3X_{111,t}(\phi,\theta)X_{112,t}(\phi,\theta)-3X_{211,t}(\phi,\theta)X_{212,t}(\phi,\theta)\\+X_{113,t}(\phi,\theta)\end{array}\right]\\+\cos\left(X_{21,t}(\phi,\theta)\right)\left[\begin{array}{l}-X_{211,t}^3(\phi,\theta)+3X_{211,t}(\phi,\theta)X_{111,t}^2(\phi,\theta)\\+3X_{111,t}(\phi,\theta)X_{212,t}(\phi,\theta)\\+3X_{211,t}(\phi,\theta)X_{112,t}(\phi,\theta)+X_{213,t}(\phi,\theta)\end{array}\right]\end{array}\right\}d\phi$$

$$-\frac{1}{\pi}e^{-r(T-t)}K\int_0^\infty \frac{e^{X_{10,t}(\phi,\theta)}}{\phi}\left\{\begin{array}{l}\sin\left(X_{20,t}(\phi,\theta)\right)\left[\begin{array}{l}X_{101,t}^3(\phi,\theta)-3X_{101,t}(\phi,\theta)X_{201,t}^2(\phi,\theta)\\+3X_{101,t}(\phi,\theta)X_{102,t}(\phi,\theta)\\-3X_{201,t}(\phi,\theta)X_{202,t}(\phi,\theta)\\+X_{103,t}(\phi,\theta)\end{array}\right]\\+\cos\left(X_{20,t}(\phi,\theta)\right)\left[\begin{array}{l}-X_{201,t}^3(\phi,\theta)+3X_{201,t}(\phi,\theta)X_{101,t}^2(\phi,\theta)\\+3X_{101,t}(\phi,\theta)X_{202,t}(\phi,\theta)\\+3X_{201,t}(\phi,\theta)X_{102,t}(\phi,\theta)+X_{203,t}(\phi,\theta)\end{array}\right]\end{array}\right\}d\phi.$$

Similar to the proof of Lemma 3.14, (3.71)-(3.72) follow.

Proof of Lemma 3.17. Without loss of generality, consider the case $\theta_i=\theta_j=\theta_k=\beta$. Noting that by definition $y_t=r+\lambda h_t(\theta_0)+h_t^{1/2}(\theta_0)z_t$, the expression for $\left(\partial^3 L_T^{S+O+E}/\partial\beta^3\right)(\theta)$ in (3.66) implies that

$\left|\frac{\partial^3 L_T^{S+O+E}(\theta)}{\partial\beta^3}\right|\le\frac{1}{T}\sum_{t=1}^{T}w_t(\theta)$, where

$$
\begin{aligned}
w_t(\theta) &= \frac{1}{h_t^3(\theta)}\left(1+\frac{3}{1-\rho^2}\frac{\left[\lambda\left(h_t(\theta_0)+h_t(\theta)\right)+h_t^{1/2}(\theta_0)z_t\right]^2}{h_t(\theta)}+\frac{15\rho}{8(1-\rho^2)}\frac{\lambda\left(h_t(\theta_0)+h_t(\theta)\right)+h_t^{1/2}(\theta_0)z_t}{h_t^{1/2}(\theta)}\frac{C_t^{HN}(\theta_0)+C_t^{HN}(\theta)+\eta u_t}{\eta}\right)h_{1t}^3(\theta) \\
&+\frac{1}{(1-\rho^2)h_t^{5/2}(\theta)}\left(\frac{9\rho}{4}\frac{C_t^{HN}(\theta_0)+C_t^{HN}(\theta)+\eta u_t}{\eta}+4\lambda\frac{\lambda\left(h_t(\theta_0)+h_t(\theta)\right)+h_t^{1/2}(\theta_0)z_t}{h_t^{1/2}(\theta)}\right)h_{1t}^3(\theta)+\frac{2\lambda^2}{(1-\rho^2)h_t^2(\theta)}h_{1t}^3(\theta) \\
&+\frac{1}{h_t^2(\theta)}\left(\frac{3}{2}+\frac{3}{1-\rho^2}\frac{\left[\lambda\left(h_t(\theta_0)+h_t(\theta)\right)+h_t^{1/2}(\theta_0)z_t\right]^2}{h_t(\theta)}+\frac{9\rho}{4(1-\rho^2)}\frac{\lambda\left(h_t(\theta_0)+h_t(\theta)\right)+h_t^{1/2}(\theta_0)z_t}{h_t^{1/2}(\theta)}\frac{C_t^{HN}(\theta_0)+C_t^{HN}(\theta)+\eta u_t}{\eta}\right)h_{2t}(\theta)h_{1t}(\theta) \\
&+\frac{3\lambda}{(1-\rho^2)h_t^{3/2}(\theta)}\left(\frac{\lambda\left(h_t(\theta_0)+h_t(\theta)\right)+h_t^{1/2}(\theta_0)z_t}{h_t^{1/2}(\theta)}+\rho\frac{C_t^{HN}(\theta_0)+C_t^{HN}(\theta)+\eta u_t}{\eta}\right)h_{2t}(\theta)h_{1t}(\theta)+\frac{3\lambda^2}{(1-\rho^2)h_t(\theta)}h_{2t}(\theta)h_{1t}(\theta) \\
&+\frac{1}{2h_t(\theta)}\left(1+\frac{1}{1-\rho^2}\frac{\left[\lambda\left(h_t(\theta_0)+h_t(\theta)\right)+h_t^{1/2}(\theta_0)z_t\right]^2}{h_t(\theta)}+\frac{\rho}{1-\rho^2}\frac{\lambda\left(h_t(\theta_0)+h_t(\theta)\right)+h_t^{1/2}(\theta_0)z_t}{h_t^{1/2}(\theta)}\frac{C_t^{HN}(\theta_0)+C_t^{HN}(\theta)+\eta u_t}{\eta}\right)h_{3t}(\theta) \\
&+\frac{\lambda}{(1-\rho^2)h_t^{1/2}(\theta)}\left(\rho\frac{C_t^{HN}(\theta_0)+C_t^{HN}(\theta)+\eta u_t}{\eta}+\frac{\lambda\left(h_t(\theta_0)+h_t(\theta)\right)+h_t^{1/2}(\theta_0)z_t}{h_t^{1/2}(\theta)}\right)h_{3t}(\theta) \\
&+\frac{\rho}{\eta(1-\rho^2)}\left(\frac{9}{4}\frac{\lambda\left(h_t(\theta_0)+h_t(\theta)\right)+h_t^{1/2}(\theta_0)z_t}{h_t^{5/2}(\theta)}+\frac{3\lambda}{h_t^{3/2}(\theta)}\right)h_{1t}^2(\theta)C_{1t}^{HN}(\theta)+\frac{3\rho}{2\eta(1-\rho^2)}\left(\frac{\lambda\left(h_t(\theta_0)+h_t(\theta)\right)+h_t^{1/2}(\theta_0)z_t}{h_t^{3/2}(\theta)}+\frac{\lambda}{h_t^{1/2}(\theta)}\right)h_{2t}(\theta)C_{1t}^{HN}(\theta) \\
&+\frac{3\rho}{2\eta(1-\rho^2)}\left(\frac{\lambda\left(h_t(\theta_0)+h_t(\theta)\right)+h_t^{1/2}(\theta_0)z_t}{h_t^{3/2}(\theta)}+\frac{2\lambda}{h_t^{1/2}(\theta)}\right)C_{2t}^{HN}(\theta)h_{1t}(\theta)+\frac{3}{2\eta^2(1-\rho^2)}C_{2t}^{HN}(\theta)C_{1t}^{HN}(\theta) \\
&+\frac{1}{\eta(1-\rho^2)}\left(\rho\frac{\lambda\left(h_t(\theta_0)+h_t(\theta)\right)+h_t^{1/2}(\theta_0)z_t}{h_t^{1/2}(\theta)}+\frac{C_t^{HN}(\theta_0)+C_t^{HN}(\theta)+\eta u_t}{\eta}\right)C_{3t}^{HN}(\theta),
\end{aligned}
$$

From Lemmas 3.4 and 3.19, we have there exists a neighborhood $N(\theta_0)$ for which $\sup_{\theta\in N(\theta_0)}\left|\frac{\partial^3 L_T^{S+O+E}(\theta)}{\partial\beta^3}\right| \le \frac{1}{T}\sum_{t=1}^{T} w_t$, where w_t is stationary and has finite moment $Ew_t = M < \infty$ such that $\frac{1}{T}\sum_{t=1}^{T} w_t \xrightarrow{a.s.} M$ by the ergodic theorem, which ends the proof of Lemma 3.17.

Appendix B

Proofs apply Brown (1971) results regarding Central Limit Theorem analogs for martingale differences.

B.1. Asymptotic behavior for ST

We seek to prove the following:

Theorem B.1. *The estimator of ST is* $\hat{\theta}_T^{ST} = \underset{\theta\in\Theta}{\arg\max} L_T^{ST}(\theta)$. *It is consistent such that* $\hat{\theta}_T^{ST} \xrightarrow{p} \theta_0$ *and asymptotically Normal such that* $H_{ST0}^{-1/2} F_{ST0} T^{1/2}(\hat{\theta}_T^{ST} - \theta_0) \xrightarrow{D} N(0, I_3)$, *where* $F_{ST0} = -E\left(\dfrac{\partial^2 L_T^{ST}(\theta_0)}{\partial\theta\partial\theta'}\right)$ *and* $H_{ST0} = E\left(T\dfrac{\partial L_T^{ST}(\theta_0)}{\partial\theta}\dfrac{\partial L_T^{ST}(\theta_0)}{\partial\theta'}\right)$.

To begin, the log-likelihood function for ST is given by (4.7) and it follows that:

$$\frac{\partial L_T^{ST}(\theta)}{\partial\theta} = -\frac{1}{2T}\sum_{t=1}^{T}\frac{1}{h_t(\theta)}\left[1 - \frac{\left(y_t - \mu_t(\theta)\right)^2}{h_t(\theta)} - \left(\lambda - h_t^{1/2}(\theta)\right)\left(\frac{y_t - \mu_t(\theta)}{h_t^{1/2}(\theta)}\right)\right]\frac{\partial h_t(\theta)}{\partial\theta}, \tag{B.1}$$

$$\begin{aligned}\frac{\partial^2 L_T^{ST}(\theta)}{\partial\theta\partial\theta'} = &-\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t^2(\theta)}\left(-1 + \frac{2\left(y_t - \mu_t(\theta)\right)^2}{h_t(\theta)} + \frac{5}{2}\left(\lambda - h_t^{1/2}(\theta)\right)\left(\frac{y_t - \mu_t(\theta)}{h_t^{1/2}(\theta)}\right) + \frac{\left(\lambda - h_t^{1/2}(\theta)\right)^2}{2}\right) + \frac{1}{2h_t^{3/2}(\theta)}\left(\frac{y_t - \mu_t(\theta)}{h_t^{1/2}(\theta)}\right)\right]\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\\ &-\frac{1}{2T}\sum_{t=1}^{T}\frac{1}{h_t(\theta)}\left[1 - \frac{\left(y_t - \mu_t(\theta)\right)^2}{h_t(\theta)} - \left(\lambda - h_t^{1/2}(\theta)\right)\left(\frac{y_t - \mu_t(\theta)}{h_t^{1/2}(\theta)}\right)\right]\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'},\end{aligned}$$

(B.2)

$$
\begin{aligned}
&\frac{\partial^3 L_T^{ST}(\theta)}{\partial\theta\partial\theta'\partial\theta} \\
&= -\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t^3(\theta)}\left(2-\frac{6\left(y_t-\mu_t(\theta)\right)^2}{h_t(\theta)}-\frac{33}{4}\left(\lambda-h_t^{1/2}(\theta)\right)\left(\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\right)-\frac{9\left(\lambda-h_t^{1/2}(\theta)\right)^2}{4}\right)-\frac{3}{4h_t^{5/2}(\theta)}\left(3\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}+\lambda-h_t^{1/2}(\theta)\right)\right]\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\frac{\partial h_t(\theta)}{\partial\theta} \\
&\quad -\frac{3}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t^2(\theta)}\left(-1+\frac{2\left(y_t-\mu_t(\theta)\right)^2}{h_t(\theta)}+\frac{5}{2}\left(\lambda-h_t^{1/2}(\theta)\right)\left(\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\right)+\frac{\left(\lambda-h_t^{1/2}(\theta)\right)^2}{2}\right)+\frac{1}{2h_t^{3/2}(\theta)}\left(\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\right)\right]\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'}\frac{\partial h_t(\theta)}{\partial\theta} \\
&\quad -\frac{1}{2T}\sum_{t=1}^{T}\frac{1}{h_t(\theta)}\left[1-\frac{\left(y_t-\mu_t(\theta)\right)^2}{h_t(\theta)}-\left(\lambda-h_t^{1/2}(\theta)\right)\left(\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\right)\right]\frac{\partial^3 h_t(\theta)}{\partial\theta\partial\theta'\partial\theta},
\end{aligned}
\tag{B.3}
$$

where $\mu_t(\theta)=r+\lambda h_t^{1/2}(\theta)-\frac{1}{2}h_t(\theta)$ and $\frac{\partial\mu_t(\theta)}{\partial\theta}=\frac{\lambda-h_t^{1/2}(\theta)}{2h_t^{1/2}(\theta)}\frac{\partial h_t(\theta)}{\partial\theta}$.

Consider the asymptotic Normality of the first derivative and the limit of the observed information matrix in (B.1) and (B.2), using Lemmas B.1 and B.2, respectively:

Lemma B.1. *The form given by (B.1) evaluated at* $\theta=\theta_0$ *is asymptotically Gaussian,*

$H_{ST0}^{-1/2}T^{1/2}\frac{\partial L_T^{ST}(\theta_0)}{\partial\theta}\xrightarrow{p}N\left(0,I_3\right)$,

where $H_{ST0}=E\left(T\frac{\partial L_T^{ST}(\theta_0)}{\partial\theta}\frac{\partial L_T^{ST}(\theta_0)}{\partial\theta'}\right)=E\left[\frac{2+(\lambda-h_t^{1/2}(\theta_0))^2}{4h_t^2(\theta_0)}\frac{\partial h_t(\theta_0)}{\partial\theta}\frac{\partial h_t(\theta_0)}{\partial\theta'}\right]$.

