EQUILIBRIUM-BASED VOLATILITY MODELS OF THE MARKET PORTFOLIO RATE OF RETURN

David Feldman[a] and Xin Xu[b,c]

Latest revision: July 4, 2014

Abstract. Volatility models of the market portfolio's return are central to financial risk management. Within an equilibrium framework, we introduce an implementation method and study two families of such models. One is deterministic volatility, represented by current popular models. Another is in the "constant elasticity of variance" family, in which we propose new models. Theoretically, we show that, together with constant expected returns, the latter family tends to have better ability to forecast. Empirically, our proposed models, while as easy to implement as the popular ones, outperform them in three out-of-sample forecast evaluations of different time periods, by standard predictability criteria. This is true particularly during high-volatility periods, whether the market rises or falls.

JEL Codes: G17, G12, C58

Key Words: Market Risk, Volatility Model, Systematic Risk, Market Portfolio, Predictive Power, Equilibrium, GARCH, RiskMetrics, Piecewise Constant Volatility, Constant Elasticity of Variance

---

[a] Banking and Finance, Australian School of Business, The University of New South Wales, UNSW Sydney 2052, Australia; telephone: +61-(0)2-6385-5748, email: d.feldman@unsw.edu.au.

[b] Commonwealth Bank of Australia, 201 Sussex Street, Sydney, Australia, email: xinxu75@gmail.com. This paper was written when Xin Xu was at University of New South Wales as part of his PhD thesis, Xu (2013). The views expressed in this paper are those of the author and do not necessarily reflect those of the employer.

[c] We thank Tony Berrada, Avi Bick, Wayne Ferson, Giampaolo Gabbi, Dan Galai, Hayette Gatfaoui, Ufuk Gucbilmez, Robert Kohn, Delphine Lautier, Jaehoon Lee, Frederik Lundtofte, Valentyn Panchenko, Linda Pesante, Robert Faff, Haim Reisman, Jonathan Reeves, Konark Saxena, Tzu Tai, and Jing Xu for helpful discussions and seminar participants at UNSW School of Banking and Finance, University of Zurich, Australasian Finance and Banking Conference, French Finance Association Conference, International Risk Management Conference, Pacific Basin Finance Economics Accounting and Management Conference, the Bicentenary Conference of the Italian Academy of Management (AIDEA), North American Winter Meetings of the Econometric Society joint with the American Economic Association, International Finance and Banking Society Conference, Portuguese Finance Network International Conference, and The Econometric Society Australasian Meeting. The SSRN link for this paper is http://ssrn.com/abstract=2138653.

1 Introduction

We introduce an implementation method of equilibrium-based volatility models for the rate of return of the market portfolio.[1] Market portfolio's return volatility models are increasingly central to financial risk management. Such models, for example, are important in determining Value at Risk or similar measures which, perhaps, all financial institutions use [see, for example, Wiener (1999), Crouhy, Galai and Mark (2001), Allen, Boudoukh and Saunders (2003), Jorion (2007)]. The models are, of course, also relevant for pricing derivatives.

Why is it important and interesting to constrain the search of volatility models to equilibrium-consistent ones? Is it not just a handicap imposed for cosmetic reasons? Peacocks signal excess strength by channeling energy from building fighting prowess to building large heavy tails, which further handicap their fighting ability. Fast gazelles, facing predators, signal their quickness by stotting rather than running away (as slower gazelles do). Predators, in turn, do not chase stotting gazelles, preserving energy, as the stotting gazelles do, by avoiding fruitless chases.[2]

While we would not argue with successful "black boxes," we advocate building ones that are equilibrium consistent. This choice, we suppose, eliminates some irrelevant and inconsistent models. We hope that this paper demonstrates that choosing equilibrium-consistent models is more similar to stotting gazelles' equilibria, where both predators and potential prey save energy, rather than to peacocks/peahens equilibria where substantial energy is spent on building handicaps.

The major role of volatility models is perceived to be prediction of market movements, particularly big ones [see, for example, Boudoukh, Richardson and Whitelaw (1997), Christoffersen and Diebold (2000)]. While one should not argue with demonstrated

---

[1] We define "volatility" as a rate of return's instantaneous variance.

[2] See, for example, Zahavi and Zahavi (1997).

empirical success of a model,[3] careful search for equilibrium-based models that are both internally consistent and supported by valid preferences, might have a good chance of success.[4]

Within the equilibrium framework adopted in Bick (1990), He and Leland (1993), Xu (2013),[5] we study two families of equilibrium-based volatility models for the market portfolio rates of return: one is a deterministic function of time, and the other is an inverse function of the ex-dividend value of the market portfolio (henceforth, "market portfolio value"). Both families of volatility models are standard in finance. They entail log-normal processes and square-root processes for the market portfolio value, respectively. The first family of volatility models is represented by three popular models in academia and industry (henceforth, "reference models"): 1) GARCH,[6] 2) RiskMetrics,[7] and 3) piecewise constant volatility. The second family is a special case of the constant elasticity of variance models (CEV); see, for example, Cox and Ross (1976) and the sequel. Within the second family, we propose three new specifications motivated by the three reference models, respectively. Each specification combines parameterization of one reference model and the CEV formulation (i.e., as an inverse function of the market portfolio value).

We show in theory that, under the common practice of assuming constant mean rates of return, the second family of volatility models has the features of allowing more flexible risk aversions and capturing "asymmetric volatility" [Black (1976)]. Thus, the second family tends to have more flexibility to approximate the market-implied risk aversion, and thus a

[3] Friedman (1953).

[4] Naturally, volatility models are also used to explain asset returns and growth. See for example Schorfheide, Song and Yaron (2013).

[5] Following Bick (1990), He and Leland (1993), Xu (2013, Chapter 3) derived necessary and sufficient conditions for the market portfolio return process to be consistent with equilibrium under market clearing stochastic interest rates, supporting a larger family of equilibrium consistent specifications of mean/volatility structures.

[6] Generalized Autoregressive Conditional Heteroskedasticity [Engle (1982) and Bollerslev (1986)]. We use GARCH(1,1).

[7] We use the 1996 version [J.P. Morgan/Reuters (1996)], which can be viewed as a special case of GARCH(1,1), not the 2006 version [Zumbach (2007)].

better ability to forecast, than the first family. To empirically test this hypothesis, we compare out-of-sample accuracy of one-day-ahead daily volatility forecasts produced by our proposed and reference models, in three evaluations of different time periods, using the Standard & Poor 500 (S&P500) Composite Index data. While our proposed models are as easy to implement, they outperform the reference models by standard predictability criteria, including R^2 of the Mincer-Zarnowitz (1969) type (henceforth MZ) regressions used by Andersen and Bollerslev (1998) (henceforth AB), and squared forecast errors, tested using Diebold and Mariano (1995) (henceforth DM) tests, during both "normal" and high-volatility periods, whether the market is rising or falling.

To address potential non-stationarity of the market portfolio rates of return implied by the proposed models, we model a difference equation of the CEV volatility,[8] and use Maximum Likelihood method to estimate parameters. In our case, the Maximum Likelihood Estimator is still asymptotically unbiased and asymptotically normally distributed (see Bollerslev, Engle and Wooldridge (1988) and references therein). We adopt two standard model comparison methodologies. The first is MZ regressions used in AB and Andersen, Bollerslev, Diebold and Labys (2003), and the second is DM tests on squared errors.[9] By the first model comparison criterion, the R^2 of regressions of realized volatility on forecasts (i.e., MZ regressions), our models outperform the reference models, typically by 5%-10%, in all three out-of-sample periods. By the second criterion, mean square error (MSE), our proposed models' MSEs are typically 5%-10% lower, which is statistically significant. The DM tests also show that the outperformance mainly comes from high-volatility periods. Moreover, as a robustness check, we plot model forecasts vis-a-vis realized volatility for all holdout periods and visually assess model underestimation levels during high-volatility periods. While similar

[8] We thank Frederik Lundtofte for suggesting that we emphasize this point.

[9] The forecast evaluation methodology of volatility models is an active research area. In this paper we employ the two classic methodologies, and we might incorporate more recent ones in future studies; for example, Patton (2011).

in "normal" periods, during high-volatility periods, the reference models typically underestimate realized volatility by 30%-60%, and ours are typically better by 10%-20%.

We advance the market risk modeling literature by introducing an approach of developing equilibrium-based volatility models of the market portfolio rates of return. We find empirical support for our approach by improving the predictive power of extant models without increasing implementation costs. Our approach paves foundations for creating new models along this line.

In addition, we provide a new way to analyze and extend current popular volatility models, within an equilibrium framework, thus facilitating insights into the predictability of volatility models. Although the families of volatility models studied here are not uncommon, to our best knowledge, we are the first to analyze them within an equilibrium framework and compare their empirical performance accordingly.

Our paper also contributes to the literature that tries to combine the CEV-type and GARCH-type volatility models; see, for example, Fornari and Mele (2006). Our approach of combining GARCH and CEV is different from theirs. It accommodates compatibility with equilibrium and its efficacy is demonstrated by empirical evidence.

Section 2 introduces our equilibrium framework and proposed models. Section 3 describes the estimation procedure, the out-of-sample forecasting procedure, the evaluation methodology, and the data. Section 4 presents empirical results. Section 5 concludes.

2 Equilibrium-Based Volatility Models

In this section, we identify two families of volatility models that are consistent with equilibrium. Recognizing that the three reference models, which are popular in academia and industry, belong to the first family, we propose three specifications in the second family, each corresponding to one of the reference models.

2.1 Equilibrium-Consistent Diffusion Return Processes of the Market Portfolio

We first describe the framework adopted in He and Leland (1993), Xu (2013). We use their necessary and sufficient conditions for the return process of the market portfolio to be consistent with equilibrium.