Proof of Lemma B.1. Evaluated at $\theta=\theta_0$ the form is given by

$$
\frac{\partial L_T^{ST}(\theta)}{\partial\theta}=-\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t(\theta)}\left(1-z_t^2-(\lambda-h_t^{1/2}(\theta_0))z_t\right)\right]\frac{\partial h_t(\theta)}{\partial\theta}=\frac{1}{T}\sum_{t=1}^{T}\nu_t
$$

such that $E\left(\nu_t\,|\,F_{t-1}\right)=0$, where $F_t=\sigma\left(z_t,z_{t-1},\ldots\right)$. Applying the central limit theorem for martingale differences in Brown (1971), consider first

$$
\frac{1}{T}\sum_{t=1}^{T}E\left(\nu_t^2\,|\,F_{t-1}\right)=\frac{1}{T}\sum_{t=1}^{T}\frac{2+(\lambda-h_t^{1/2}(\theta))^2}{4h_t^2(\theta)}\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\to E\left[\frac{2+(\lambda-h_t^{1/2}(\theta))^2}{4h_t^2(\theta)}\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\right]
$$

imply that $H_{ST0}^{-1}\frac{1}{T}\sum_{t=1}^{T}E\left(\nu_t^2\,|\,F_{t-1}\right)\to I_3$ in probability as $T\to\infty$, using the ergodic theorem.

Hence we complete the proof of Lemma B.1.

Lemma B.2. *The observed information matrix given by (B.2) evaluated at* $\theta=\theta_0$ *converges in probability to* $-F_{ST0}^{-1}\dfrac{\partial^2 L_T^{ST}(\theta_0)}{\partial\theta\partial\theta'}\xrightarrow{p} I_3$, *where* $F_{ST0}=-E\left(\dfrac{\partial^2 L_T^{ST}(\theta_0)}{\partial\theta\partial\theta'}\right)$.

Proof of Lemma B.2. For $\theta=\theta_0$ the observed information is given by

$$
\begin{aligned}
&\frac{\partial^2 L_T^{ST}(\theta)}{\partial\theta\partial\theta'}\\
&=-\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t^2(\theta)}\left(-1+2z_t^2+\frac{5}{2}(\lambda-h_t^{1/2}(\theta))z_t+\frac{(\lambda-h_t^{1/2}(\theta))^2}{2}\right)+\frac{z_t}{2h_t^{3/2}(\theta)}\right]\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\\
&\quad-\frac{1}{2T}\sum_{t=1}^{T}\frac{1}{h_t(\theta)}\left(1-z_t^2-(\lambda-h_t^{1/2}(\theta))z_t\right)\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'}.
\end{aligned}
$$

The first term on the right-hand side converges by the ergodic theorem to $-E\left[\dfrac{2+(\lambda-h_t^{1/2}(\theta))^2}{4h_t^2(\theta)}\dfrac{\partial h_t(\theta)}{\partial\theta}\dfrac{\partial h_t(\theta)}{\partial\theta'}\right]$; second term on the right-hand side converges in probability to zero. Hence we can complete the proof of Lemma B.2.

Then, evaluating the third derivative of the likelihood function in (B.3), we seek to show that it is uniformly bounded in a neighborhood around the true parameter value θ_0. The neighborhood $N(\theta_0)$ around the true value θ_0 defined as

$$N(\theta_0)=\left\{\theta\,|\,0<\omega_L\le\omega_0\le\omega_U,0<\alpha_L\le\alpha_0\le\alpha_U,0<\beta_L\le\beta_0\le\beta_U,\alpha_U(1+\lambda^2)+\beta_U<1\right\}. \tag{B.4}$$

Lemma B.3. *There exists* $N(\theta_0)$, *for all* $1\le i,j,k\le 3$, *for which*

$\sup_{\theta\in N(\theta_0)}\left|\dfrac{\partial^3 L_T^{ST}(\theta)}{\partial\theta_i\partial\theta_j\partial\theta_k}\right|\le g(\omega_L,\omega_U,\alpha_L,\alpha_U,\beta_L,\beta_U,T)\xrightarrow{a.s.}M<\infty$ *as* $T\to\infty$ *where* M *is constant.*

In order to prove Lemma B.3, without loss of generality, consider the case $\theta_i = \theta_j = \theta_k = \beta$. The next lemma shows that the individual terms of the third derivative $\left(\partial^3 L_T^{ST} / \partial\beta^3\right)(\theta)$ in (B.3) are uniformly bounded in the neighborhood $N(\theta_0)$.

Lemma B.4. *With $N(\theta_0)$ defined in (B.4), then for any t and $i = 1,2,3$,*

$$\omega_L \le \sup_{\theta\in N(\theta_0)} h_t(\theta) \le H_{tU} \tag{B.5}$$

and

$$\sup_{\theta\in N(\theta_0)} h_{it}(\theta) \le H_{it}, \tag{B.6}$$

where $h_{it}(\theta) \equiv \dfrac{\partial^i h_t(\theta)}{\partial\beta^i}$, $H_{0U} = \omega_U + \alpha_U(y_0 - r - \lambda h_0^{1/2} + \frac{1}{2}h_0)^2 + \beta_U h_0$,

$$H_{tU} = \omega_U + \alpha_U(y_{t-1}-r)^2 + \left[\alpha_U(y_{t-1}-r)^2 + \alpha_U(1+\lambda^2) + \beta_U\right]H_{t-1U} + 2\alpha_U\left[\lambda^2(y_{t-1}-r)^2+1\right]H_{t-1U}^{1/2}$$
$$+\alpha_U\left(1+\lambda^2\right)H_{t-1U}^{3/2} + \frac{\alpha_U H_{t-1U}^2}{4},$$

$$H_{1t} = H_{t-1U} + \sum_{i=1}^{t-1} H_{t-2-iU} \times$$
$$\prod_{j=0}^{i}\left[\alpha_U(y_{t-1-j}-r)^2 + \frac{\alpha_U\left(\lambda^2(y_{t-1-j}-r)^2+1\right)}{\omega_L^{1/2}} + \frac{3\alpha_U(1+\lambda^2)}{2}H_{t-1-jU}^{1/2} + \frac{\alpha_U}{2}H_{t-1-jU} + \alpha_U(1+\lambda^2) + \beta_U\right],$$

$$H_{2t} = 2H_{1t-1} + \frac{\alpha_U}{2}\left[1 + \frac{3(\lambda^2+1)}{2\omega_L^{1/2}} + \frac{\lambda^2(y_{t-1}-r)^2+1}{\omega_L^{3/2}}\right]H_{1t-1}^2$$
$$+\sum_{i=1}^{t-1}\left[2H_{1t-2-i} + \frac{\alpha_U}{2}\left(1 + \frac{3(\lambda^2+1)}{2\omega_L^{1/2}} + \frac{\lambda^2(y_{t-2-i}-r)^2+1}{\omega_L^{3/2}}\right)H_{1t-2-i}^2\right]\times$$
$$\prod_{j=0}^{i}\left[\alpha_U(y_{t-1-j}-r)^2 + \frac{\alpha_U\left(\lambda^2(y_{t-1-j}-r)^2+1\right)}{\omega_L^{1/2}} + \frac{3\alpha_U(1+\lambda^2)}{2}H_{t-1-jU}^{1/2} + \frac{\alpha_U}{2}H_{t-1-jU} + \alpha_U(1+\lambda^2) + \beta_U\right],$$

and

$$H_{3t} = 3H_{2t-1} + \frac{3\alpha_U}{4}\left[\frac{\lambda^2+1}{2\omega_L^{3/2}} + \frac{\lambda^2(y_{t-1}-r)^2+1}{\omega_L^{5/2}}\right]H_{1t-1}^3 + \frac{3\alpha_U}{2}\left[1+\frac{3(\lambda^2+1)}{2\omega_L^{1/2}} + \frac{\lambda^2(y_{t-1}-r)^2+1}{\omega_L^{3/2}}\right]H_{2t-1}H_{1t-1}$$
$$+3\sum_{i=1}^{t-1}\left[H_{2t-2-i} + \frac{\alpha_U}{4}\left(\frac{\lambda^2+1}{2\omega_L^{3/2}} + \frac{\lambda^2(y_{t-2-i}-r)^2+1}{\omega_L^{5/2}}\right)H_{1t-2-i}^3 + \frac{\alpha_U}{2}\left(1+\frac{3(\lambda^2+1)}{2\omega_L^{1/2}} + \frac{\lambda^2(y_{t-2-i}-r)^2+1}{\omega_L^{3/2}}\right)H_{2t-2-i}H_{1t-2-i}\right]\times$$
$$\prod_{j=0}^{i}\left[\alpha_U(y_{t-1-j}-r)^2 + \frac{\alpha_U\left(\lambda^2(y_{t-1-j}-r)^2+1\right)}{\omega_L^{1/2}} + \frac{3\alpha_U(1+\lambda^2)}{2}H_{t-1-jU}^{1/2} + \frac{\alpha_U}{2}H_{t-1-jU} + \alpha_U(1+\lambda^2)+\beta_U\right].$$

Proof of Lemma B.4. Applying (B.4), $h_t(\theta)\ge\omega$, and $|x|\le x^2+1$, if $t=1$, then $h_1(\theta)\le\omega_U+\alpha_U(y_0-r-\lambda h_0^{1/2}+\frac{1}{2}h_0)^2+\beta_U h_0$ holds. Suppose that $t=k$ holds. When $t=k+1$,

$$\begin{aligned}
h_{k+1}(\theta) &= \omega+\alpha(y_k-r)^2+\alpha(y_k-r)h_k(\theta)-2\alpha\lambda(y_k-r)h_k^{1/2}(\theta)+\frac{\alpha}{4}h_k^2(\theta)-\alpha\lambda h_k^{3/2}(\theta)+(\alpha\lambda^2+\beta)h_k(\theta)\\
&\le \omega+\alpha(y_k-r)^2+\alpha\left[(y_k-r)^2+1\right]h_k(\theta)+2\alpha h_k^{1/2}(\theta)\left[\lambda^2(y_k-r)^2+1\right]+\frac{\alpha}{4}h_k^2(\theta)\\
&\quad+\alpha h_k^{3/2}(\theta)\left[\lambda^2+1\right]+(\alpha\lambda^2+\beta)h_k(\theta)\\
&\le \omega_U+\alpha_U(y_k-r)^2+\alpha_U\left[(y_k-r)^2+1\right]H_{kU}+2\alpha_U H_{kU}^{1/2}\left[\lambda^2(y_k-r)^2+1\right]+\frac{\alpha_U H_{kU}^2}{4}\\
&\quad+\alpha_U H_{kU}^{3/2}\left[\lambda^2+1\right]+(\alpha_U\lambda^2+\beta_U)H_{kU}\\
&= H_{k+1U}.
\end{aligned}$$

By induction, (B.5) follows.

Here, the first-, second-, and third-order derivatives of $h_t(\theta)$ are

$$h_{1t}(\theta)=h_{t-1}(\theta)+\left[\alpha\left(y_{t-1}-r-\lambda h_{t-1}^{1/2}(\theta)+\frac{1}{2}h_{t-1}(\theta)\right)(1-\lambda h_{t-1}^{-1/2}(\theta))+\beta\right]h_{1t-1}(\theta),$$

$$\begin{aligned}
h_{2t}(\theta) &= 2h_{1t-1}(\theta)+\frac{\alpha}{2}\left[\frac{\lambda}{h_{t-1}^{3/2}(\theta)}\left(y_{t-1}-r-\lambda h_{t-1}^{1/2}(\theta)+\frac{1}{2}h_{t-1}(\theta)\right)+\left(1-\lambda h_{t-1}^{-1/2}(\theta)\right)^2\right]h_{1t-1}^2(\theta)\\
&\quad+\left[\alpha\left(y_{t-1}-r-\lambda h_{t-1}^{1/2}(\theta)+\frac{1}{2}h_{t-1}(\theta)\right)(1-\lambda h_{t-1}^{-1/2}(\theta))+\beta\right]h_{2t-1}(\theta),
\end{aligned}$$

and

$$
\begin{aligned}
h_{3t}(\theta) = {} & 3h_{2t-1}(\theta) + \frac{3\alpha\lambda}{4}\left[\frac{1}{2h_{t-1}^{3/2}(\theta)} - \frac{y_{t-1}-r}{h_{t-1}^{5/2}(\theta)}\right]h_{1t-1}^{3}(\theta) \\
& + \frac{3\alpha}{2}\left[\frac{\lambda}{h_{t-1}^{3/2}(\theta)}\left(y_{t-1} - r - \lambda h_{t-1}^{1/2}(\theta) + \frac{1}{2}h_{t-1}(\theta)\right) + \left(1-\lambda h_{t-1}^{-1/2}(\theta)\right)^2\right]h_{2t-1}(\theta)h_{1t-1}(\theta) \\
& + \left[\alpha\left(y_{t-1} - r - \lambda h_{t-1}^{1/2}(\theta) + \frac{1}{2}h_{t-1}(\theta)\right)(1-\lambda h_{t-1}^{-1/2}(\theta)) + \beta\right]h_{3t-1}(\theta).
\end{aligned}
$$