In a continuous-time, frictionless, fully divisible securities market, over a finite time horizon, there is a single source of uncertainty, Z, $Z \triangleq \{Z_t, \mathcal{F}_t\}$, $Z_0 = 0$, a Wiener process defined over a complete probability space $(\Omega, \mathcal{F}, P)$ with a non-decreasing right continuous family of sub-σ-algebras $\{\mathcal{F}_t, 0 \le t \le T\}$. There is a riskless asset, at zero net supply, and a risky asset, at unit net supply (the market portfolio), whose ex-dividend value at time t, $t \in [0,T]$, M_t, is $\{\mathcal{F}_t\}$–adapted and evolves as an Itô diffusion. Thus, it is a (strong) solution to the following stochastic differential equation:

$$\frac{dM_t}{M_t} = \mu^E(M_t,t)dt + \sigma(M_t,t)dZ_t, \tag{1}$$

with an initial condition M_0, where, $\mu^E(M_t,t)$ and $\sigma^2(M_t,t)$ the instantaneous mean and volatility of the ex-dividend rate of return, are twice continuously differentiable scalar functions with respect to M_t and t. We also assume that the market portfolio pays a continuous dividend yield $\delta(M_t,t)$,[10] which is continuously differentiable with respect to M_t and t, and satisfies $\delta(M_t,t) \in [0,1)$, $\forall t < T$ and $\delta(M_T,T) = 1$. We require that $\mu^E(M_t,t)$, $\sigma^2(M_t,t)$, and $\delta(M_t,t)$ satisfy Lipschitz and growth conditions [see Liptser and Shiryaev (2001, pp. 134-135)] to guarantee a solution to Equation (1) and that this solution be non-negative; that is, $M_t \ge 0$, $\forall t$. Note that this market is dynamically complete.

A representative investor who lives on $[0,T]$ is endowed with an initial wealth of W_0, $W_0 = M_0$, and maximizes a time-additive, state-independent Von Neumann-Morgenstern

---

[10] Thus, $(\mu_t^E + \delta_t)$ is the instantaneous expected total rate of return of the market portfolio.

utility of lifetime consumption, $U(C_t, t)$, by allocating his/her wealth, W_t (the total wealth in the economy) to consumption and investments in the risky and riskless assets. This is a version of Merton's (1971, 1973) problem.

We set the instantaneous riskless interest rate, r (henceforth, "interest rate"), to clear the market where the investor's optimal weights in the risky and riskless asset are (always) 1 and 0 respectively, and set $\delta(M_t, t)$ to equal the investor's optimal consumption rate. In other words, the investor liquidates his/her wealth, in the form of dividends, at his/her optimal consumption rate. In equilibrium, r is a function of (W_t, M_t, t) [see, for example, Cox, Ingersoll, and Ross (1985a,b)]. Because W_t is a deterministic function of M_t, (particularly, $W_t = M_t$), we can write r as a function of M_t only; that is, $r = r(W_t, M_t, t) = r(M_t, M_t, t)$. We define f, $f \triangleq \frac{(\mu^E + \delta - r)}{\sigma^2}$, and a similar argument allows us to also write it as a function of M_t only, $f = f(M_t, t)$.

The necessary and sufficient conditions for the process $\{M_t,\ M_t \geq 0,\ \forall t \in [0,T]\}$ to be consistent with equilibrium are [see He and Leland (1993), Xu (2013, Chapter 3)]

1. $$\sigma\sigma_M [f_M M + f^2 - f] + \frac{1}{2}\sigma^2 f_{MM} M + \mu^E f_M - \delta_M f + \frac{f_t}{M} + r_M (f - 1) = 0, \tag{2}$$

and

2. there exists an increasing and concave utility function $U(C_t, t)$ such that f satisfies

$$\begin{cases} f(M,t) = -\dfrac{U_{CC}(M,t)}{U_C(M,t)} M(\delta_M M + \delta) \\ f(M,T) = -\dfrac{U_{CC}(M,T)}{U_C(M,T)} M \end{cases}. \tag{3}$$

where in Equations (2) and (3) subscripts denote partial derivatives and, for simplicity, we suppress the time subscript of M_t.

2.2 Two Families of Equilibrium-Based Volatility Models

The above equilibrium framework and conditions per se do not entail any particular functional form of $\mu^E(M_t,t)$, $\sigma^2(M_t,t)$, $\delta(M_t,t)$, and $r(M_t,M_t,t)$. They only impose (minimum) restrictions that any specification needs to abide by. Hence, this framework provides much freedom for econometricians to create volatility specifications. In order to conduct empirical study, econometricians need to specify, at least, $\mu^E(M_t,t)$ and $\sigma^2(M_t,t)$. In this paper, we consider constant instantaneous mean rates of return (i.e., $\mu^E(M_t,t)=\mu$, where μ is a constant over time), and the following two families of volatility models:

1. $\sigma^2(M_t,t)=S(t)$, i.e., volatility as a deterministic function of time, and
2. $\sigma^2(M_t,t)=\dfrac{S(t)}{M_t}$, i.e., volatility as an inverse function of the market portfolio value,

where $S(t)$ is a positive, bounded and differentiable function of time.

The "constant mean" assumption is consistent with standard practices in the market risk measurement literature. These practices demean the return time series and estimate the volatility as if the conditional mean rates of return are zero [see, for example, Christoffersen, Hahn and Inoue (2001)]. In order to have fair model comparisons, we make this assumption for all volatility models studied here.

Both families of volatility models are standard in finance. As we will show in the next section (Section 2.3), the first family, as a deterministic function of time, is represented by three popular models in academia and industry, GARCH, RiskMetrics, and piecewise constant volatility. The second family is a special case of the constant elasticity of variance (CEV) models; see, for example, Cox and Ross (1976) and the sequel, which are also not uncommon.

Both families can be consistent with our equilibrium framework. For example, the first family can induce preferences with Constant Relative Risk Aversion (CRRA). In particular, if the dividend yield is also a deterministic function of time, $\delta(M_t, t) = \delta(t)$, in view of Equation (3),

$$\frac{\mu^E + \delta(t) - r}{\sigma^2(t)} = f = -\frac{U_{CC}(M,t)}{U_C(M,t)} M\delta = -\frac{U_{CC}}{U_C} C = \eta \,, \tag{4}$$

where η is the coefficient of RRA. Thus, as long as $\left(\mu^E - r\right)$ is a deterministic function of time, then $\eta = \eta(t)$ is not a function of C. In other words, this family of volatility models can be supported by CRRA preferences.[11] The second family can induce preferences with Constant Absolute Risk Aversion (CARA), as shown in the next proposition.

Proposition 1. The following specifications of the market portfolio rates of return are consistent with equilibrium in the economy, described in Section 2.1, where the investor has CARA preferences.

The volatility of the market's portfolio rates of return is

$$\sigma^2(M_t, t) = \frac{S(t)}{M_t}. \tag{5}$$

The (ex-dividend) mean of the market portfolio rates of return, interest rates, and dividend yields are, respectively,

$$\begin{cases} \mu^E(M_t, t) = \mu^E(t) = \frac{1}{2}\gamma(t)S(t)\delta(t) - \frac{d\log\gamma}{dt} - \frac{d\log\delta}{dt} \\ \delta(M_t, t) = \delta(t) \\ r(M_t, M_t, t) = r(t) = \mu^E(t) + \delta(t)[1 - S(t)\gamma] \end{cases}, \tag{6}$$

where $\delta(t)$ is a function of t satisfying $\delta(t) \in [0,1)$, $\forall t < T$ and $\delta(T) = 1$ (which becomes the equilibrium consumption rate), and $\gamma(t)$ is a positive function of time (which becomes the absolute risk aversion (ARA) coefficient of the utility function).

---

[11] In this case, in light of Equation (2), $r_M = \frac{\eta_t}{M(1-\eta)}$. Hence, r and μ^E can both be some logarithmic functions of M_t.

Proof. It suffices to show that the specifications in Equations (5) and (6) satisfy the conditions in Equations (2) and (3). One can verify that Equation (2) holds by direct substitution. Substituting the definition of f in Equation (3), recalling that $\delta(T)=1$, implies that the ARA coefficient of the utility function is $\gamma(t),\ \forall t\in[0,T]$, which is independent of wealth. See Equation (7) below.

$$f(M_t,t)=\frac{(\mu^E+\delta-r)}{\sigma^2}=\delta(t)\gamma(t)M_t \ \Rightarrow\ -\frac{U_{CC}(M_t,t)}{U_C(M_t,t)}=\gamma(t),\quad \forall t\in[0,T], \tag{7}$$

where subscripts of U denote partial derivatives. □

While both families can be consistent with equilibrium and can induce reasonable preferences, when used together with the "constant mean" assumption, the second family tends to gain more advantages, in terms of the flexibility to approximate the market-implied risk aversion and thus the ability to forecast.

The first family together with the "constant mean" assumption might induce more restricted preferences. For example, in Equation (4), if $\mu^E(M,t)=\mu$, then $r=r(t)$, and the equilibrium condition in Equation (2) becomes $\frac{\eta_t}{M(1-\eta)}=r_M=0\Rightarrow\frac{d\eta}{dt}=0$. Hence, the RRA coefficient, η, must be constant over time. In contrast, as shown in Proposition 1, even if dividend yields (which are also the optimal consumption-to-wealth rates) and interest rates are deterministic, the second family together with constant mean still allows the ARA coefficient, $\gamma(t)$, to be (deterministically) time-varying.[12]

Moreover, the second family features "asymmetric volatility" [Black (1976)], a well-known stylized fact that can be observed under the "constant mean" assumption.[13] The

---

[12] When interest rates are deterministic, as in these cases, the specifications also satisfy the conditions of He and Leland (1993).