From h_0 is a constant such that $h_{i0} = 0$, $i = 1,2,3$, and applying simple recursions,

$$
h_{1t}(\theta) \le h_{t-1}(\theta) + \sum_{i=1}^{t-1} h_{t-2-i}(\theta)\prod_{j=0}^{i}\left[\alpha(y_{t-1-j}-r)^2 + \frac{\alpha\left(\lambda^2(y_{t-1-j}-r)^2+1\right)}{\omega^{1/2}} + \frac{3\alpha(1+\lambda^2)}{2}h_{t-1-j}^{1/2}(\theta) + \frac{\alpha}{2}h_{t-1-j}(\theta) + \alpha(1+\lambda^2) + \beta\right],
$$

$$
\begin{aligned}
h_{2t}(\theta) \le {} & 2h_{1,t-1}(\theta) + \frac{\alpha}{2}\left[1 + \frac{3(\lambda^2+1)}{2\omega^{1/2}} + \frac{\lambda^2(y_{t-1}-r)^2+1}{\omega^{3/2}}\right]h_{1,t-1}^{2}(\theta) \\
& + \sum_{i=1}^{t-1}\left[2h_{1,t-2-i}(\theta) + \frac{\alpha}{2}\left(1 + \frac{3(\lambda^2+1)}{2\omega^{1/2}} + \frac{\lambda^2(y_{t-2-i}-r)^2+1}{\omega^{3/2}}\right)h_{1,t-2-i}^{2}(\theta)\right]\times \\
& \prod_{j=0}^{i}\left[\alpha(y_{t-1-j}-r)^2 + \frac{\alpha\left(\lambda^2(y_{t-1-j}-r)^2+1\right)}{\omega^{1/2}} + \frac{3\alpha(1+\lambda^2)}{2}h_{t-1-j}^{1/2}(\theta) + \frac{\alpha}{2}h_{t-1-j}(\theta) + \alpha(1+\lambda^2) + \beta\right],
\end{aligned}
$$

and

$$
\begin{aligned}
h_{3t}(\theta) \le {} & 3h_{2,t-1}(\theta) + \frac{3\alpha}{4}\left[\frac{\lambda^2+1}{2\omega^{3/2}} + \frac{\lambda^2(y_{t-1}-r)^2+1}{\omega^{5/2}}\right]h_{1,t-1}^{3}(\theta) + \frac{3\alpha}{2}\left[1 + \frac{3(\lambda^2+1)}{2\omega^{1/2}} + \frac{\lambda^2(y_{t-1}-r)^2+1}{\omega^{3/2}}\right]h_{2,t-1}(\theta)h_{1,t-1}(\theta) \\
& + 3\sum_{i=1}^{t-1}\left[H_{2,t-2-i}(\theta) + \frac{\alpha}{4}\left(\frac{\lambda^2+1}{2\omega^{3/2}} + \frac{\lambda^2(y_{t-2-i}-r)^2+1}{\omega^{5/2}}\right)h_{1,t-2-i}^{3}(\theta) + \frac{\alpha}{2}\left(1 + \frac{3(\lambda^2+1)}{2\omega^{1/2}} + \frac{\lambda^2(y_{t-2-i}-r)^2+1}{\omega^{3/2}}\right)h_{2,t-2-i}(\theta)h_{1,t-2-i}(\theta)\right]\times \\
& \prod_{j=0}^{i}\left[\alpha(y_{t-1-j}-r)^2 + \frac{\alpha\left(\lambda^2(y_{t-1-j}-r)^2+1\right)}{\omega^{1/2}} + \frac{3\alpha(1+\lambda^2)}{2}h_{t-1-j}^{1/2}(\theta) + \frac{\alpha}{2}h_{t-1-j}(\theta) + \alpha(1+\lambda^2) + \beta\right].
\end{aligned}
$$

Then, (B.6) follows by applying (B.5).

Proof of Lemma B.3. Without loss of generality, consider the case $\theta_i = \theta_j = \theta_k = \beta$. Noting that by definition $y_t = r + \lambda h_t^{1/2}(\theta_0) - \frac{1}{2}h_t(\theta_0) + h_t^{1/2}(\theta_0)z_t$, the expression for $\left(\partial^3 L_T^{ST} / \partial\beta^3\right)(\theta)$ in (B.3) implies that

$\left|\dfrac{\partial^3 L_T^{ST}(\theta)}{\partial\beta^3}\right| \le \dfrac{1}{T}\sum_{t=1}^{T} w_t(\theta)$, where

$$
\begin{aligned}
& w_t(\theta) \\
&= \left[\begin{array}{l} \dfrac{1}{h_t^3(\theta)} + \dfrac{3[\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t]^2}{h_t^4(\theta)} + \dfrac{33(\lambda+h_t^{1/2}(\theta))[\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t]}{8h_t^{7/2}(\theta)} \\ + \dfrac{9[(\lambda+h_t^{1/2}(\theta))^2+\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t]}{8h_t^3(\theta)} + \dfrac{3[\lambda+h_t^{1/2}(\theta)]}{8h_t^{5/2}(\theta)} \end{array} \right] h_{1t}^3(\theta) \\
&\quad + \left[\begin{array}{l} \dfrac{3}{2h_t^2(\theta)} + \dfrac{3[\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t]^2}{h_t^3(\theta)} + \dfrac{15(\lambda+h_t^{1/2}(\theta))[\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t]}{4h_t^{5/2}(\theta)} \\ + \dfrac{3[(\lambda+h_t^{1/2}(\theta))^2+\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t]}{4h_t^2(\theta)} \end{array} \right] h_{2t}(\theta)h_{1t}(\theta) \\
&\quad + \left[\dfrac{1}{2h_t(\theta)} + \dfrac{[\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t]^2}{2h_t^2(\theta)} + \dfrac{(\lambda+h_t^{1/2}(\theta))[\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t]}{2h_t^{3/2}(\theta)} \right] h_{3t}(\theta).
\end{aligned}
$$

Lemma B.4 suffices to show that there exists a neighborhood $N(\theta_0)$ for which

$$
\sup_{\theta \in N(\theta_0)} \left| \frac{\partial^3 L_T^{ST}(\theta)}{\partial \beta^3} \right| \le \frac{1}{T}\sum_{t=1}^{T} w_t,
$$

where w_t is stationary and has finite moment $Ew_t = M < \infty$ such that $\frac{1}{T}\sum_{t=1}^{T} w_t \xrightarrow{a.s.} M$ by the ergodic theorem, which ends the proof of Lemma B.3.

Proof of Theorem B.1. From Lemma B.3, we have that

$$
\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \frac{\partial L_T^{ST}(\hat{\theta}_T^{ST})}{\partial \theta} \approx \frac{\partial L_T^{ST}(\theta_0)}{\partial \theta} + \frac{\partial^2 L_T^{ST}(\theta_0)}{\partial \theta \partial \theta'}(\hat{\theta}_T^{ST} - \theta_0)
$$

and

$$
\hat{\theta}_T^{ST} - \theta_0 \approx -\left(\frac{\partial^2 L_T^{ST}(\theta_0)}{\partial \theta \partial \theta'} \right)^{-1} \frac{\partial L_T^{ST}(\theta_0)}{\partial \theta}.
$$

Combining with Lemmas B.1 and B.2, we complete the proof of Theorem B.1.

B.2. Asymptotic behavior for S+O+E

We seek to prove the following:

Theorem B.2. *The estimator of S+O+E is* $\hat{\theta}_T^{S+O+E} = \arg\max_{\theta\in\Theta} L_T^{S+O+E}(\theta)$. *It is consistent such that* $\hat{\theta}_T^{S+O+E} \xrightarrow{p} \theta_0$ *and asymptotically Normal such that* $H_{S+O+E0}^{-1/2} F_{S+O+E0} T^{1/2}(\hat{\theta}_T^{S+O+E} - \theta_0) \overset{A}{\sim} N(0, I_3)$, *where* $F_{S+O+E0} = -E\left(\dfrac{\partial^2 L_T^{S+O+E}(\theta_0)}{\partial\theta\partial\theta'}\right)$ *and* $H_{S+O+E0} = E\left(T\dfrac{\partial L_T^{S+O+E}(\theta_0)}{\partial\theta}\dfrac{\partial L_T^{S+O+E}(\theta_0)}{\partial\theta'}\right)$.

The log-likelihood function for S+O+E is given by (4.11) and it follows that:

$$
\begin{aligned}
\frac{\partial L_T^{S+O+E}(\theta)}{\partial\theta} ={}& -\frac{1}{2T}\sum_{t=1}^{T}\frac{1}{h_t(\theta)}\left[1-\frac{1}{1-\rho^2}\frac{(y_t-\mu_t(\theta))^2}{h_t(\theta)}+\frac{\rho}{1-\rho^2}\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\frac{C_t-C_t^D(\theta)}{\eta}+\frac{\lambda-h_t^{1/2}(\theta)}{1-\rho^2}\left(\rho\frac{C_t-C_t^D(\theta)}{\eta}-\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\right)\right]\frac{\partial h_t(\theta)}{\partial\theta} \\
&-\frac{1}{T}\sum_{t=1}^{T}\frac{1}{\eta(1-\rho^2)}\left(\rho\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}-\frac{C_t-C_t^D(\theta)}{\eta}\right)\frac{\partial C_t^D(\theta)}{\partial\theta},
\end{aligned}
\tag{B.7}
$$

$$
\begin{aligned}
\frac{\partial^2 L_T^{S+O+E}(\theta)}{\partial\theta\partial\theta'} ={}& -\frac{1}{2T}\sum_{t=1}^{T}\left\{\begin{array}{l}\dfrac{1}{h_t^2(\theta)}\left[-1+\dfrac{2}{1-\rho^2}\dfrac{(y_t-\mu_t(\theta))^2}{h_t(\theta)}-\dfrac{3\rho}{2(1-\rho^2)}\dfrac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\dfrac{C_t-C_t^D(\theta)}{\eta}-\dfrac{\lambda-h_t^{1/2}(\theta)}{2(1-\rho^2)}\left(3\rho\dfrac{C_t-C_t^D(\theta)}{\eta}-5\dfrac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}-\lambda+h_t^{1/2}(\theta)\right)\right] \\ -\dfrac{1}{2(1-\rho^2)h_t^{3/2}(\theta)}\left(\rho\dfrac{C_t-C_t^D(\theta)}{\eta}-\dfrac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\right)\end{array}\right\}\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'} \\
&-\frac{1}{2T}\sum_{t=1}^{T}\frac{1}{h_t(\theta)}\left[1-\frac{1}{1-\rho^2}\frac{(y_t-\mu_t(\theta))^2}{h_t(\theta)}+\frac{\rho}{1-\rho^2}\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\frac{C_t-C_t^D(\theta)}{\eta}+\frac{\lambda-h_t^{1/2}(\theta)}{1-\rho^2}\left(\rho\frac{C_t-C_t^D(\theta)}{\eta}-\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\right)\right]\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'} \\
&+\frac{1}{T}\sum_{t=1}^{T}\frac{\rho}{\eta(1-\rho^2)h_t(\theta)}\left(\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}+\lambda-h_t^{1/2}(\theta)\right)\frac{\partial C_t^D(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}-\frac{1}{T}\sum_{t=1}^{T}\frac{1}{\eta^2(1-\rho^2)}\frac{\partial C_t^D(\theta)}{\partial\theta}\frac{\partial C_t^D(\theta)}{\partial\theta'}-\frac{1}{T}\sum_{t=1}^{T}\frac{1}{\eta(1-\rho^2)}\left(\rho\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}-\frac{C_t-C_t^D(\theta)}{\eta}\right)\frac{\partial^2 C_t^D(\theta)}{\partial\theta\partial\theta'},
\end{aligned}
\tag{B.8}
$$

$$
\begin{aligned}
\frac{\partial^3 L_T^{S+O+E}(\theta)}{\partial\theta\partial\theta'\partial\theta} ={}& -\frac{1}{2T}\sum_{t=1}^{T}\left\{\begin{array}{l}\dfrac{1}{h_t^3(\theta)}\left[2-\dfrac{6}{1-\rho^2}\dfrac{(y_t-\mu_t(\theta))^2}{h_t(\theta)}+\dfrac{15\rho}{4(1-\rho^2)}\dfrac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\dfrac{C_t-C_t^D(\theta)}{\eta}+\dfrac{3(\lambda-h_t^{1/2}(\theta))}{4(1-\rho^2)}\left(5\rho\dfrac{C_t-C_t^D(\theta)}{\eta}-11\dfrac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}-3(\lambda-h_t^{1/2}(\theta))\right)\right] \\ +\dfrac{3}{4(1-\rho^2)h_t^{5/2}(\theta)}\left(2\rho\dfrac{C_t-C_t^D(\theta)}{\eta}-3\dfrac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}-\lambda+h_t^{1/2}(\theta)\right)\end{array}\right\}\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\frac{\partial h_t(\theta)}{\partial\theta} \\
&-\frac{3}{2T}\sum_{t=1}^{T}\left\{\begin{array}{l}\dfrac{1}{h_t^2(\theta)}\left[-1+\dfrac{2}{1-\rho^2}\dfrac{(y_t-\mu_t(\theta))^2}{h_t(\theta)}-\dfrac{3\rho}{2(1-\rho^2)}\dfrac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\dfrac{C_t-C_t^D(\theta)}{\eta}-\dfrac{\lambda-h_t^{1/2}(\theta)}{2(1-\rho^2)}\left(3\rho\dfrac{C_t-C_t^D(\theta)}{\eta}-5\dfrac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}-\lambda+h_t^{1/2}(\theta)\right)\right] \\ -\dfrac{1}{2(1-\rho^2)h_t^{3/2}(\theta)}\left(\rho\dfrac{C_t-C_t^D(\theta)}{\eta}-\dfrac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\right)\end{array}\right\}\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'}\frac{\partial h_t(\theta)}{\partial\theta} \\
&-\frac{3}{4T}\sum_{t=1}^{T}\frac{\rho}{\eta(1-\rho^2)}\left[\frac{3}{h_t^2(\theta)}\left(\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}+\lambda-h_t^{1/2}(\theta)\right)+\frac{1}{h_t^{3/2}(\theta)}\right]\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\frac{\partial C_t^D(\theta)}{\partial\theta}+\frac{3}{2T}\sum_{t=1}^{T}\frac{\rho}{\eta(1-\rho^2)h_t(\theta)}\left(\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}+\lambda-h_t^{1/2}(\theta)\right)\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'}\frac{\partial C_t^D(\theta)}{\partial\theta} \\
&-\frac{1}{2T}\sum_{t=1}^{T}\frac{1}{h_t(\theta)}\left[1-\frac{1}{1-\rho^2}\frac{(y_t-\mu_t(\theta))^2}{h_t(\theta)}+\frac{\rho}{1-\rho^2}\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\frac{C_t-C_t^D(\theta)}{\eta}+\frac{\lambda-h_t^{1/2}(\theta)}{1-\rho^2}\left(\rho\frac{C_t-C_t^D(\theta)}{\eta}-\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}\right)\right]\frac{\partial^3 h_t(\theta)}{\partial\theta\partial\theta'\partial\theta} \\
&+\frac{3}{2T}\sum_{t=1}^{T}\frac{\rho}{\eta(1-\rho^2)h_t(\theta)}\left(\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}+\lambda-h_t^{1/2}(\theta)\right)\frac{\partial^2 C_t^D(\theta)}{\partial\theta\partial\theta'}\frac{\partial h_t(\theta)}{\partial\theta}-\frac{3}{T}\sum_{t=1}^{T}\frac{1}{\eta^2(1-\rho^2)}\frac{\partial^2 C_t^D(\theta)}{\partial\theta\partial\theta'}\frac{\partial C_t^D(\theta)}{\partial\theta}-\frac{1}{T}\sum_{t=1}^{T}\frac{1}{\eta(1-\rho^2)}\left(\rho\frac{y_t-\mu_t(\theta)}{h_t^{1/2}(\theta)}-\frac{C_t-C_t^D(\theta)}{\eta}\right)\frac{\partial^3 C_t^D(\theta)}{\partial\theta\partial\theta'\partial\theta}.
\end{aligned}
\tag{B.9}
$$

Once again, consider the asymptotic Normality of the score and observed information matrix in (B.7) and (B.8), using Lemmas B.5 and B.6, respectively.