[13] This stylized fact is that the volatility moves asymmetrically with the security value (or price), i.e., the volatility is higher when the value falls than when it rises. This phenomenon was explained by Black (1976) as the "leverage effect." Nonetheless, to our best knowledge, it is unknown in the literature whether it can be observed when stochastic mean rates of return are modeled.

asymmetry of volatility is captured by the inverse function of M_t , even though this functional form is motivated by the compatibility with CARA preferences, not by “asymmetric volatility” alone.

Finally, the major problem with CARA preferences is lack of wealth effect in ARA. This is less a problem over short time intervals, like days or weeks, because the wealth of the representative investor does not change much, and hence does not impact ARA much. Thus, CARA preferences, especially with time-varying ARA, might approximate real-world markets reasonably well.

To shed further light on both families of volatility, substituting $\mu^E(M,t) = \mu$ and $\sigma^2(M_t, t)$ into the dynamics of the market portfolio value, M_t, we note that under constant mean rates of return, M_t follows log-normal processes using volatility in the first family, and follows “square-root” processes using volatility in the second, as shown below,

$$dM_t = \mu M_t dt + \sqrt{S(t)}\sqrt{M_t} dZ_t \,. \tag{8}$$

Equation (8) ensures the non-negativity of M_t, as desired. Therefore, here we effectively compare the empirical performances of log-normal and square-root processes of M_t, with the former featuring symmetric volatility and the latter featuring asymmetric volatility, under the “constant mean” assumption.

The above theoretical analysis yields empirically testable hypotheses regarding the forecast performance of the two families of volatility models, which are the major subjects of Sections 3 and 4. To test empirical forecast performance, we need to provide plausible parametric specifications in the second family.

2.3 Parametric Specifications

We first recognize that the three reference models are, in fact, parameterizations in the first family of volatility models. Denoting $y_t \triangleq \frac{M_t - M_{t-1}}{M_{t-1}}$ as the realized rate of return at time t, the reference models are

Reference Model 1, GARCH(1,1): $\sigma_{t+1}^2 = \alpha_0 + \alpha_1 \left(y_t - \mu \right)^2 + \alpha_2 \sigma_t^2$, (9)

where μ, α_0, α_1 and α_2 are positive constants estimated from the data.

Reference Model 2, RiskMetrics: $\sigma_{t+1}^2 = 0.06\left(y_t - \hat{\mu} \right)^2 + 0.94\sigma_t^2$, $\hat{\mu} = \frac{1}{D} \sum_{s=t-D+1}^{t} y_s$, (10)

where D is the estimation sample size (in number of periods).

Reference Model 3, "Moving Average": $\hat{\sigma}_{t+1}^2 = \frac{1}{N} \sum_{s=t-N+1}^{t} (y_s - \hat{\mu})^2$, $\hat{\mu} = \frac{1}{N} \sum_{s=t-N+1}^{t} y_s$, (11)

where $N \le D$ is the sub-sample size for estimation (in number of periods). RiskMetrics can be viewed as a special case of GARCH(1,1), with $\alpha_1 + \alpha_2 = 1$ and $\alpha_0 = 0$. Because of this property, RiskMetrics is also known as an "Integrated GARCH(1,1)" model or the "Exponentially Weighted Moving Average" (EWMA) model. We call the third reference model "Moving Average" because it is effectively the N-day-moving average of the squared (demeaned) returns.

These models are standard benchmarks used for model comparisons; see Boudoukh, Richardson and Whitelaw (1997), Christoffersen and Diebold (2000), Christoffersen, Hahn and Inoue (2001). They are also widely used in practice; see, for example, Phelan (1995), Litterman and Winkelmann (1998), and BCBS (2011). As shown below, they can all be viewed as discretized versions of volatility models in the first family, $\sigma^2(M_t, t) = S(t)$, with certain functional forms of $S(t)$.

For the first reference model, Corradi (2000) showed that the continuous time limit of the GARCH(1,1) process is the following ordinary differential equation (ODE):[14]

[14] With appropriate initial conditions.

$$dS(t) = \left[\alpha_0 + (\alpha_1 + \alpha_2 - 1)S(t)\right]dt\,, \qquad (12)$$

whose solution is obviously a deterministic function of time.[15] The second reference model, RiskMetrics, can be simply viewed as a special case of GARCH(1,1) with certain pre-defined parameter values, $\alpha_0 = 0$, $\alpha_1 = 0.06$, $\alpha_2 = 0.94$, and $\mu = \frac{1}{D}\sum_{s=t-D+1}^{t} y_s$. For the third reference model, expressed in Equation (11), we recognize that it is the Maximum Likelihood Estimator (MLE) of piecewise constant volatility, using a sub-sample of N periods, i.e., $\sigma^2(M_t, t) = S(t) = S$,where S is a positive constant.

The above analysis of the reference models has direct implications for parameterizations in the second family of volatility models, i.e., a CEV model. Instead of thinking of $S(t)$ as volatility, we take it as a component of volatility in the second family, $\sigma^2(M_t, t) = \frac{S(t)}{M_t}$. Thus, the functional forms of $S(t)$ implied by the three reference models directly lead to three new parametric specifications in the second family. They are our proposed models, as presented below.

Model 1: $$\sigma_t^2 = \frac{\alpha_0}{M_{t-1}} + \alpha_1\left(y_{t-1} - \mu\right)^2 + \alpha_2\sigma_{t-1}^2 + \sigma_{t-1}^2\left[y_{t-1}{}^2 - y_{t-1}\right], \qquad (13)$$

Model 2: $$\sigma_t^2 = 0.06\left(y_{t-1} - \frac{1}{D}\sum_{s=1}^{D} y_s\right)^2 + 0.94\sigma_{t-1}^2 + \sigma_{t-1}^2\left[y_{t-1}{}^2 - y_{t-1}\right], \qquad (14)$$

Model 3: $$\sigma_t^2 = \frac{S}{M_{t-1}}\,. \qquad (15)$$

Model 1 is created by combining a CEV formulation and GARCH(1,1):

[15] The limit of the GARCH(1,1) process may not be unique. For example, the well-known result of Nelson (1990) showed that in the limit it becomes a stochastic volatility process. We use Corradi's (2000) result because the Euler discretization of Equation (12) leads back to GARCH(1,1). For more detailed discussion, see Singleton (2006, Chapter 7, pp. 177-178).

$$\begin{cases} \sigma^2(M_t,t) = \dfrac{S(t)}{M_t} \\ dS(t) = \left[\alpha_0 + (\alpha_1+\alpha_2-1)S(t)\right]dt \end{cases}. \qquad (16)$$

We can simplify Equation (16), by differentiating $\sigma^2(M_t,t)$ using Itô's formula, substituting Equation (12), and rearranging:

$$\begin{aligned} d\sigma_t^2 &= \frac{dS(t)}{M_t} - \frac{S(t)}{M_t}\frac{dM_t}{M_t} + \frac{S(t)}{M_t^3}(dM_t)^2 \\ &= \frac{\alpha_0}{M_t}dt + (\alpha_1+\alpha_2-1)\sigma_t^2 dt + \sigma_t^2\left[\left(\frac{dM_t}{M_t}\right)^2 - \frac{dM_t}{M_t}\right]. \end{aligned} \qquad (17)$$

Noting that $\sigma_t^2 dt = \left(\sigma_t dZ_t\right)^2 = \left(\dfrac{dM_t}{M_t} - \mu dt\right)^2$, we can further rewrite Equation (17) as

$$d\sigma_t^2 = \frac{\alpha_0}{M_t}dt + \alpha_1\left(\frac{dM_t}{M_t} - \mu dt\right)^2 + (\alpha_2-1)\sigma_t^2 dt + \sigma_t^2\left[\left(\frac{dM_t}{M_t}\right)^2 - \frac{dM_t}{M_t}\right]. \qquad (18)$$

Applying Euler discretization to Equation (18) using the lagged information set,[16] we obtain Model 1 of Equation (13), in discrete time, where all parameters are to be estimated from the data. Note that Model 1 is obtained from a first-order difference equation of the CEV volatility, Equation (18), rather than the CEV volatility itself, Equation (5). This treatment addresses potential concerns on the existence of the long-run (or unconditional) mean of the CEV volatility processes.[17]

If we set μ, α_0, α_1 and α_2 as pre-defined constants, as in RiskMetrics, $\alpha_0 = 0$, $\alpha_1 = 0.06$, $\alpha_2 = 0.94$ and $\mu = \dfrac{1}{D}\sum_{s=t-D+1}^{t} y_s$, we obtain Model 2 of Equation (14).[18]

---

[16] Note that in discretized versions, M_t becomes M_{t-1}.

[17] Equation (13) might be used to identify the conditions under which the long-run mean of Model 1 exists, see Appendix A. We also note that the volatility in all the models considered in this paper is conditionally deterministic. In other words, conditioning on the lagged information set, the (following) volatilty is known and the error terms are i.i.d. Therefore, the non-stationarity of the volatility process is of less a concern when Maximum Likelihood method is used, as described in the next section.

[18] Note that we could have fine-tuned the parameter values of Model 2 because of its different functional form from RiskMetrics. Here we adopt the values from RiskMetrics for simplicity.

For Model 3, in the spirit of the "moving average" model, we choose $S(t)$ to be a constant, $S(t) = S$, over a sub-sample of N periods, giving rise to the parameterization in Equation (15).

Parameterizations are of course not unique. Here we propose three specifications in a way that is directly analogous to the reference models, and requires no additional implementation cost due to the functional similarity between the proposed and reference models. Indeed, the number of parameters in our models is the same as that in the reference models, and the econometric methods required to estimate parameters are also the same.[19] However, the estimated parameter values, the volatility forecasts, and their empirical implications are greatly different. We provide further insights into our proposed models using Model 1 as an example.