Lemma B.5. *The score given by (B.7) evaluated at* $\theta=\theta_0$ *is asymptotically Gaussian,* $H_{S+O+E0}^{-1/2}T^{1/2}\dfrac{\partial L_T^{S+O+E}(\theta)}{\partial\theta}\xrightarrow{D}N(0,I_3)$, *where*

$$
\begin{aligned}
H_{S+O+E0} &= E\left(T\frac{\partial L_T^{S+O+E}(\theta_0)}{\partial\theta}\frac{\partial L_T^{S+O+E}(\theta_0)}{\partial\theta'}\right)\\
&= E\left[\frac{1}{h_t^2(\theta_0)}\left(\frac{2-\rho^2+(\lambda-h_t^{1/2}(\theta_0))^2}{4(1-\rho^2)}\right)\frac{\partial h_t(\theta_0)}{\partial\theta}\frac{\partial h_t(\theta_0)}{\partial\theta'}\right]-\frac{\rho}{\eta(1-\rho^2)}E\left[\frac{\lambda-h_t^{1/2}(\theta_0)}{h_t(\theta_0)}\frac{\partial C_t^D(\theta_0)}{\partial\theta}\frac{\partial h_t(\theta_0)}{\partial\theta'}\right]+\frac{1}{\eta^2(1-\rho^2)}E\left[\frac{\partial C_t^D(\theta_0)}{\partial\theta}\frac{\partial C_t^D(\theta_0)}{\partial\theta'}\right].
\end{aligned}
$$

Proof of Lemma B.5. Evaluated at $\theta=\theta_0$ the form is given by

$$
\begin{aligned}
\frac{\partial L_T^{S+O+E}(\theta)}{\partial\theta} &= \frac{1}{T}\sum_{t=1}^{T}\left\{\left[\frac{-1}{2h_t(\theta)}\left(1-\frac{z_t^2}{1-\rho^2}+\frac{\rho z_t u_t}{1-\rho^2}+\frac{(\lambda-h_t^{1/2}(\theta))(\rho u_t-z_t)}{1-\rho^2}\right)\right]\frac{\partial h_t(\theta)}{\partial\theta}-\frac{\rho z_t-u_t}{\eta(1-\rho^2)}\frac{\partial C_t^D(\theta)}{\partial\theta}\right\}\\
&= \frac{1}{T}\sum_{t=1}^{T}v_t
\end{aligned}
$$

such that $E(v_t\,|\,F_{t-1})=0$, where $F_t=\sigma(z_t,u_t,z_{t-1},u_{t-1},\ldots)$. Applying the central limit theorem for martingale differences in Brown (1971), consider first

$$\frac{1}{T}\sum_{t=1}^{T}E(v_t^2\,|\,F_{t-1})\to H_{S+O+E0}$$

imply that $H_{S+O+E0}^{-1}\dfrac{1}{T}\sum_{t=1}^{T}E(v_t^2\,|\,F_{t-1})\to I_3$ in probability as $T\to\infty$, using the ergodic theorem. Hence we complete the proof of Lemma B.5.

Lemma B.6. *The observed information matrix given by (B.8) evaluated at* $\theta=\theta_0$ *converges in probability to* $-F_{S+O+E0}^{-1}\dfrac{\partial^2 L_T^{S+O+E}(\theta)}{\partial\theta\partial\theta'}\xrightarrow{p}I_3$, *where* $F_{S+O+E0}=-E\left(\dfrac{\partial^2 L_T^{S+O+E}(\theta_0)}{\partial\theta\partial\theta'}\right)$.

Proof of Lemma B.6. For $\theta=\theta_0$ the observed information is given by

$$
\begin{aligned}
&\frac{\partial^2 L_T^{S+O+E}(\theta)}{\partial\theta\partial\theta'} \\
&= -\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t^2(\theta)}\left(-1+\frac{2z_t^2}{1-\rho^2}-\frac{3\rho z_t u_t}{2(1-\rho^2)}-\frac{(\lambda-h_t^{1/2}(\theta))\left(3\rho u_t-5z_t-\lambda+h_t^{1/2}(\theta)\right)}{2(1-\rho^2)}\right)-\frac{(\rho u_t-z_t)}{2(1-\rho^2)h_t^{3/2}(\theta)}\right]\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'} \\
&+\frac{1}{T}\sum_{t=1}^{T}\frac{\rho}{\eta(1-\rho^2)h_t(\theta)}\left(z_t+\lambda-h_t^{1/2}(\theta)\right)\frac{\partial C_t^D(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}-\frac{1}{T}\sum_{t=1}^{T}\frac{1}{\eta^2(1-\rho^2)}\frac{\partial C_t^D(\theta)}{\partial\theta}\frac{\partial C_t^D(\theta)}{\partial\theta'} \\
&-\frac{1}{2T}\sum_{t=1}^{T}\left[\frac{1}{h_t(\theta)}\left(1-\frac{z_t^2}{1-\rho^2}+\frac{\rho z_t u_t}{1-\rho^2}+\frac{(\lambda-h_t^{1/2}(\theta))\left(\rho u_t-z_t\right)}{1-\rho^2}\right)\right]\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'}-\frac{1}{T}\sum_{t=1}^{T}\frac{\rho z_t-u_t}{\eta(1-\rho^2)}\frac{\partial^2 C_t^D(\theta)}{\partial\theta\partial\theta'}.
\end{aligned}
$$

The first term on the right-hand side converges by the ergodic theorem to $-E\left[\frac{1}{h_t^2(\theta)}\left(\frac{2-\rho^2+(\lambda-h_t^{1/2}(\theta))^2}{4(1-\rho^2)}\right)\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\right]$; second term on the right-hand side converges to $E\left[\frac{\rho}{\eta(1-\rho^2)}\frac{\lambda-h_t^{1/2}(\theta)}{h_t(\theta)}\frac{\partial C_t^D(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta'}\right]$; third term on the right-hand side converges to $-E\left[\frac{1}{\eta^2(1-\rho^2)}\frac{\partial C_t^D(\theta)}{\partial\theta}\frac{\partial C_t^D(\theta)}{\partial\theta'}\right]$; last two terms on the right-hand side converges in probability to zero. Hence we can complete the proof of Lemma B.6.

Lemma B.7. *There exists* $N(\theta_0)$ *defined in (B.4), for all* $1\le i,j,k\le 3$*, for which*

$\sup_{\theta\in N(\theta_0)}\left|\frac{\partial^3 L_T^{S+O+E}(\theta)}{\partial\theta_i\partial\theta_j\partial\theta_k}\right|\le g(\omega_L,\omega_U,\alpha_L,\alpha_U,\beta_L,\beta_U,T)\xrightarrow{a.s.}M<\infty$ *as* $T\to\infty$ *where* M *is constant.*

Without loss of generality, consider the case $\theta_i=\theta_j=\theta_k=\beta$. The next lemma shows that the individual terms of the third derivative $\left(\partial^3 L_T^{S+O+E}/\partial\beta^3\right)(\theta)$ in (B.9) are uniformly bounded in the neighborhood $N(\theta_0)$.

Lemma B.8. *With* $N(\theta_0)$ *defined in (B.4), for all t and* $i=1,2,3$*,*

$$\sup_{\theta\in N(\theta_0)} Y_{t,T}(\theta)\le Y_{t,T}, \qquad \textit{(B.10)}$$

$$\sup_{\theta\in N(\theta_0)} Y_{it,T}(\theta)\le Y_{it,T}, \qquad \textit{(B.11)}$$

$$\sup_{\theta\in N(\theta_0)} C_t^D(\theta) \le C_t^D, \tag{B.12}$$

and

$$\sup_{\theta\in N(\theta_0)} C_{it}^D(\theta) \le C_{it}^D, \tag{B.13}$$

where $Y_{it,T}(\theta) = \dfrac{\partial^i Y_{t,T}(\theta)}{\partial \beta^i}$, $C_{it}^D(\theta) = \dfrac{\partial^i C_t^D(\theta)}{\partial \beta^i}$,

$$Y_{t,T} = (T-t)r + \frac{1}{2}\sum_{s=t+1}^{T} H_{sU} + \sum_{s=t+1}^{T} H_{sU}^{1/2}\left[(z_s^Q)^2 + 1\right],$$

$$Y_{1t,T} = \frac{1}{2\omega_L}\sum_{s=t+1}^{T}(z_s^Q)^2 H_{1s},\ Y_{2t,T} = \frac{1}{2\omega_L}\sum_{s=t+1}^{T}(z_s^Q)^2 H_{2s} + \frac{1}{4\omega_L^{3/2}}\sum_{s=t+1}^{T}\left[(z_s^Q)^2+1\right]H_{1s}^2,$$

$$Y_{3t,T} = \frac{1}{2\omega_L}\sum_{s=t+1}^{T}(z_s^Q)^2 H_{3s} + \frac{3}{4\omega_L^{3/2}}\sum_{s=t+1}^{T}\left[(z_s^Q)^2+1\right]H_{2s}H_{1s} + \frac{3}{8\omega_L^{5/2}}\sum_{s=t+1}^{T}\left[(z_s^Q)^2+1\right]H_{1s}^3,$$

$$C_t^D = e^{-r(T-t)}E_t^Q\left[S_t\,\mathrm{e}^{Y_{t,T}} + K \mid S_T > K\right],\ C_{1t}^D = e^{-r(T-t)}E_t^Q\left[S_t\,\mathrm{e}^{Y_{t,T}}\,Y_{1t,T} \mid S_T > K\right],$$

$$C_{2t}^D = e^{-r(T-t)}E_t^Q\left[S_t\,\mathrm{e}^{Y_{t,T}}\left(Y_{1t,T}^2 + Y_{2t,T}\right) \mid S_T > K\right], \textit{ and}$$

$$C_{3t}^D = e^{-r(T-t)}E_t^Q\left[S_t\,\mathrm{e}^{Y_{t,T}}\left(Y_{1t,T}^3 + 3Y_{2t,T}Y_{1t,T} + Y_{3t,T}\right) \mid S_T > K\right].$$

Proof of Lemma B.8. Here, the first-, second-, and third-order derivatives of $Y_{t,T}(\theta)$ and $C_t^D(\theta)$ are

$$Y_{1t,T}(\theta) = \frac{1}{2}\sum_{s=t+1}^{T}\left(\frac{z_s^Q}{h_s^{1/2}(\theta)} - 1\right)h_{1s}(\theta),$$

$$Y_{2t,T}(\theta) = \frac{1}{2}\sum_{s=t+1}^{T}\left(\frac{z_s^Q}{h_s^{1/2}(\theta)} - 1\right)h_{2s}(\theta) - \frac{1}{4}\sum_{s=t+1}^{T}\frac{z_s^Q}{h_s^{3/2}(\theta)}h_{1s}^2(\theta),$$

$$Y_{3t,T}(\theta) = \frac{1}{2}\sum_{s=t+1}^{T}\left(\frac{z_s^Q}{h_s^{1/2}(\theta)} - 1\right)h_{3s}(\theta) - \frac{3}{4}\sum_{s=t+1}^{T}\frac{z_s^Q}{h_s^{3/2}(\theta)}h_{2s}(\theta)h_{1s}(\theta) + \frac{3}{8}\sum_{s=t+1}^{T}\frac{z_s^Q}{h_s^{5/2}(\theta)}h_{1s}^3(\theta),$$

$$C_{1t}^{D}(\theta)=e^{-r(T-t)}E_t^Q\left[S_t\,\mathrm{e}^{Y_{t,T}(\theta)}\,Y_{1t,T}(\theta)\mid S_T>K\right],$$

$$C_{2t}^{D}(\theta)=e^{-r(T-t)}E_t^Q\left[S_t\,\mathrm{e}^{Y_{t,T}(\theta)}\left(Y_{1t,T}^2(\theta)+Y_{2t,T}(\theta)\right)\mid S_T>K\right], \text{ and}$$

$$C_{3t}^{D}(\theta)=e^{-r(T-t)}E_t^Q\left[S_t\,\mathrm{e}^{Y_{t,T}(\theta)}\left(Y_{1t,T}^3(\theta)+3Y_{2t,T}(\theta)Y_{1t,T}(\theta)+Y_{3t,T}(\theta)\right)\mid S_T>K\right].$$

By $|x|\le x^2+1$, (B.4), and Lemma B.4, (B.10)-(B.13) follow.

Proof of Lemma B.7. Without loss of generality, consider the case $\theta_i=\theta_j=\theta_k=\beta$. Noting that by definition $y_t=r+\lambda h_t^{1/2}(\theta_0)-\frac{1}{2}h_t(\theta_0)+h_t^{1/2}(\theta_0)z_t$, the expression for $\left(\partial^3 L_T^{S+O+E}/\partial\beta^3\right)(\theta)$ in (B.7) implies that

$\left|\dfrac{\partial^3 L_T^{S+O+E}(\theta)}{\partial\beta^3}\right|\le\dfrac{1}{T}\sum_{t=1}^{T}w_t(\theta)$, where

$$\begin{aligned}
w_t(\theta)={}&\left\{\begin{array}{l}
\dfrac{1}{2h_t^3(\theta)}\left[\begin{array}{l}
2+\dfrac{6}{1-\rho^2}\dfrac{\left(\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t\right)^2}{h_t(\theta)}+\dfrac{15\rho}{4(1-\rho^2)}\dfrac{\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t}{h_t^{1/2}(\theta)}\dfrac{C_t^D(\theta_0)+C_t^D(\theta)+\eta u_t}{\eta}\\
+\dfrac{3(\lambda+h_t^{1/2}(\theta))}{4(1-\rho^2)}\left(5\rho\dfrac{C_t^D(\theta_0)+C_t^D(\theta)+\eta u_t}{\eta}+11\dfrac{\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t}{h_t^{1/2}(\theta)}+3(\lambda+h_t^{1/2}(\theta))\right)
\end{array}\right]\\
+\dfrac{3}{8(1-\rho^2)h_t^{5/2}(\theta)}\left(2\rho\dfrac{C_t^D(\theta_0)+C_t^D(\theta)+\eta u_t}{\eta}+3\dfrac{\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t}{h_t^{1/2}(\theta)}+\lambda+h_t^{1/2}(\theta)\right)
\end{array}\right\}h_{1t}^3(\theta)\\
&+\left\{\begin{array}{l}
\dfrac{3}{2h_t^2(\theta)}\left[\begin{array}{l}
1+\dfrac{2}{1-\rho^2}\dfrac{\left(\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t\right)^2}{h_t(\theta)}+\dfrac{3\rho}{2(1-\rho^2)}\dfrac{\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t}{h_t^{1/2}(\theta)}\dfrac{C_t^D(\theta_0)+C_t^D(\theta)+\eta u_t}{\eta}\\
+\dfrac{\lambda+h_t^{1/2}(\theta)}{2(1-\rho^2)}\left(3\rho\dfrac{C_t^D(\theta_0)+C_t^D(\theta)+\eta u_t}{\eta}+5\dfrac{\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t}{h_t^{1/2}(\theta)}+\lambda+h_t^{1/2}(\theta)\right)
\end{array}\right]\\
-\dfrac{3}{4(1-\rho^2)h_t^{3/2}(\theta)}\left(\rho\dfrac{C_t^D(\theta_0)+C_t^D(\theta)+\eta u_t}{\eta}+\dfrac{\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t}{h_t^{1/2}(\theta)}\right)
\end{array}\right\}h_{2t}(\theta)h_{1t}(\theta)\\
&+\dfrac{1}{2h_t(\theta)}\left[\begin{array}{l}
1+\dfrac{1}{1-\rho^2}\dfrac{\left(\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t\right)^2}{h_t(\theta)}+\dfrac{\rho}{1-\rho^2}\dfrac{\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t}{h_t^{1/2}(\theta)}\dfrac{C_t^D(\theta_0)+C_t^D(\theta)+\eta u_t}{\eta}\\
+\dfrac{\lambda+h_t^{1/2}(\theta)}{1-\rho^2}\left(\rho\dfrac{C_t^D(\theta_0)+C_t^D(\theta)+\eta u_t}{\eta}+\dfrac{\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t}{h_t^{1/2}(\theta)}\right)
\end{array}\right]h_{3t}(\theta)\\
&+\dfrac{3\rho}{4\eta(1-\rho^2)}\left[\dfrac{3}{h_t^2(\theta)}\left(\dfrac{\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t}{h_t^{1/2}(\theta)}+\lambda+h_t^{1/2}(\theta)\right)+\dfrac{1}{h_t^{3/2}(\theta)}\right]h_{1t}^2(\theta)C_{1t}^D(\theta)\\
&+\dfrac{3\rho}{2\eta(1-\rho^2)h_t(\theta)}\left(\dfrac{\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t}{h_t^{1/2}(\theta)}+\lambda+h_t^{1/2}(\theta)\right)h_{2t}(\theta)C_{1t}^D(\theta)\\
&+\dfrac{3\rho}{2\eta(1-\rho^2)h_t(\theta)}\left(\dfrac{\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t}{h_t^{1/2}(\theta)}+\lambda+h_t^{1/2}(\theta)\right)C_{2t}^D(\theta)h_{1t}(\theta)+\dfrac{3}{\eta^2(1-\rho^2)}C_{2t}^D(\theta)C_{1t}^D(\theta)\\
&+\dfrac{1}{\eta(1-\rho^2)}\left(\rho\dfrac{\mu_t(\theta)+\mu_t(\theta_0)+h_t^{1/2}(\theta_0)z_t}{h_t^{1/2}(\theta)}+\dfrac{C_t^D(\theta_0)+C_t^D(\theta)+\eta u_t}{\eta}\right)C_{3t}^D(\theta).
\end{aligned}$$

From Lemmas B.4 and B.8, we have there exists a neighborhood $N(\theta_0)$ for which $\sup_{\theta \in N(\theta_0)} \left| \frac{\partial^3 L_T^{S+O+E}(\theta)}{\partial \beta^3} \right| \le \frac{1}{T}\sum_{t=1}^{T} w_t$, where w_t is stationary and has finite moment $Ew_t = M < \infty$ such that $\frac{1}{T}\sum_{t=1}^{T} w_t \xrightarrow{a.s.} M$ by the ergodic theorem, which ends the proof of Lemma B.7.

Proof of Theorem B.2. From Lemma B.7, we have that

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \frac{\partial L_T^{S+O+E}(\hat{\theta}_T^{S+O+E})}{\partial \theta} \approx \frac{\partial L_T^{S+O+E}(\theta_0)}{\partial \theta} + \frac{\partial^2 L_T^{S+O+E}(\theta_0)}{\partial \theta \partial \theta'}(\hat{\theta}_T^{S+O+E} - \theta_0)$$

and

$$\hat{\theta}_T^{S+O+E} - \theta_0 \approx -\left(\frac{\partial^2 L_T^{S+O+E}(\theta_0)}{\partial \theta \partial \theta'} \right)^{-1} \frac{\partial L_T^{S+O+E}(\theta_0)}{\partial \theta}.$$

Combining with Lemmas B.5 and B.6, we complete the proof of Theorem B.2.

References

Aït-sahalia, Y., Kimmel, R., 2007. Maximum likelihood estimation of stochastic volatility models. Journal of Financial Economics 83, 413-452.

Bakshi, G., Cao, C., Chen, Z., 1997. Empirical performance of alternative option pricing models. Journal of Finance 52, 2003-2049.

Bollerslev, T., 1986. Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31, 307-327.

Bossaerts, P., Hillion, P., 1997. Local parametric analysis of hedging in discrete time. Journal of Econometrics 81, 243-272.

Bougerol, P., Picard, N., 1992. Stationarity of GARCH processes and of some nonnegative time series. Journal of Econometrics 52, 115-127.

Brown, B.M., 1971. Martingale Central Limit Theorems. Annals of Mathematical Statistics 42, 59-66.

Chernov, M., Ghysels, E., 2000. Towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation. Journal of Financial Economics 56, 407–458.

Christoffersen, P., Jacobs, K., 2004. Which GARCH model for options valuation? Management Science 50, 1204-1221.

Duan, J.C., 1995. The GARCH options pricing model. Mathematical Finance 5, 13-32.

Duffie, D., Pedersen, L.H., Singleton, K.J., 2003. Modeling Sovereign Yield Spreads: A

Case Study of Russian Debt. Journal of Finance 58, 119-159.

Engle, R.F., Mustafa, C., 1992. Implied ARCH models from options prices. Journal of Econometrics 52, 289-311.

Engle, R.F., Ng, V.K., 1993. Measuring and Testing the Impact of News on Volatility. Journal of Finance 48, 1749-1778.

Engle, R.F., Rosenberg, J.V., 1995. GARCH Gamma. Journal of Derivatives 4, 47-59.

Eraker, B., Johannes, M.S., Polson, N.G., 2003. The impact of jumps in returns and volatility, Journal of Finance 53, 1269-1300.

Eraker, B., 2004. Do stock prices and volatility jump? Reconciling evidence from spot and options prices. Journal of Finance 59, 1367-1403.

Heston, S., 1993. A closed form solution for options with stochastic volatility with applications to bond and currency options. Review of Financial Studies 6, 327-343.

Heston, S., Nandi, S., 2000. A Closed-Form GARCH Option Valuation Model. Review of Financial Studies 13, 585-625.

Jacquier, E., Jarrow, R., 2000. Bayesian analysis of contingent claim model error. Journal of Econometrics 94, 145–180.

Johannes, M., Polson, N., Stroud, J., 2009. Optimal filtering of jump-diffusions: extracting latent states from asset prices. Review of Financial Studies 22, 2759-2799.

Jones, C.S., 2003. The dynamics of stochastic volatility: evidence from underlying and options markets. Journal of Econometrics 116, 181-224.

Lee, S.W., Hansen, B.E., 1994. Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. Econometric Theory 10, 29-52.

Lehar, A., Scheicher, M., Schittenkopf, C., 2002. GARCH vs. stochastic volatility: options pricing and risk management. Journal of Banking & Finance 26, 323-345.

Lumsdaine, R.L., 1996. Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models. Econometrica 64, 575-596.

Nelson, D.B., 1990. Stationary and Persistence in the GARCH(1,1) Model. Econometric Theory 6, 318-334.

Pan, J., 2002. The jump-risk premia implicit in options: evidence from an integrated time-series study. Journal of Financial Economics 63, 3-50.

Platen, E., Schweizer, M., 1994. On smile and skewness. Discussion Paper Serie B 302, University of Bonn, Germany.

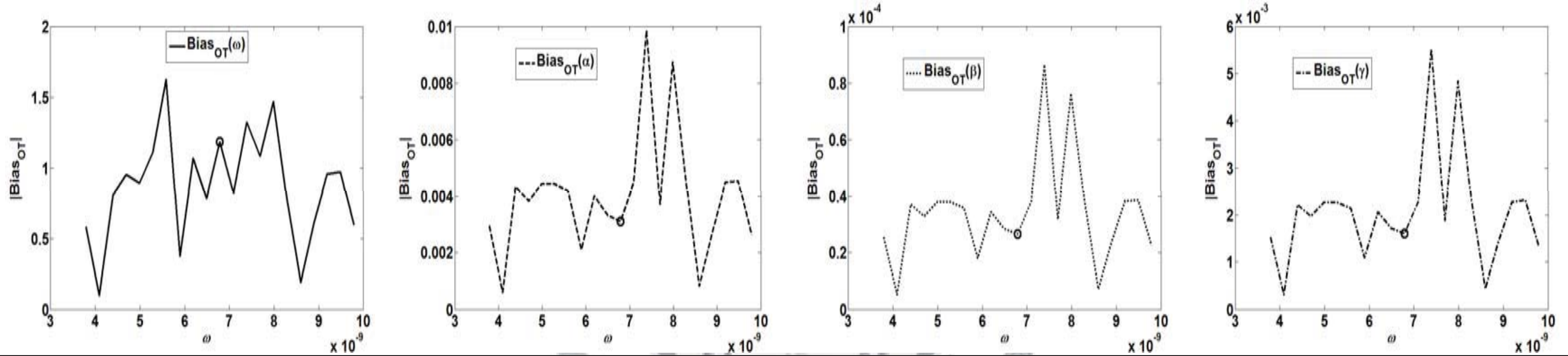
Renault, E., 1997. Econometric Models of Option Pricing Errors. In Advances in Economics and Econometrics, Seventh World Congress, edited by D.M. Kreps and K.F. Wallis, Econometric Society Monographs, Cambridge University Press, 223-278.

Renault, E., Touzi, N., 1996. Option Hedging and Implied Volatilities in a Stochastic Volatility Model. Mathematical Finance 6, 279-302.

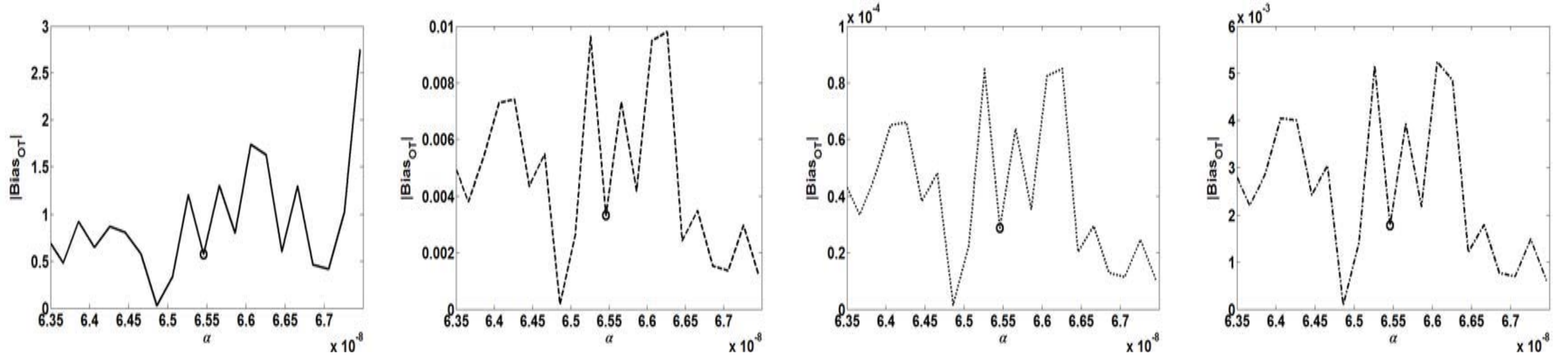
Figure 1: |*Bias* $_{OT}$| Graphs

Presented are graphs of $|Bias_{OT}|$ for ω, α, β, and γ. The bias value uses relative error (the absolute value of estimate bias divided by true value). True parameters are $(\lambda, \omega_0, \alpha_0, \beta_0, \gamma_0) = (0.1746, 6.792\times10^{-9}, 6.546\times10^{-8}, 0.9914, 351.945)$. In panel a, α_0, β_0, and γ_0 are given and ω is varied; in panel b, ω_0, β_0, and γ_0 are given and α is varied; in panel c, ω_0, α_0, and γ_0 are given and β is varied; and in panel d, ω_0, α_0, and β_0 are given and γ is varied. Graphs on the left to right of each panel show *Bias* for ω, α, β, and γ. True parameters calibrations are circled.