Model 1 looks similar to GARCH(1,1). However, we stress that it is *not* created by nesting GARCH, but by "borrowing" GARCH as one component in its specification. This is a new, nonstandard approach to invent models. One distinction between Model 1 and GARCH(1,1) is the extra addend of $\sigma_{t-1}^2\left[y_{t-1}{}^2 - y_{t-1}\right]$, which is a consequence of the second-order Taylor expansion of the inverse function of M_t in Itô's formula. This addend has the feature of capturing "asymmetric volatility." To see why, we rewrite it as $\left[\sigma_{t-1}^2(y_{t-1} - 0.5)^2 - 0.25\sigma_{t-1}^2\right]$. Therefore, as long as the lagged realized rates of return, y_{t-1}, is below 50%, which is true at daily or weekly frequencies, *ceteris paribus*, the higher the return, the lower the next-period volatility, and vice versa. Furthermore, in this model, we require $\alpha_2 \geq 0.25$ to ensure the positivity of the volatility forecasts. We find in empirical study that the unconstrained estimate of α_2 always satisfies this condition. Finally, there may

[19] Fornari and Mele (2006), for instance, used Nelson's (1990) interpretation of GARCH(1,1) and, thus, required more sophisticated econometric methods like Indirect Inference.

also be a concern about the term of $\frac{\alpha_0}{M_{t-1}}$, which can cause volatility to explode if $M_t \to 0$. When this happens, the squared-root processes in Equation (8) imply that the economy stops. Otherwise, the assumed Lipschitz and growth conditions ensure that $\sigma^2(M_t, t) < \infty$ when the economy is running. In empirical study, the absolute level of M_t is meaningless. Hence, we always normalize it to 1 at beginning of each estimation period. This treatment also avoids the potential explosion issue.

The analogs in the parameterizations, in terms of $S(t)$, provide a natural correspondence between our proposed models and the reference models: Model 1 vs. GARCH(1,1), Model 2 vs. RiskMetrics, and Model 3 vs. the "Moving Average" model. Therefore, in empirical study, we compare our models' forecast performance to that of their corresponding reference models. Such model comparisons are effectively between the two families of volatility models because the specifications of $S(t)$ are the same across, correspondingly, our model and the reference models.

It might also be interesting to compare the empirical performances between our models and models that are purely designed to capture "asymmetric volatility" and are not necessarily consistent with equilibrium. This comparison requires separate, full-blown empirical work, which we defer to future study.[20]

Our approach, to create new models in the second family, can be generalized. If a volatility process is recognized as a deterministic function of time, $S(t)$, we can easily create a corresponding new model using Equation (5), as we have done here.

[20] However, we stress that our focus is on equilibrium-consistency of modeling practices rather than capturing "asymmetric volatility" (leverage effect) alone, because the latter may have less significant implications. For example, it is still an open question whether leverage effect exists if stochastic conditional mean rates of return are modeled.

3 Estimation Procedure, Forecast Evaluation Methodology, and Data

In this section we describe the volatility models' parameter estimation methodology, forecast evaluation methodology, and the data we use.

3.1 Parameter Estimation

We use the Maximum Likelihood method to estimate parameters in our proposed models, where required, because, in our case, the MLE can be shown asymptotically unbiased and asymptotically normally distributed (with the asymptotic covariance matrix), under standard regularity conditions [see Bollerslev, Engle and Wooldridge (1988) and the references therein].

For Model 1's reference model, GARCH(1,1), we use the well-known Maximum Likelihood estimation procedure [Engle (1982), Bollerslev (1986)]. For Model 1, we use a similar procedure to identify the MLEs, described in Appendix A.

For Model 2 and its reference model, RiskMetrics, no estimation is required.

For Model 3's reference model, the "Moving Average" model is in fact the MLE of piecewise constant volatility using a sub-sample of N periods, as discussed earlier. We derive a closed-form expression, shown in Appendix A, for Model 3's MLE of the parameters, $\hat{S}^{MLE}$ and $\hat{\mu}^{MLE}$. Using them, Model 3 of Equation (15) becomes

$$\textbf{Model 3}: \quad \hat{\sigma}_t^2 = \frac{\sum_{s=t-N}^{t-1} (y_s - \hat{\mu})^2 M_{s-1}}{N M_{t-1}}, \qquad (19)$$

where $\hat{\mu} = \frac{\sum_{s=t-N}^{t-1} y_s M_{s-1}}{\sum_{s=t-N}^{t-1} M_{s-1}}$. We set $N = 10$ for both Model 3 and the "Moving Average" model.[21]

[21] The choice of N is arbitrary. We conjecture that it is less relevant to the model comparison we are interested in here because both models adopt the same value.

As an example, we present MLEs using the daily return data of S&P500 Composite Index during the period from July 1, 1997 to June 29, 2007. The data will be described in Section 3.3 below. Note that Model 3 and its reference model use only the data of the last 10 trading days in the estimation sample (i.e., the data from June 18, 2007, to June 29, 2007).

The MLEs for the reference models are as follows, with standard errors in parentheses:

GARCH(1,1): $$\sigma_{t+1}^2 = \underset{(7.09\times10^{-7})}{1.74\times10^{-6}} + \underset{(0.0169)}{0.0915}\left(y_t - \underset{(0.0002)}{0.0005}\right)^2 + \underset{(0.0163)}{0.8972}\sigma_t^2, \tag{20}$$

"Moving Average": $$\hat{\sigma}_{t+1}^2 = \underset{(2.0771\times10^{-5})}{4.6445\times10^{-5}}, \quad \hat{\mu} = \underset{(0.0022)}{-0.0019}. \tag{21}$$

The MLEs for the proposed volatility models are[22]

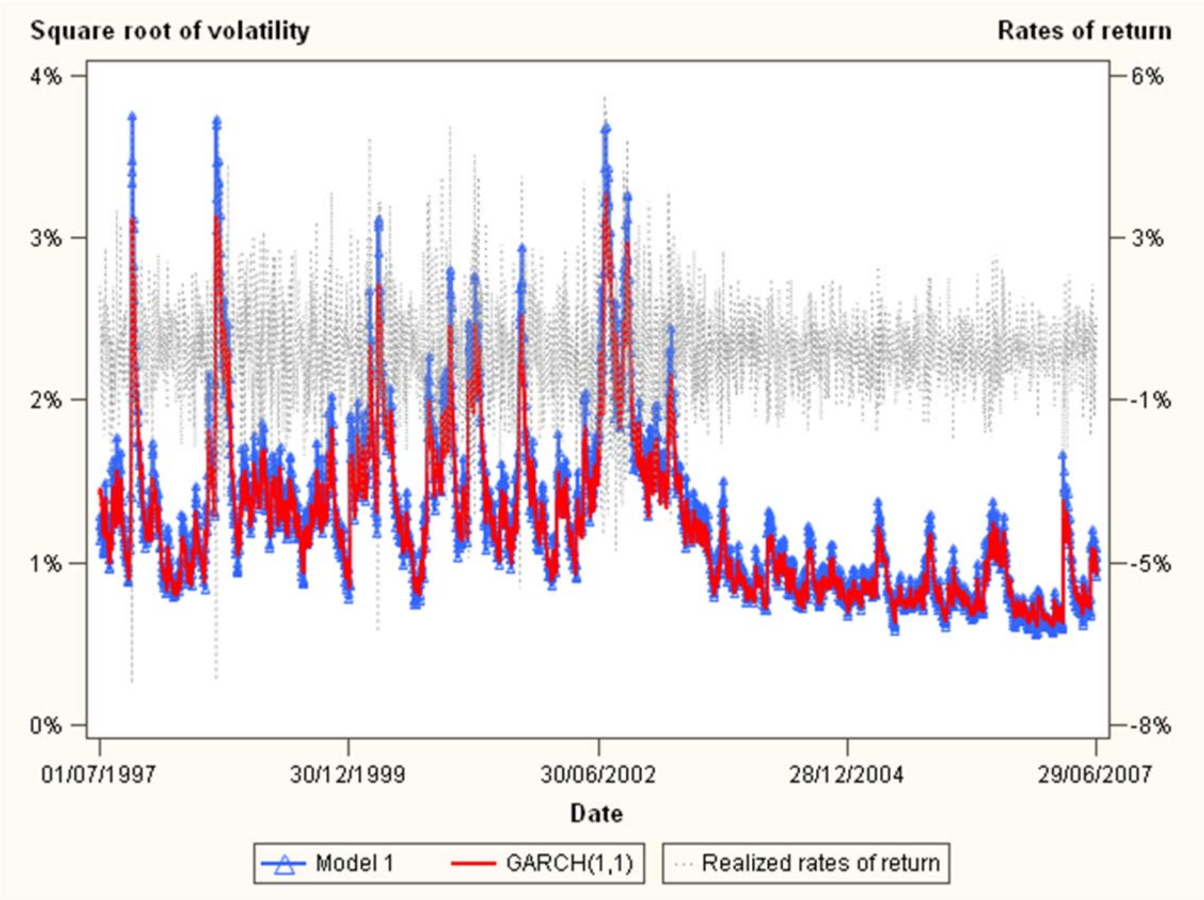
Model 1: $$\sigma_{t+1}^2 = \frac{\underset{(9.92\times10^{-7})}{3.18\times10^{-6}}}{M_{t-1}} + \underset{(0.0208)}{0.1455}\left(y_{t-1} - \underset{(0.0002)}{0.0005}\right)^2 + \underset{(0.0208)}{0.8540}(\sigma_{t-1}^2) + \sigma_{t-1}^2\left[y_{t-1}{}^2 - y_{t-1}\right], \tag{22}$$

Model 3: $$\hat{\sigma}_{t+1}^2 = \underset{(2.0956\times10^{-5})}{4.6859\times10^{-5}}, \quad \hat{\mu} = \underset{(0.0022)}{-0.002}. \tag{23}$$

All the parameters (except for $\hat{\mu}$ in Model 3 and the "Moving Average" model) are significantly different from zero at the 5% significance level using the standard t- or z-tests. We also note that, like GARCH(1,1), the sum of the three parameter values in Model 1 is also close to 1; nevertheless, they are quite different from those of GARCH(1,1). The difference can also be seen from the in-sample estimates of volatility during this period in Figure 1, where the estimates of both models are plotted with vertical axis labeling on the left (red and blue with triangle markers and lines, respectively) and actual daily rates of return are plotted labeling on the right (grey dotted line). Therefore, one expects that the two models produce different daily volatility forecasts, as we will demonstrate later.

[22] In Model 1, we show $\hat{\sigma}_{t+1}^2$ instead of $\hat{S}$ to draw an analog to the "Moving Average" model.

Figure 1: In-sample estimates of daily volatility, Model 1 vs. GARCH(1,1)



3.2 Forecast Evaluation[23]

We compare all models' performance of one-day-ahead daily volatility forecasts. We adopt the setup of the forecast evaluations as in AB and Andersen, Bollerslev, Diebold and Labys (2003) and describe it below.

We first create holdout samples using the daily realized rate of returns of the market portfolio. Next, we use each model to produce a one-day-ahead forecast for each day in the holdout samples.[24] Last, we evaluate the forecasts using the realized volatility of the holdout samples.

The realized volatility is the sum of squared *intraday* returns, obtained by aggregating

[23] "Forecast evaluation" is also frequently referred to in the literature as "back-testing." We use this terminology following AB and Andersen, Bollerslev, Diebold and Labys (2003).

[24] We give each model a sample of D daily (actual) returns immediately prior to each holdout observation (i.e., a rolling time window of size D). Each model then estimates the parameters using the D -trading day sample and produces a volatility forecast for the next day. We set $D = 2500$.

five-minute rates of return within each day and overnight. Where the intraday data is not available, we use the squared daily rates of return.

To evaluate volatility forecasts, we adopt two standard evaluation criteria: R^2 of the evaluation regressions (henceforth "R^2") and mean square error (MSE).

R^2 is obtained by regressing the square roots of realized volatilities onto a constant and the square roots of forecasts (i.e., MZ regressions):

$$(v_{t+1})^{1/2} = b_0 + b_1 (v_{t+1|t,\ Model})^{1/2} + u_{t+1}, \tag{24}$$

where v_{t+1} is the realized volatility as described above, $v_{t+1|t,\ Model}$ is the one-day-ahead volatility forecast of a model, and u_{t+1} is the regression error term. This is our first criterion to evaluate the forecast performance of the volatility models (see AB's justification for that).[25]

Our second criterion is the forecast MSE,[26] $MSE \triangleq \frac{1}{H}\sum_{t=1}^{H}\left[(v_{t+1|t,\ Model})^{1/2} - (v_{t+1})^{1/2}\right]^2$, where H is the holdout sample size. This criterion assesses overall forecast errors and thus provides a measure of how "far away" the forecasts are from realized volatility. It corresponds to a special case of R^2, where the regression coefficients, b_0 and b_1, are set to zero and one, respectively. While R^2 tends to ignore the average bias, the MSE reflects both bias and variance; according to the standard decomposition, $MSE = Bias^2 + Variance$.[27] The decomposition also allows us to trace the sources of MSE differences. We test MSE differences using a Diebold-Mariano [Diebold and Mariano (2002), DM] test.[28] We also conduct DM tests, using the "squared error" loss function, to assess the statistical significance

[25] The population b_0 and b_1 should be zero and one, respectively. However, "Errors-In-Variables" issues bias the regression coefficients estimators in MZ regressions, Equation (24). Following AB we use only the R^2 criterion.

[26] The MSE of the out-of-sample forecasts (i.e., the forecasting errors), while commonly used in statistical predictive modeling, is not commonly used in the volatility literature.

[27] We thank Wayne Ferson for suggesting the motivation of using the MSE.

[28] We thank Jonathan Reeves for suggesting the DM tests.

of performance difference during high-volatility and "normal" periods, respectively.

Moreover, as a robustness check, we plot model forecasts vis-a-vis realized volatility for all holdout periods and visually assess model underestimation levels during periods with the highest realized volatility. We pay particular attention to this property because of current volatility models' tendency to underestimate volatility during those crucial periods [see, for example, Christoffersen and Diebold (2000)].

3.3 Data

We conduct forecast evaluations using ex-dividend prices of S&P500, which is commonly used in the risk-measurement literature [see, for example, Christoffersen, Hahn and Inoue (2001)]. Also, as a broad value-weighted index, it is an appropriate choice for our equilibrium framework.

To minimize data-snooping bias, we use three holdout samples, in non-overlapping periods. Table 1 summarizes the size and time coverage of the three pre-holdout and holdout samples. As discussed before, we normalize the S&P500 "price" (or index level) to 1 at the beginning of each pre-holdout period.

We summarize the statistics of the unconditional distributions of realized rates of return in Table 2. The return distributions in the holdout samples have high excess kurtosis (>0), a manifestation of the heavy tails reflecting big market movements (see Figure 2). The three holdout samples capture most of the major market swings in the last three decades. We number these swings in Figure 2's Panel A, B and C and describe them in Panel D. Our holdout samples are, thus, a good test field for evaluating volatility models both in general and under big market movements.

To calculate realized volatility, we need high-frequency data on the S&P500 Index, which we find for only the second and third evaluation periods. It is the intraday prices of the

SPDR S&P500 Exchange-Traded Fund (SPY).[29] For the first period, we use the squared daily returns as realized volatility.

We use DataStream's S&P500 daily price data and the intraday price data of SPY from NYSE-TAQ (New York Stock Exchange Trade and Quote) in Wharton Research Data Services (WRDS).

Table 1: Holdout and pre-holdout samples: time coverage and sample size

Holdout sample	Pre-holdout sample coverage/size	Holdout sample coverage/size
1	July 1, 1977 – June 30, 1987: 2,515	July 1, 1987 – June 30, 1990: 757
2	July 1, 1987 – June 30, 1997: 2,525	July 1, 1997 – December 31, 1998: 380
3	July 1, 1997 – June 30, 2007: 2,515	July 1, 2007 – December 31, 2008: 380

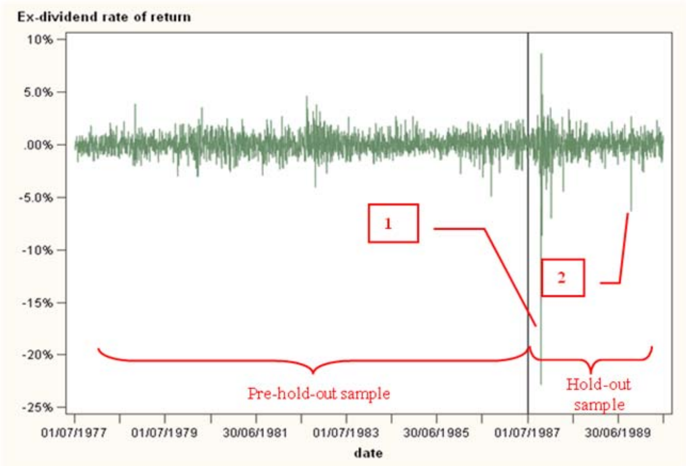
Table 2: Holdout samples' unconditional daily return distributions summary statistics

Holdout sample	Mean	Std. dev.	Skewness	Excess kurtosis
1	0.0216%	1.433%	-5.656	89.764
2	0.0864%	1.286%	-0.679	5.729
3	-0.1344%	2.228%	-0.102	5.224

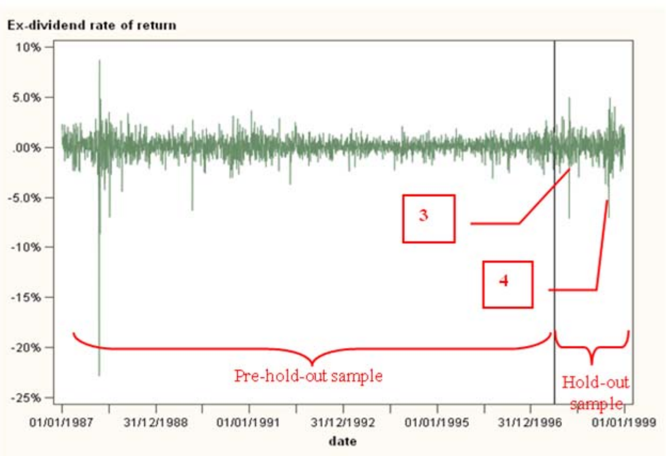
[29] SPY has been considered the most liquid ETF and with relatively small tracking errors. We use the mid-point of bid-ask prices.

Figure 2: Return time series for holdout and pre-holdout samples

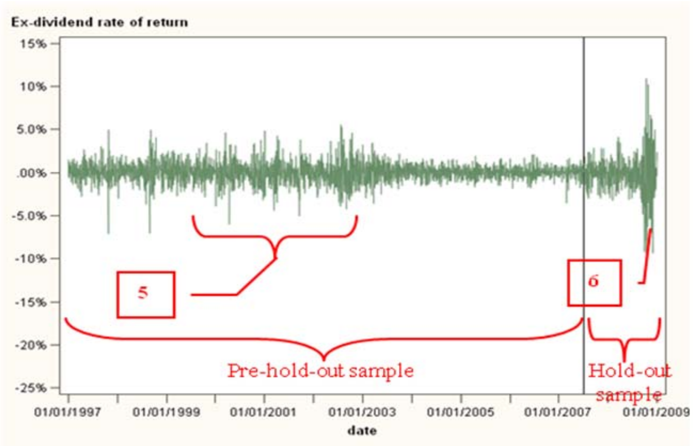
A. July 1, 1977 – June 30, 1990



B. July 1, 1987 – December 31, 1998



C. July 1, 1997 – December 31, 2008



D. Major market turbulences

ID	Time / periods	Description
1	October 19, 1987 and afterwards	"Black Monday" market crash
2	October 13, 1989 and afterwards	"Black Friday" market crash
3	Around October 27-28, 1997	Asian financial crisis
4	Around August 31, 1998	Russian crisis / LTCM failure
5	2000 – 2002	"Internet bubble"
6	2008	Global financial crisis

4 Empirical Results

In this section, we present R^2, DM test results and, as a robustness check, visualized results on forecasts and underestimation levels.