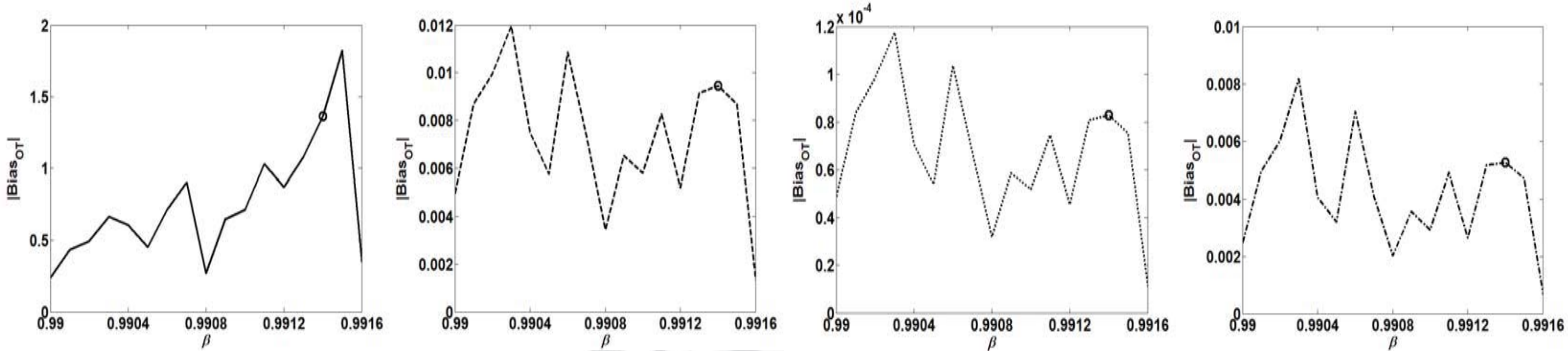
Panel A: ω $_0$ Unknown



Panel B: α $_0$ Unknown



Panel C: β_0 Unknown



Panel D: γ_0 Unknown

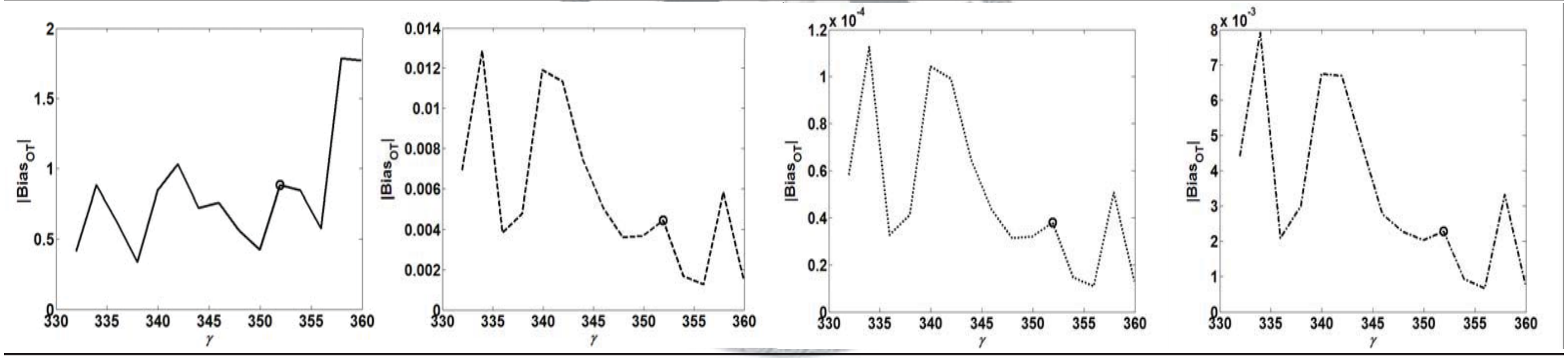
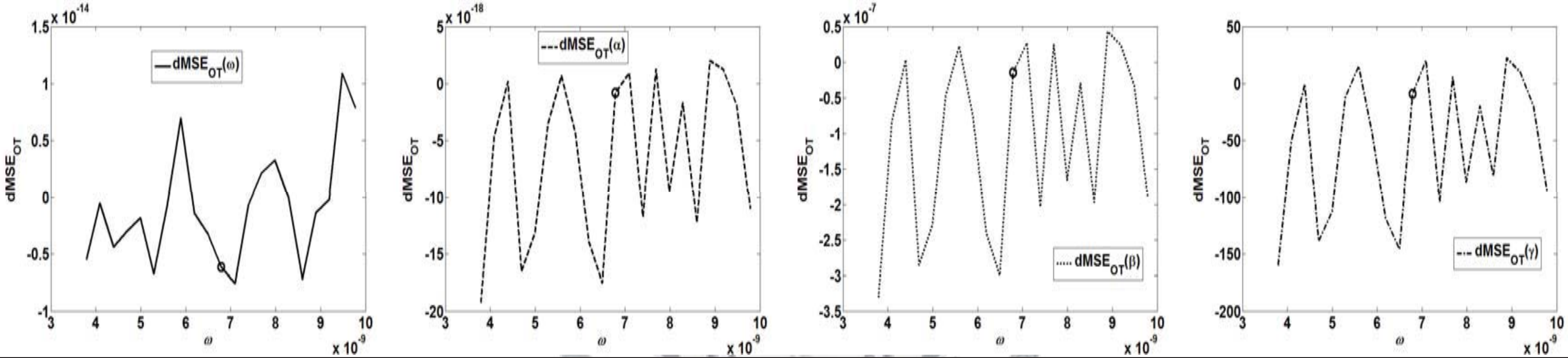


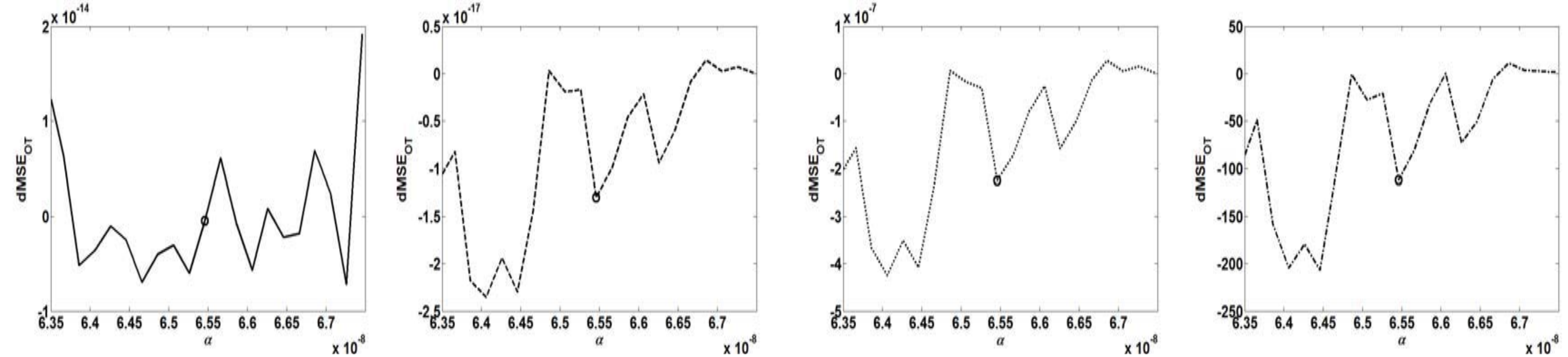
Figure 2: $dMSE_{OT}$ Graphs

Presented are graphs of $dMSE_{OT}$ for ω, α, β, and γ. True parameters are $(\lambda, \omega_0, \alpha_0, \beta_0, \gamma_0) = (0.1746, 6.792\times10^{-9}, 6.546\times10^{-8}, 0.9914, 351.945)$. In panel a, α_0, β_0, and γ_0 are given and ω is varied; in panel b, ω_0, β_0, and γ_0 are given and α is varied; in panel c, ω_0, α_0, and γ_0 are given and β is varied; and in panel d, ω_0, α_0, and β_0 are given and γ is varied. Graphs on the left to right of each panel show $dMSE_{OT}$ for ω, α, β, and γ. True parameters calibrations are circled.

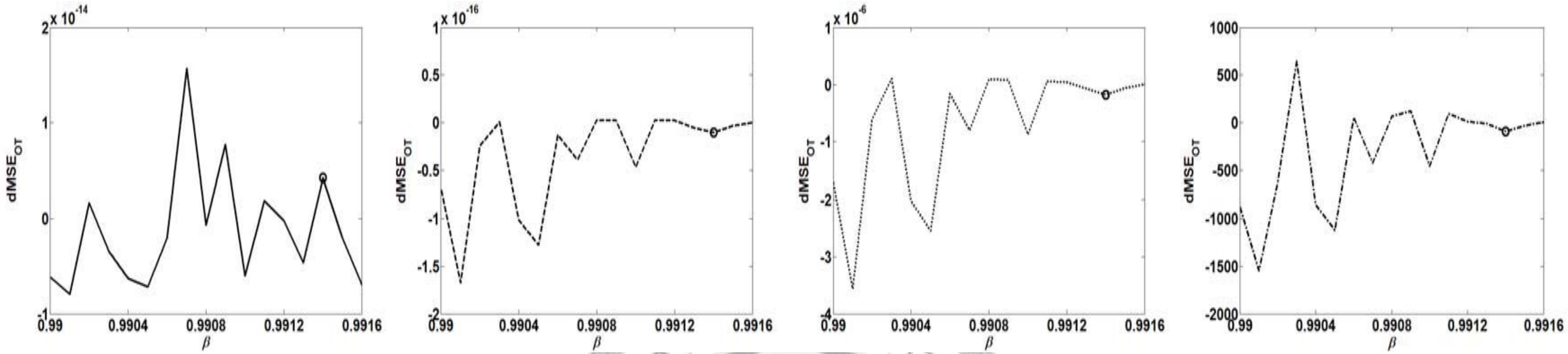
Panel A: ω_0 Unknown



Panel B: α_0 Unknown



Panel C: β_0 Unknown



Panel D: γ_0 Unknown

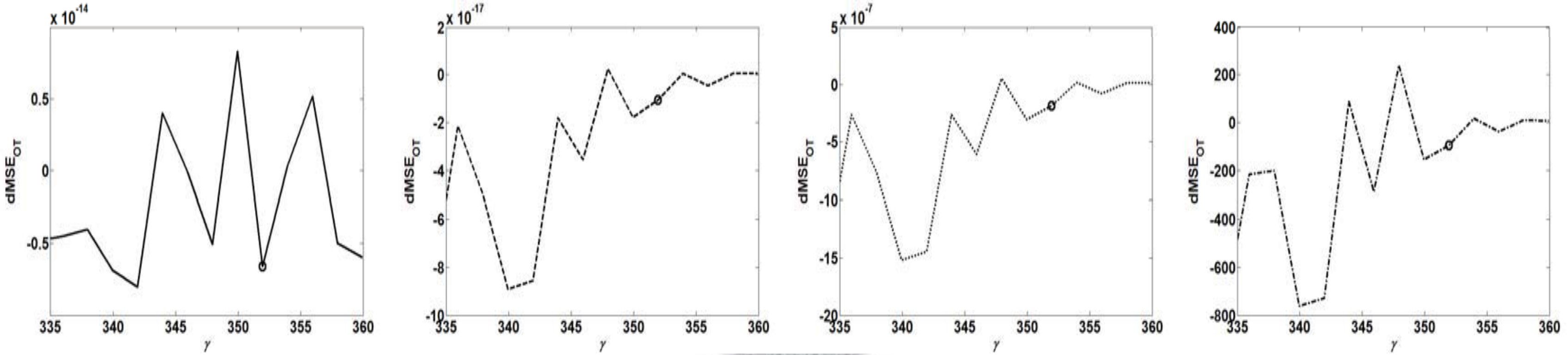
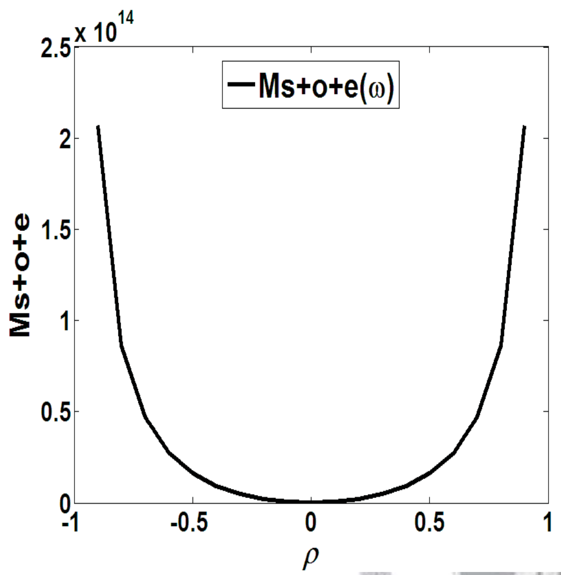


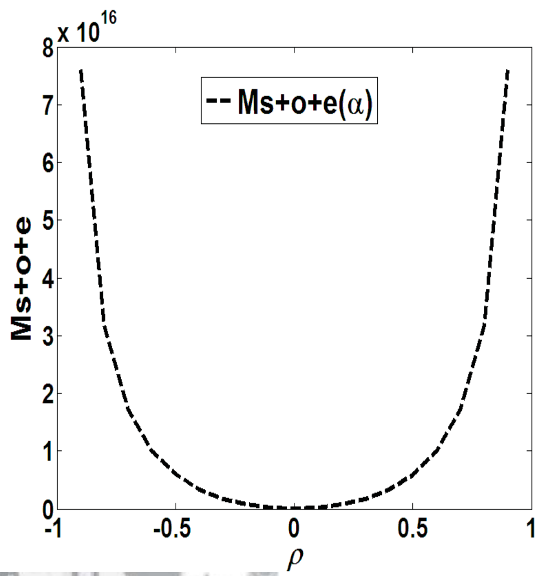
Figure 3: M_{S+O+E} Graphs with Different ρ

Presented are graphs of M_{S+O+E} for ω, α, β, and γ with ρ from -0.9 to 0.9. True parameters are $(\lambda, \omega_0, \alpha_0, \beta_0, \gamma_0) = (0.1746, 6.792\times10^{-9}, 6.546\times10^{-8}, 0.9914, 351.945)$. In panel a, α_0, β_0, and γ_0 are given and ω is unknown; in panel b, ω_0, β_0, and γ_0 are given and α is unknown; in panel c, ω_0, α_0, and γ_0 are given and β is unknown; and in panel d, ω_0, α_0, and β_0 are given and γ is unknown.