4.1 Main Results

Table 3 reports R^2, regression coefficients, b_0 and b_1, and their heteroskedasticity robust standard errors (in parentheses) of the forecast evaluation regressions. The "Relative Difference" rows indicate by how much higher, in relative scale, our models' R^2's are than those of the corresponding reference models.

As can be seen in Table 3, the R^2's for all the proposed models, in all time periods, are higher than those of their corresponding reference models. We note that this is true whether we compute realized volatility using daily returns (as in holdout sample 1) or use the intraday returns (as in holdout samples 2 and 3).[30] Typically, the R^2's of our models are around 5% to 10% higher.[31]

We also compare b_0 and b_1 in our models and the corresponding models to assess model biases [similar to, for example, Andersen, Bollerslev, Diebold and Labys (2003)]. We find that our models have no more forecast biases than the corresponding reference models.[32]

Table 3: The forecast evaluation regression results

Model	R^2	b_0	b_1
Holdout Sample 3: July 1, 2007 – December 31, 2008			
Model 1	**0.566**	0.003 (0.001)	0.906 (0.064)
GARCH(1,1)	**0.523**	0.0035 (0.001)	0.913 (0.070)
Relative Difference	**+8%**		
Model 2	**0.505**	0.0041 (0.00098)	0.86 (0.067)
RiskMetrics	**0.479**	0.0041 (0.001)	0.88 (0.069)
Relative Difference	**+5%**		
Model 3	**0.562**	0.0043 (0.00093)	0.88 (0.066)
Moving Average	**0.532**	0.0047 (0.00090)	0.86 (0.065)
Relative Difference	**+6%**		

[30] Our result is consistent with that of AB, in that R^2's for the volatility models are higher when intraday returns are used for realized volatility. For our purpose, however, the absolute magnitude of the R^2's are not of concern.

[31] We are not aware of any test that can assess the statistical significance of the difference in R^2's within the volatility literature. One can interpret the reported differences in the R^2's as the differences in economic significance of the volatility forecasts [see AB and Andersen, Bollerslev, Diebold and Labys (2003)].

[32] They provide three pieces of evidence: 1) The t-tests on b_0 and b_1's differences from 0 and 1, respectively, within the evaluation regressions give similar conclusions for our models and the corresponding reference models (although b_0 and b_1 may not accurately evaluate model bias for the reason discussed in Footnote 25); 2) The z-tests on *differences* between b_0 and b_1 of our and corresponding models, assuming model independence [see Clogg, Petkova and Haritou (1995)], show no significant differences; and 3) The tests of Gujarati (2003, pp. 306-310) with heteroskedasticity robust standard errors show no significant differences between b_0 and b_1 of our and corresponding models.

Holdout Sample 2: July 1, 1997 – December 31, 1998			
Model 1	**0.386**	0.0036 (0.0008)	0.78 (0.08)
GARCH(1,1)	**0.290**	0.0026 (0.0012)	0.86 (0.11)
Relative Difference	**+32%**	+0.001	-0.08
Model 2	**0.368**	0.0018 (0.0009)	0.89 (0.087)
RiskMetrics	**0.347**	0.002 (0.0009)	0.86 (0.086)
Relative Difference	**+6%**	-0.0002	+0.03
Model 3	**0.313**	0.0054 (0.00076)	0.63 (0.08)
Moving Average	**0.299**	0.0055 (0.00076)	0.62 (0.08)
Relative Difference	**+5%**	-0.0001	+0.01
Holdout Sample 1: July 1, 1987 – June 30, 1990			
Model 1	**0.1495**	0.0011 (0.0015)	0.71 (0.179)
GARCH(1,1)	**0.1385**	0.0005 (0.0015)	0.72 (0.163)
Relative Difference	**+8%**	+0.0006	-0.01
Model 2	**0.126**	0.0025 (0.00085)	0.486 (0.091)
RiskMetrics	**0.124**	0.0026 (0.00085)	0.477 (0.090)
Relative Difference	**+1.6%**	-0.0001	+0.01
Model 3	**0.122**	0.0037 (0.00087)	0.42 (0.1)
Moving Average	**0.119**	0.0036 (0.00090)	0.44 (0.1)
Relative Difference	**+2.5%**	+0.0001	-0.02

Next, we conduct the Diebold-Mariano [Diebold and Mariano (2002), DM] tests on the significance of "squared error" differences between our models and the reference models. For the sake of brevity, we only include the results for Model 1 and GARCH(1,1).[33] For high-volatility periods' forecasts, due to small sample size, we conduct exact finite-sample tests, the Sign test, and Wilcoxon's Signed-Rank test (with Binomial and Normal sampling distributions, respectively, reported in Panel B of Table 4). Otherwise, we use the asymptotic test (reported in Panels A and C of Table 4).

The MSE for Model 1 (Table 4, Panel A) are smaller in all holdout samples, typically by 5%-10%. In holdout samples 2 and 3, they are statistically significant at 1% and 10% confidence levels, respectively. The MSE is smaller, but not significantly, in holdout sample 1, which uses squared daily returns as realized volatility (because intraday returns are not

[33] The tables and graphs for other models are qualitatively similar.

available).

We further split the holdout samples into "normal" and high-volatility periods.[34] During high-volatility periods (Table 4, Panel B), the outperformance of Model 1 is statistically significant in holdout samples 2 and 3. There is a tie in holdout sample 1, whose sample size is very small. During "normal" volatility periods (Table 4.C), Model 1 performs better (by the negative sign of the asymptotic test statistics), but not significantly, except for holdout sample 3. Thus, the DM tests show that Model 1's outperformance comes mainly from the high-volatility periods, which are exactly when volatility models are needed most for risk management. Overall, these tests confirm that Model 1's forecasts are "closer" to the true realized volatility than GARCH(1,1)'s, both in general and in extreme cases.

Table 4: Squared errors of volatility forecasts, Model 1 vs. GARCH(1,1)

A. Full holdout sample MSE of the forecasts

Holdout Samples/Periods	Model 1 MSE	GARCH(1,1) MSE	Difference (%)	DM Asymptotic Test Statistics	p-value
3: July 1, 2007 – December 31, 2008	0.963×10^{-4}	1.072×10^{-4}	-10.17%	-3.52	<0.001
2: July 1, 1997 – December 31, 1998	0.276×10^{-4}	0.302×10^{-4}	-8.61%	-1.44	0.07
1: July 1, 1987 – June 30, 1990	1.27×10^{-4}	1.306×10^{-4}	-2.76%	-0.54	0.30

B. DM tests on forecasts during high-volatility periods

Holdout Samples/Periods	Sample Size	DM Sign Test Statistics	p-value	DM Signed-Rank Test Statistics	p-value
3: July 1, 2007 – December 31, 2008	65	19	<0.001	-3.35	<0.001
2: July 1, 1997 – December 31, 1998	10	2	0.05	-2.50	0.006
1: July 1, 1987 – June 30, 1990	7	4	0.77	0.17	0.57

[34] We define a day as "high-volatility" if, in holdout samples 2 and 3, the square root of an intraday realized volatility is $\geq 3\%$, or, in holdout sample 1, the absolute value of a realized rate of return is $\geq 5\%$. Results are robust if we slightly change the thresholds.

C. DM tests on forecasts excluding high-volatility periods

Holdout Samples/Periods	Sample Size	DM Asymptotic Test Statistics	p-value
3: July 1, 2007 – December 31, 2008	315	-1.70	0.044
2: July 1, 1997 – December 31, 1998	370	-0.26	0.39
1: July 1, 1987 – June 30, 1990	750	-0.57	0.28

4.2 Robustness Check

To ensure that the results in Section 4.1 are not statistical artifacts, we perform a robustness check by plotting, in Figure 3, the model forecasts vis-a-vis realized volatility for all holdout periods.[35] This also facilitates evaluation of model underestimation levels during high-volatility periods.

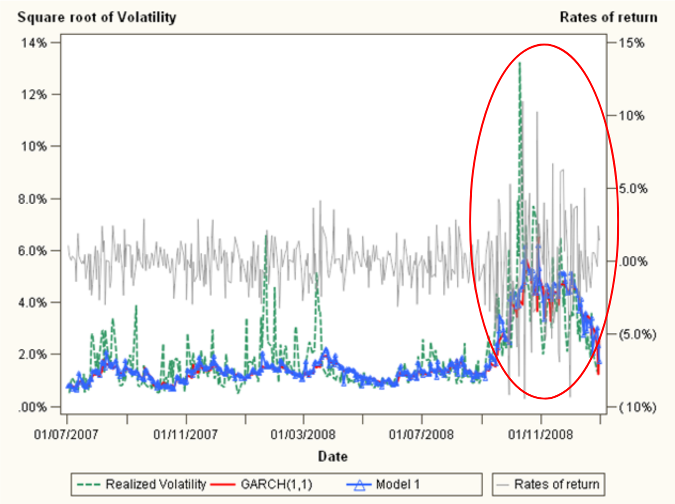
Throughout Figure 3, the forecasts of Model 1 are the blue solid line with triangle markers; those of GARCH(1,1) are the red solid line (with no markers), and realized volatility is the green dashed line. Within each period, the graph on the left plots both models' forecasts and realized volatility with vertical axis labeling on the left, and plots the actual daily rates of return (as a grey line) with vertical axis labeling on the right. We highlight the periods with high volatility by (red) circles. The graph on the right depicts the forecasts during the periods with the highest 5% realized volatility. We provide data of forecasts, realized volatility, underestimation levels, and actual daily rate of returns during high-volatility periods, at the bottom.