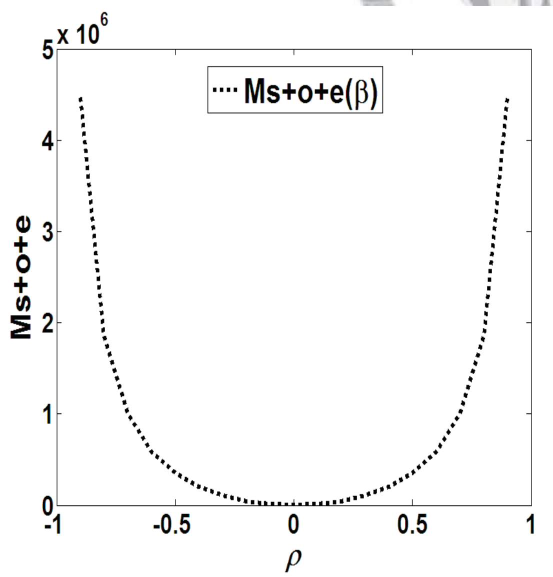
Panel A: ω_0 Unknown



Panel B: α_0 Unknown



Panel C: β_0 Unknown



Panel D: γ_0 Unknown

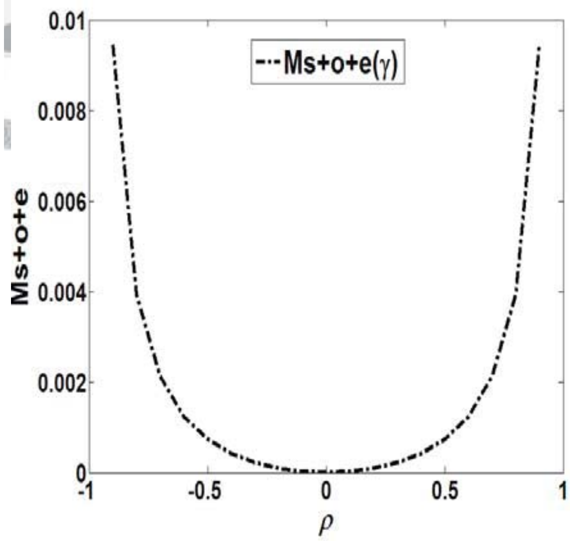
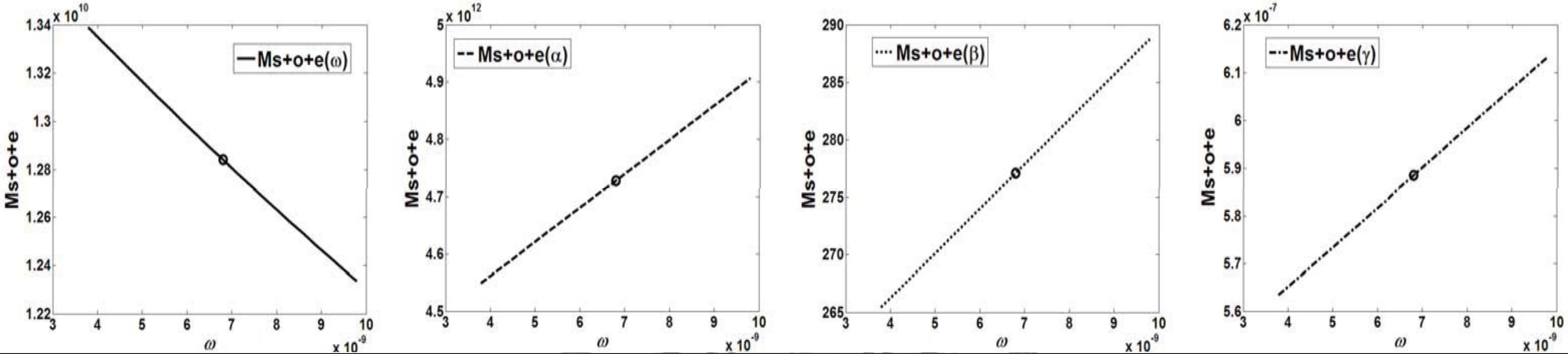


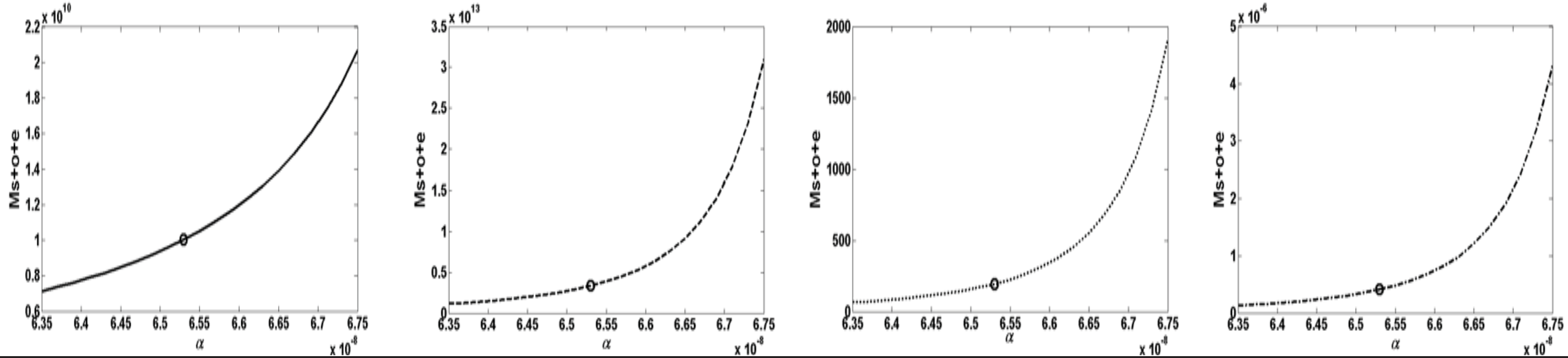
Figure 4: M_{S+O+E} Graphs with $\rho = 0$

Presented are graphs of M_{S+O+E} for ω, α, β, and γ with $\rho = 0$. True parameters are $(\lambda, \omega_0, \alpha_0, \beta_0, \gamma_0) = (0.1746, 6.792\times10^{-9}, 6.546\times10^{-8}, 0.9914, 351.945)$. In panel a, α_0, β_0, and γ_0 are given and ω is varied; in panel b, ω_0, β_0, and γ_0 are given and α is varied; in panel c, ω_0, α_0, and γ_0 are given and β is varied; and in panel d, ω_0, α_0, and β_0 are given and γ is varied. Graphs on the left to right of each panel show M_{S+O+E} for ω, α, β, and γ. True parameters calibrations are circled.

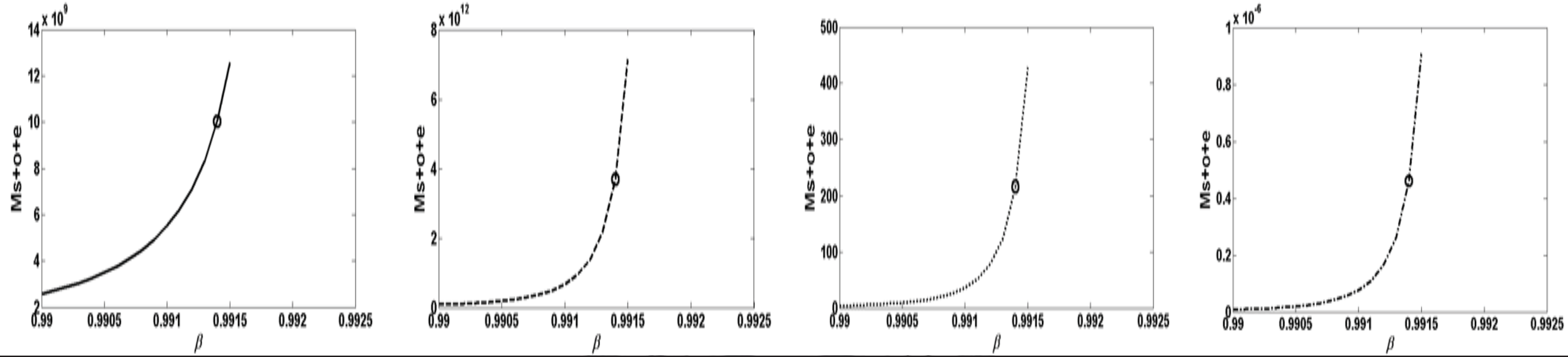
Panel A: ω_0 Unknown



Panel B: α_0 Unknown



Panel C: β_0 Unknown



Panel D: γ_0 Unknown

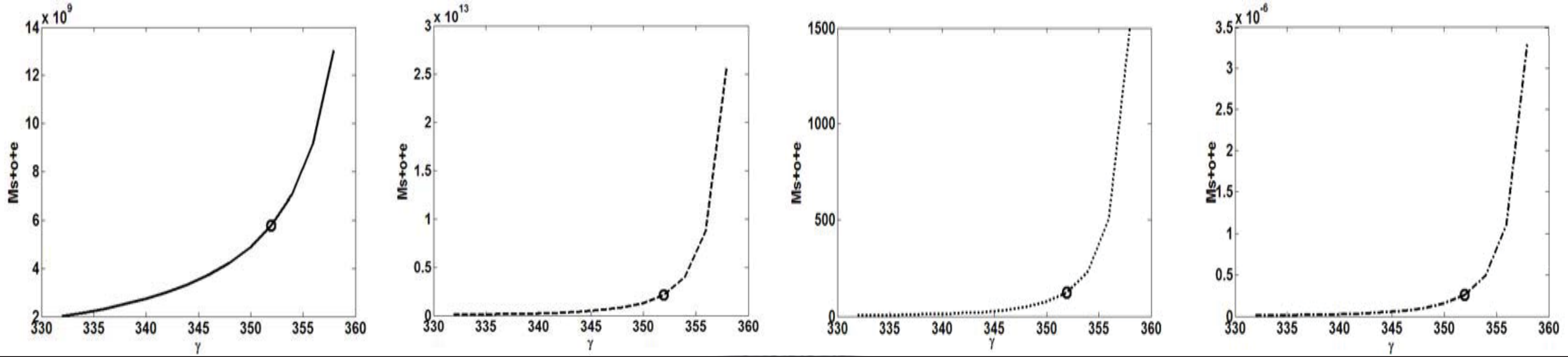


Figure 5: M_{S+O+E} Graphs with Different ρ

Presented are graphs of M_{S+O+E} for ω, α, and β with ρ from -0.9 to 0.9. True parameters are (λ, ω_0, α_0, β_0) =(0.0116, 9.228×10^{-7}, 0.068, 0.9248). In panel a, α_0 and β_0 are given and ω is unknown; in panel b, ω_0 and β_0 are given and α is unknown; in panel c, ω_0 and α_0 are given and β is unknown.

Panel A: ω_0 Unknown | *Panel B: α_0 Unknown* | *Panel C: β_0 Unknown*

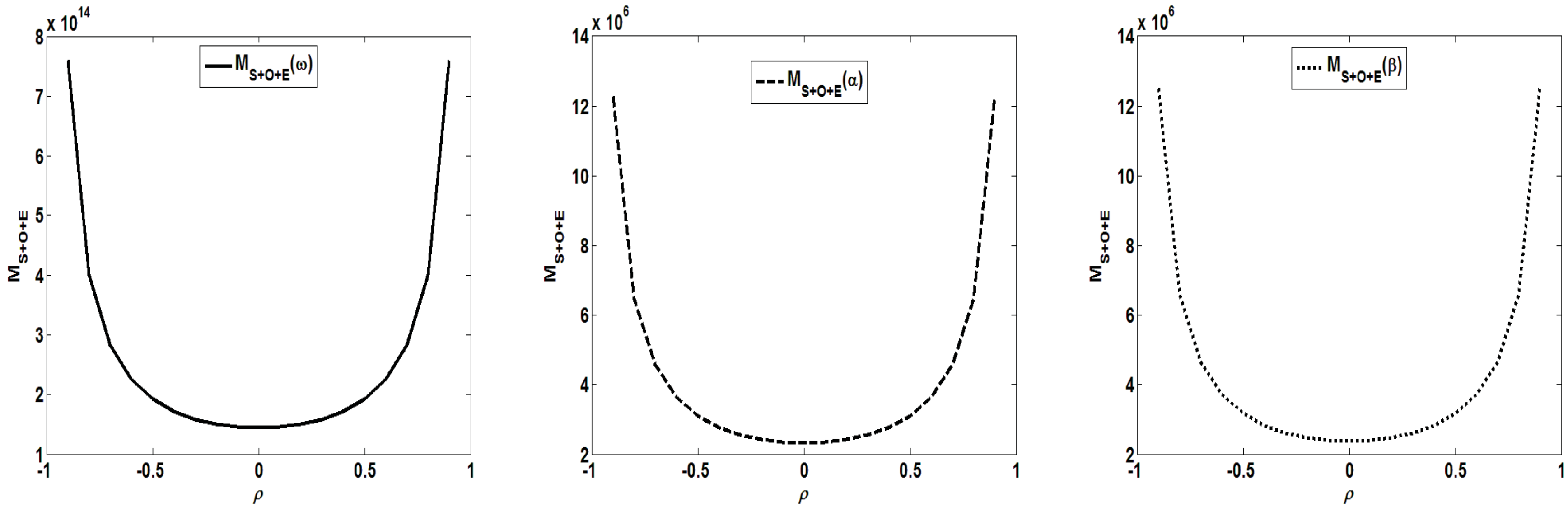
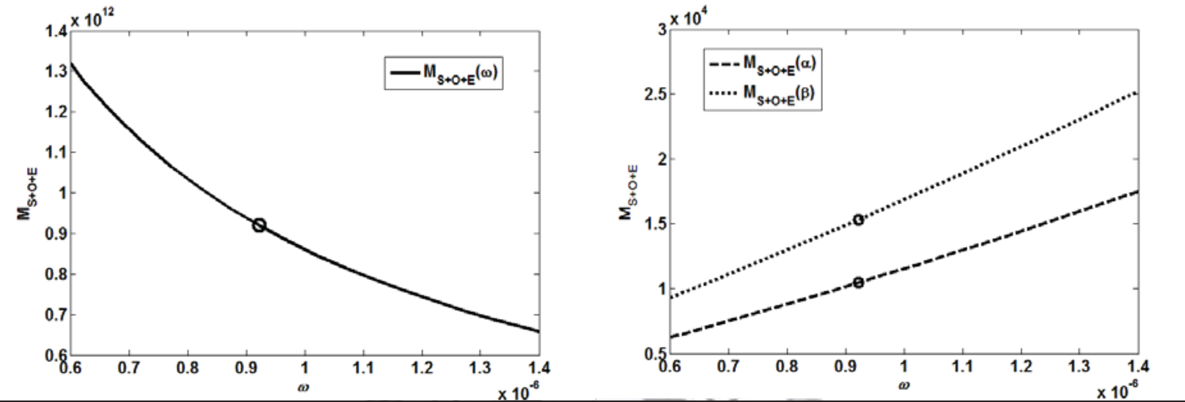