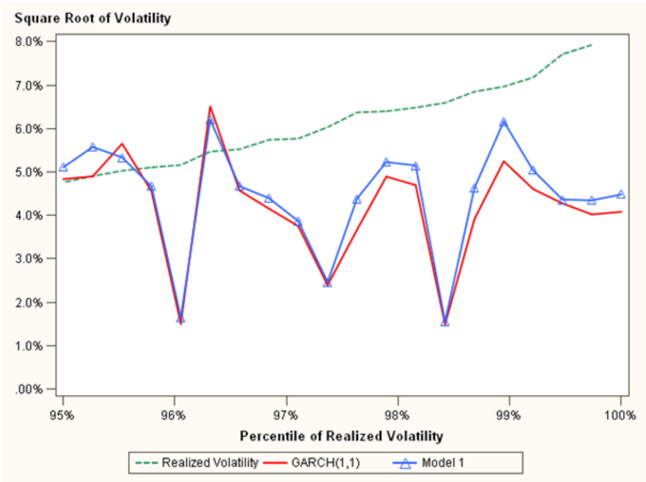
As shown in Figure 3, while Model 1's forecasts are close to those of GARCH(1,1) in "normal" periods, during high-volatility periods they are consistently higher, typically better by 10%-20%, than the GARCH(1,1) forecasts that typically underestimate realized volatility by 30%-60%. We note that this is true whether the realized daily return is positive or negative.

[35] We also did another robustness check, by using averages of squared daily returns in the following 5 days as a measure of realized volatility, following Boudoukh, Richardson and Whitelaw (1997). The results (not shown here) are similar.

Figure 3: Forecasts in holdout samples, Model 1 vs. GARCH(1,1)

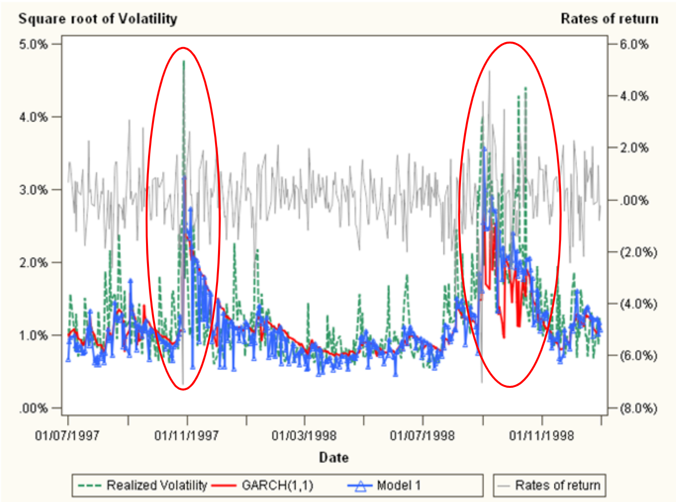
A: Holdout Sample 3, July 1, 2007 – December 31, 2008

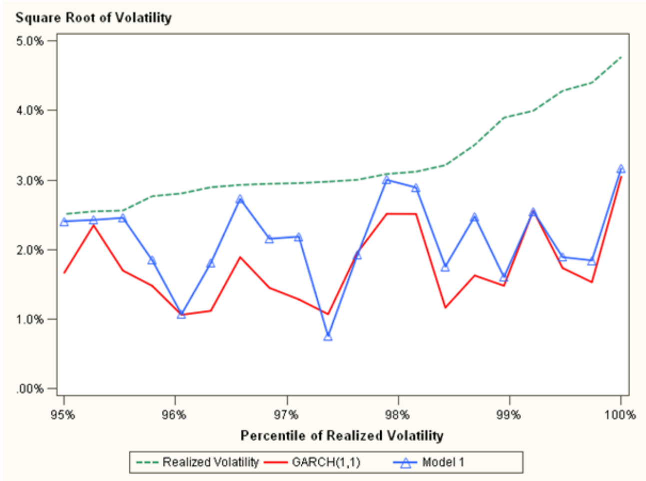




Date	(Model 1 Forecast)$^{1/2}$	Underesti-mate by	[GARCH(1,1) Forecast]$^{1/2}$	Underestim ate by	(Realized Volatility)$^{1/2}$ (intraday return)	Realized Rate of Return
Oct. 10, 2008	4.50%	65%	4.10%	68%	13%	-1.18%
Oct. 08, 2008	4.30%	46%	4.00%	49%	7.90%	-1.14%
Oct. 24, 2008	4.40%	43%	4.30%	44%	7.70%	-3.51%
Oct. 27, 2008	5.00%	31%	4.60%	36%	7.20%	-3.23%
Oct. 16, 2008	6.20%	11%	5.20%	26%	7%	4.16%
Oct. 13, 2008	4.60%	32%	3.90%	43%	6.80%	10.96%
Jan. 22, 2008	1.60%	76%	1.50%	77%	6.60%	-1.11%
Nov. 21, 2008	5.10%	22%	4.70%	28%	6.50%	6.13%
Oct. 23, 2008	5.20%	19%	4.90%	23%	6.40%	1.26%
Oct. 28, 2008	4.40%	31%	3.70%	42%	6.40%	10.25%
Sep. 29, 2008	2.50%	58%	2.40%	60%	6%	-9.20%

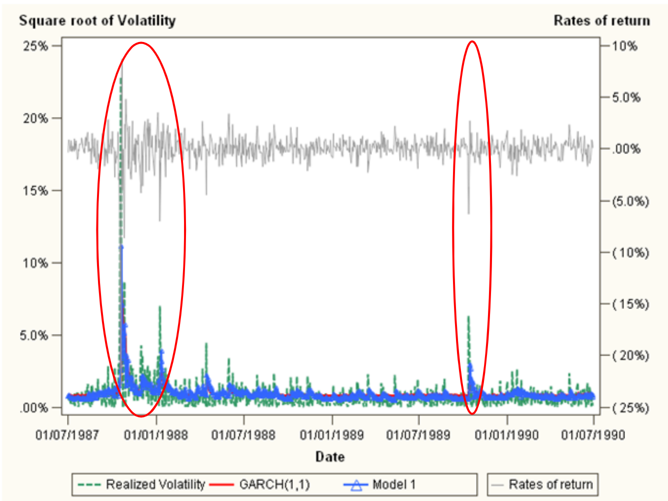
B: Holdout Sample 2, July 1, 1997 – December 31, 1998

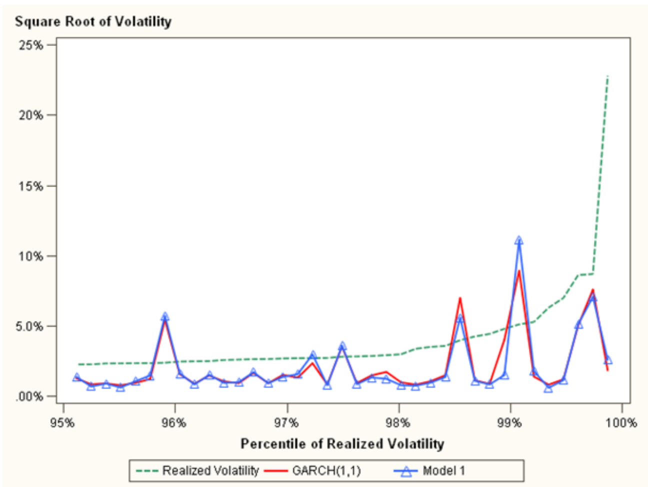




Date	(Model 1 Forecast)$^{1/2}$	Underestimate by	[GARCH(1,1) Forecast]$^{1/2}$	Underestimate by	(Realized Volatility)$^{1/2}$ (intraday return)	Realized Rate of Return
Oct. 28, 1997	3.20%	33%	3.00%	38%	4.80%	4.99%
Oct. 15, 1998	1.80%	59%	1.50%	66%	4.40%	4.09%
Oct. 08, 1998	1.90%	56%	1.70%	60%	4.30%	-1.16%
Sep. 01, 1998	2.60%	35%	2.60%	35%	4.00%	3.79%
Aug. 31, 1998	1.60%	59%	1.50%	62%	3.90%	-7.04%

C: Holdout Sample 1, July 1, 1987 – June 30, 1990





Date	(Model 1 Forecast)$^{1/2}$	Underestimate by	[GARCH(1,1) Forecast]$^{1/2}$	Underestimate by	(Realized Volatility)$^{1/2}$ (intraday return)	Realized Rate of Return
Oct. 19, 1987	2.60%	89%	1.90%	92%	--	-22.83%
Jan. 08, 1988	1.20%	83%	1.20%	83%	--	-7.01%
Oct. 13, 1989	0.61%	90%	0.84%	87%	--	-6.32%

5 Conclusions and Future Work

Market portfolio's return volatility models are vital to financial risk management. Within an equilibrium framework, we analyze two families of such models: one is a deterministic function of time, and the other is an inverse function of the market portfolio value (i.e., a CEV model). We show in theory that, under the common practice of assuming constant mean rates of return, the second family of volatility models has the features of allowing more flexible risk aversions and capturing “asymmetric volatility” [Black (1976)]. Thus, the second family tends to have more flexibility to approximate markets and, thus, better ability to forecast than the first family.

We further introduce an implementation method of proposing new parametric specifications in the second family (the proposed models), by combining the CEV formulation and the parameterizations of popular volatility models representing the first family (the reference models). We estimate them using Maximum Likelihood. We assess their forecasts by three out-of-sample evaluations and find that, while as easy to implement as the corresponding reference models, the proposed models have higher predictive power during both normal and high-volatility periods.

This paper highlights the importance of modeling the volatility and mean of the market portfolio’s rates of return in internally and equilibrium consistent manners. We find empirical evidence that substantiates both our theoretical analysis and our approach of identifying new, predictive, equilibrium-based volatility models of the market portfolio rates of return.