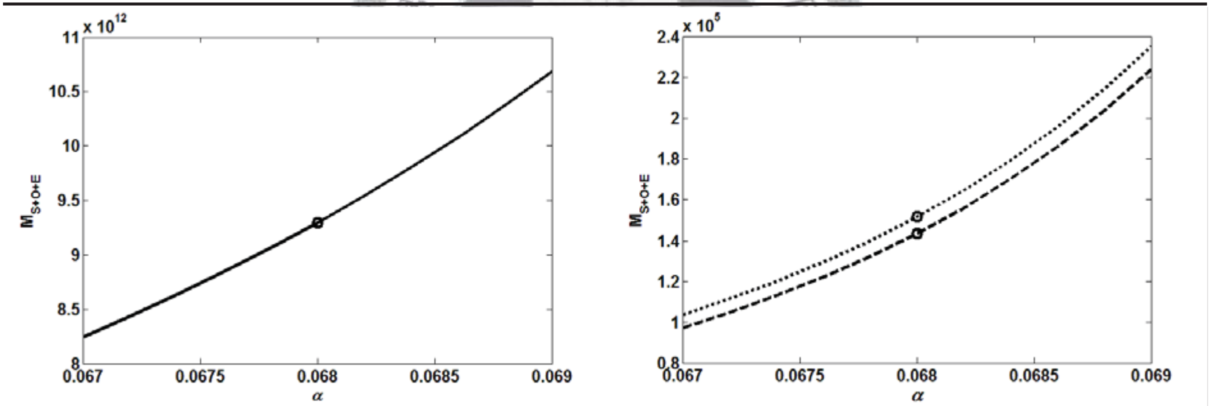
Figure 6: M_{S+O+E} Graphs with $\rho = 0$

Presented are graphs of M_{S+O+E} for ω, α, and β with $\rho = 0$. True parameters are (λ, ω_0, α_0, β_0) = (0.0116, 9.228×10^{-7}, 0.068, 0.9248). In panel a, α_0 and β_0 are given and ω is varied; in panel b, ω_0 and β_0 are given and α is varied; and in panel c, ω_0 and α_0 are given and β is varied. Graphs on the left of each panel show M_{S+O+E} for ω while those on the right show M_{S+O+E} for α and β. True parameters calibrations are circled.

Panel A: ω_0 Unknown



Panel B: α_0 Unknown



Panel C: β_0 Unknown

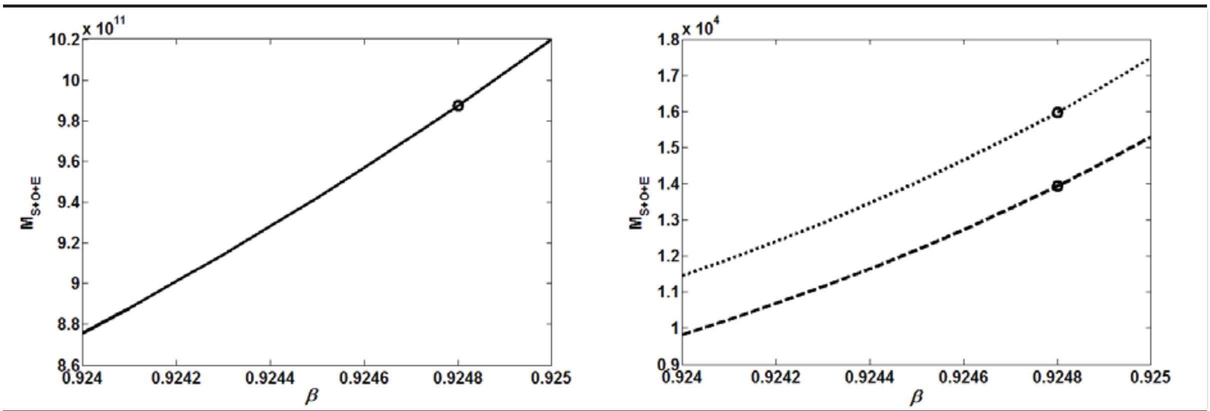


Table 1: Simulated Parameter Estimates and Errors

This table shows the bias, standard deviations (SD), and mean square errors (MSE) of the parameter estimates for different exercise prices and estimation methods. 1,000 Monto Carlo iterations of 30 daily observations are run where true parameters values are (λ, ω_0, α_0, β_0, γ_0) = (0.1746, 6.792×10^{-9}, 6.546×10^{-8}, 0.9914, 351.945). ST uses stock data only, OT uses options data only, and S+O+E uses both stock and options data includes an error term and assumes that $\eta = 1$, $\rho = 0$. The riskless rate is fixed at 5%. Values shown for ω are adjusted by 10^9 for Bias and SD, and 10^{17} for MSE. Values for α are adjusted by 10^{10} for Bias and SD, and 10^{19} for MSE. Values for β are adjusted by 10^4 for Bias and SD, and 10^8 for MSE. Values for γ are adjusted by 1 for Bias, SD, and MSE

	ST	OT			S+O+E		
S_0 / K		0.9	1.0	1.1	0.9	1.0	1.1
ω unknown							
Bias	4.955	4.700	4.930	5.355	4.355	2.175	1.230
SD	1.628	1.810	1.813	1.366	2.134	2.103	1.258
MSE	2.718	2.534	2.756	3.052	2.347	0.911	0.308
α unknown							
Bias	3.744	5.528	5.256	3.276	2.360	1.120	0.604
SD	2.919	3.629	3.510	1.885	1.153	1.097	0.600
MSE	2.245	4.360	3.982	1.425	0.689	0.244	0.072
β unknown							
Bias	0.570	1.380	1.030	0.840	0.440	0.081	0.030
SD	6.237	7.886	7.029	6.470	4.989	2.815	1.714
MSE	0.710	2.520	1.550	1.120	0.440	0.086	0.030
γ unknown							
Bias	1.164	1.620	1.300	1.328	0.828	0.276	0.124
SD	0.722	0.589	0.671	0.694	0.654	0.424	0.218
MSE	1.870	2.968	2.136	2.240	1.109	0.254	0.062

Table 2: Risk Management Metrics under Different Data Specifications

Using 10,000 simulation iterations and 12 months of stock and options data, we calculate Black-Scholes (B-S) and GARCH options risk metrics. Presented first are deltas (first partial with respect to stock price), then gammas (the second partial). Each is calculated using stock data only in columns (I) ST and both stock and options data including an error term in columns (II) S+O+E. Difference columns (III) calculate quotients I / II - 1 for respective entries. For GARCH specifications, $h_0 = (\omega + \alpha) / (1 - \alpha\gamma^2 - \beta)$.

		(I) ST		(II) S+O+E		(III) I / II - 1	
	S_0/K	B-S	GARCH	B-S	GARCH	B-S	GARCH
Delta							
T = 30	0.90	0.0099	0.0101	0.0618	0.0634	-0.8398	-0.8412
	0.95	0.1308	0.1371	0.2320	0.2455	-0.4364	-0.4415
	1.00	0.5089	0.5305	0.5134	0.5413	-0.0086	-0.0200
	1.05	0.8669	0.8749	0.7769	0.7909	0.1158	0.1062
	1.10	0.9842	0.9867	0.9274	0.9345	0.0613	0.0559
T = 90	0.90	0.0935	0.1000	0.1976	0.2247	-0.5267	-0.5549
	0.95	0.2668	0.2871	0.3505	0.3911	-0.2386	-0.2660
	1.00	0.5155	0.5449	0.5231	0.5657	-0.0146	-0.0367
	1.05	0.7478	0.7752	0.6839	0.7244	0.0935	0.0701
	1.10	0.8975	0.9092	0.8104	0.8396	0.1074	0.0829
T = 180	0.90	0.1826	0.2083	0.2877	0.3392	-0.3653	-0.3858
	0.95	0.3399	0.3749	0.4088	0.4676	-0.1685	-0.1982
	1.00	0.5219	0.5580	0.5327	0.5884	-0.0203	-0.0517
	1.05	0.6913	0.7218	0.6478	0.7002	0.0672	0.0308
	1.10	0.8221	0.8480	0.7463	0.7932	0.1016	0.0691
Gamma							
T = 30	0.90	0.0011	0.0012	0.0034	0.0036	-0.6760	-0.6631
	0.95	0.0083	0.0086	0.0080	0.0082	0.0410	0.0455
	1.00	0.0148	0.0149	0.0099	0.0099	0.4960	0.4968
	1.05	0.0076	0.0074	0.0071	0.0069	0.0776	0.0780
	1.10	0.0013	0.0012	0.0031	0.0030	-0.5721	-0.5814
T = 90	0.90	0.0040	0.0042	0.0044	0.0047	-0.1010	-0.1127
	0.95	0.0074	0.0076	0.0056	0.0058	0.3265	0.3241
	1.00	0.0086	0.0086	0.0057	0.0057	0.4969	0.4955
	1.05	0.0065	0.0064	0.0049	0.0047	0.3418	0.3422
	1.10	0.0035	0.0034	0.0035	0.0034	-0.0138	0.0057
T = 180	0.90	0.0045	0.0047	0.0038	0.0040	0.1611	0.1714
	0.95	0.0059	0.0061	0.0042	0.0043	0.4104	0.4081
	1.00	0.0061	0.0060	0.0040	0.0040	0.4983	0.4984
	1.05	0.0051	0.0049	0.0036	0.0035	0.4186	0.4144
	1.10	0.0036	0.0034	0.0030	0.0028	0.2162	0.2236

Table 3: Simulated Parameter Estimates and Errors

This table shows the bias, standard deviations (SD), and mean square errors (MSE) of the parameter estimates for different exercise prices and estimation methods. 1,000 Monto Carlo iterations of 30 daily observations are run where true parameters values are (λ, ω_0, α_0, β_0) = (0.0116, 9.228×10^{-7}, 0.068, 0.9248). ST uses stock data only, S+O+E uses both stock and options data includes an error term and assumes that $\eta = 1$, $\rho = 0$. The riskless rate is fixed at 5%. Values shown for ω are adjusted by 10^7 for Bias and SD, and 10^{13} for MSE. Values for α and β are adjusted by 10^3 for Bias and SD, and 10^5 for MSE.

	ST	S+O+E		
S_0 / K		0.9	1.0	1.1
ω unknown				
Bias	3.398	2.816	2.181	1.673
SD	2.832	2.753	2.485	2.038
MSE	1.956	1.550	1.093	0.695
α unknown				
Bias	2.651	2.019	1.087	0.430
SD	2.532	2.517	1.985	1.209
MSE	1.343	1.041	0.512	0.164
β unknown				
Bias	2.634	2.183	1.108	0.314
SD	2.337	2.363	1.865	0.987
MSE	1.239	1.034	0.470	0.107

Table 4: Risk Management Metrics under Different Data Specifications

Using 10,000 simulation iterations and 12 months of stock and options data, we calculate Black-Scholes (B-S) and GARCH options risk metrics. Presented first are deltas (first partial with respect to stock price), then gammas (the second partial). Each is calculated using stock data only in columns (I) ST and both stock and options data including an error term in columns (II) S+O+E. Difference columns (III) calculate quotients I / II - 1 for respective entries. For GARCH specifications, $h_0 = \omega / (1 - \alpha - \beta)$.

		(I) ST		(II) S+O+E		(III) I / II - 1	
	S_0/K	B-S	GARCH	B-S	GARCH	B-S	GARCH
Delta							
T = 30	0.80	0.0021	0.0060	0.0077	0.0135	-0.7312	-0.5582
	0.90	0.0911	0.0839	0.1312	0.1211	-0.3058	-0.3067
	1.00	0.5153	0.5106	0.5180	0.5191	-0.0052	-0.0164
	1.10	0.8997	0.9127	0.8643	0.8741	0.0410	0.0443
	1.20	0.9921	0.9909	0.9804	0.9787	0.0119	0.0125
T = 90	0.80	0.0535	0.0533	0.0889	0.0823	-0.3983	-0.3521
	0.90	0.2339	0.2179	0.2759	0.2470	-0.1523	-0.1177
	1.00	0.5265	0.5352	0.5312	0.5354	-0.0088	-0.0004
	1.10	0.7833	0.8000	0.7540	0.7780	0.0388	0.0284
	1.20	0.9248	0.9280	0.8931	0.9072	0.0354	0.0230
T = 180	0.80	0.1372	0.1186	0.1847	0.1649	-0.2570	-0.2809
	0.90	0.3205	0.2971	0.3574	0.3226	-0.1033	-0.0790
	1.00	0.5374	0.5383	0.5441	0.5451	-0.0121	-0.0124
	1.10	0.7261	0.7500	0.7059	0.7300	0.0286	0.0274
	1.20	0.8563	0.8764	0.8250	0.8457	0.0379	0.0363
Gamma							
T = 30	0.80	0.0002	0.0001	0.0005	0.0008	-0.6368	-0.9194
	0.90	0.0040	0.0036	0.0044	0.0046	-0.0941	-0.2179
	1.00	0.0087	0.0097	0.0074	0.0084	0.1775	0.1644
	1.10	0.0035	0.0034	0.0037	0.0035	-0.0499	-0.0110
	1.20	0.0004	0.0004	0.0007	0.0009	-0.4630	-0.5846
T = 90	0.80	0.0017	0.0020	0.0021	0.0025	-0.2040	-0.2050
	0.90	0.0043	0.0047	0.0040	0.0041	0.0795	0.1471
	1.00	0.0050	0.0061	0.0042	0.0046	0.1781	0.3423
	1.10	0.0033	0.0035	0.0031	0.0032	0.0968	0.0850
	1.20	0.0015	0.0015	0.0016	0.0012	-0.0932	0.1771
T = 180	0.80	0.0024	0.0027	0.0025	0.0028	-0.0308	-0.0358
	0.90	0.0035	0.0034	0.0031	0.0033	0.1287	0.0394
	1.00	0.0035	0.0050	0.0030	0.0045	0.1791	0.1153
	1.10	0.0027	0.0024	0.0024	0.0028	0.1377	-0.1293
	1.20	0.0017	0.0015	0.0016	0.0011	0.0345	0.4385