Our approach allows us to progress with exploring the family of equilibrium-consistent models. Particularly, we can model a stochastic process for equilibrium-based mean rates of return together with a volatility process. We can also model equilibrium interest rates and dividend yields, as observable or implied variables, and require them to be internally consistent. A notable example of this modeling strategy is the “GARCH-in-Mean” model in Bollerslev, Engle and Wooldridge (1988). For these purposes, we might also require other evaluation methodologies. Moreover, we can model market indices, such as S&P500, as a noisy proxy of the market portfolio within an Incomplete Information Equilibrium [IIE, see

Feldman (2007)]. This approach also specifies endogenously equilibrium consistent volatility structures. Finally, we can extend evaluations to forecasts within longer time horizons.[36]

Appendix A. MLE for Model 1 and Model 3

For Model 3, we obtain MLEs for (μ, S) as follows. We first write down the process for the rate of return, y_{t+1},

$$\begin{aligned} & y_{t+1} = \mu + \sigma_{t+1}\varepsilon_{t+1}; \quad \varepsilon_{t+1} \sim N(0,1); \quad E(\varepsilon_i \varepsilon_j) = 0, \quad \forall\ i \neq j \\ & \sigma_{t+1}^2 = \frac{S}{M_t} \end{aligned} \quad . \tag{A1}$$

Ignoring the constant terms, the log-likelihood functions for the parameters given one observation, and for the parameters given the N-day sample, respectively, are

$$\begin{aligned} l^s &= -\frac{1}{2}\log\sigma_s^2 - \frac{(y_s-\mu)^2}{2\sigma_s^2} \\ &= \frac{1}{2}\log M_{s-1} - \frac{1}{2}\log S - \frac{(y_s-\mu)^2 M_{s-1}}{2S}, \\ L &= -\frac{N}{2}\log S - \frac{\sum_{s=1}^{N}\left[(y_s-\mu)^2 M_{s-1}\right]}{2S}. \end{aligned} \tag{A2}$$

From the first-order condition of maximizing L, we have

$$\hat{\mu} = \frac{\sum_{s=t-N+1}^{t} y_s M_{s-1}}{\sum_{s=t-N+1}^{t} M_{s-1}}, \qquad \hat{S} = \frac{\sum_{s=t-N+1}^{t} (y_s-\hat{\mu})^2 M_{s-1}}{N}, \tag{A3}$$

and, consequently, obtain Model 3, described in Equation (19). We could have fine-tuned N (for instance, using cross-validation) to produce better forecasts, but in this study we set $N=10$ for simplicity. We note that Model 3 can be thought of as a "weighted moving average" model, weighted by the past prices, and thus is very similar to the MLE of its

[36] For example, we could test the "10-(trading) days-ahead" volatility forecasts [see, for example, Christoffersen and Diebold (2000)].

reference model, the "moving average" model (i.e., constant σ^2 and μ).

We obtain the standard error of the estimators, from their asymptotic covariance matrix, $(N\hat{\mathcal{I}})^{-1}$, where $\hat{\mathcal{I}}$ is the estimator of the information matrix. We obtain $\hat{\mathcal{I}}$ from the sample version of the expected Hessian matrix,

$$\hat{\mathcal{I}} \overset{a}{=} \begin{bmatrix} \dfrac{\sum_{s=t-N+1}^{t} M_{s-1}}{N\hat{S}} & 0 \\ 0 & \dfrac{1}{2\hat{S}^2} \end{bmatrix}, \tag{A4}$$

where $\overset{a}{=}$ denotes asymptotic equivalence. We note the similarity of the above asymptotic covariance matrix, Equation (A4), to that of the constant volatility model. From the standard error of $\hat{S}$, we can find the standard error of $\hat{\sigma}_{t+1}^{2\ \text{MLE}}$, which is $\left(\sqrt{\dfrac{2}{N}}\right)\hat{\sigma}_{t+1}^{2\ \text{MLE}}$.

For Model 1, we basically follow the estimation procedures developed in Engle (1982) and Bollerslev (1986). In a general form, denoting Θ as the unknown parameters to be estimated, we have the process for the rate of return, y_{t+1}:[37]

$$\begin{aligned} & y_{t+1} = \mu_{t+1} + \sigma_{t+1}\varepsilon_{t+1}; \quad \varepsilon_{t+1} \sim N(0,1); \quad E(\varepsilon_i \varepsilon_j) = 0, \quad \forall\ i \neq j, \\ & \mu_{t+1} = m(\Theta, \mu_t, \sigma_t^2), \\ & \sigma_{t+1}^2 = s(\Theta, \mu_t, \sigma_t^2), \end{aligned} \tag{A5}$$

where $m(\cdot)$ and $s(\cdot)$ are differentiable functions. Using Equation (A5), we can take the first derivative of μ_{t+1} and σ_{t+1}^2 w.r.t. Θ,

$$\begin{aligned} \frac{\partial \mu_{t+1}}{\partial \Theta} &= \frac{\partial m}{\partial \Theta} + \frac{\partial m}{\partial \mu_t}\frac{\partial \mu_t}{\partial \Theta} + \frac{\partial m}{\partial \sigma_t^2}\frac{\partial \sigma_t^2}{\partial \Theta}, \\ \frac{\partial \sigma_{t+1}^2}{\partial \Theta} &= \frac{\partial s}{\partial \Theta} + \frac{\partial s}{\partial \mu_t}\frac{\partial \mu_t}{\partial \Theta} + \frac{\partial s}{\partial \sigma_t^2}\frac{\partial \sigma_t^2}{\partial \Theta}. \end{aligned} \tag{A6}$$

Ignoring the constant terms, the conditional log-likelihood functions for Θ given one

[37] In this study we only consider the first-order relationships, i.e., $(\mu_{t+1}, \sigma_{t+1}^2)$ depend on (μ_t, σ_t^2) only.

observation and for Θ given the whole sample, respectively, are

$$l^t = -\frac{1}{2}\left[\log\sigma_t^2 + (y_t - \mu_t)^2(\sigma_t^2)^{-1}\right], \qquad L = \sum_{t=1}^{D} l^t , \tag{A7}$$

where D is the sample size. Taking the partial derivative of l^t w.r.t. Θ gives the score (vector) for one observation:

$$\frac{\partial l^t}{\partial\Theta} = -\frac{1}{2}\Big(\frac{\partial\sigma_t^2}{\partial\Theta}\Big)\left[(\sigma_t^2)^{-1} - (\sigma_t^2)^{-2}(y_t - \mu_t)^2\right] + \Big(\frac{\partial\mu_t}{\partial\Theta}\Big)(\sigma_t^2)^{-1}(y_t - \mu_t)\,. \tag{A8}$$

Equations (A6) and (A8) altogether define an iterative procedure to find the score (vector) for the whole sample.

We use a numerical procedure to maximize the log-likelihood function L in Equation (A7). As in Engle (1982) and Bollerslev (1986), we apply the Berndt, Hall, Hall and Hausman (1974, BHHH) algorithm to find the MLEs.[38] By the law of iterated expectations and after simplifying, the expected Hessian matrix is

$$\mathrm{E}H = -D\mathrm{E}\Big[\frac{(\sigma_t^2)^{-2}}{2}\Big(\frac{\partial\sigma_t^2}{\partial\Theta}\Big)\Big(\frac{\partial\sigma_t^2}{\partial\Theta}\Big)' + (\sigma_t^2)^{-1}\Big(\frac{\partial\mu_t}{\partial\Theta}\Big)\Big(\frac{\partial\mu_t}{\partial\Theta}\Big)'\Big], \tag{A9}$$

where " ' " is the "transpose" operator. The sample version of (A9) is $\hat{H} = -\sum_{t=1}^{D}\left[\frac{(\sigma_t^2)^{-2}}{2}\Big(\frac{\partial\sigma_t^2}{\partial\Theta}\Big)\Big(\frac{\partial\sigma_t^2}{\partial\Theta}\Big)' + (\sigma_t^2)^{-1}\Big(\frac{\partial\mu_t}{\partial\Theta}\Big)\Big(\frac{\partial\mu_t}{\partial\Theta}\Big)'\right]$. In the implementation, we use the robust asymptotic covariance matrix (i.e., the "sandwich" form). Denoting $\hat{V} = \sum_{t=1}^{D}\Big(\frac{\partial l^t}{\partial\Theta}\Big)\Big(\frac{\partial l^t}{\partial\Theta}\Big)'$ as the "Outer Product of Gradient" estimator of the Hessian matrix, the robust asymptotic covariance matrix is $\hat{H}^{-1}\hat{V}\hat{H}^{-1}$.

We set the initial value, σ_1^2, to be the long-run mean of σ_t^2, $\mathrm{E}\sigma_t^2$.[39] By the

[38] We implement the BHHH algorithm by supplying an "Outer Product of Gradient" estimator of the Hessian matrix to a built-in optimization procedure in MATLAB.

[39] In the implementation of GARCH(1,1), we set the initial value as $\sigma_1^2 = \frac{\alpha_0}{1-\alpha_1-\alpha_2}$, which is standard in practice.

assumptions of i.i.d. ε_t, $\forall t$ and continuous compounding, we take the unconditional expectation on both sides of Equation (13) and obtain

$$\mathrm{E}\sigma_t^2 = \frac{\alpha_0}{M_0}\exp\left[(-t+1)\mu + \frac{1}{2}\sum_{s=1}^{t-1}\mathrm{E}\sigma_s^2\right] + (\alpha_1+\alpha_2+\mu^2-\mu)\mathrm{E}\sigma_{t-1}^2 + \mathrm{E}(\sigma_{t-1}^2)^2. \quad \text{(A10)}$$

Further assuming that M_0 is finite, that the unconditional variance of σ_t^2 exists, and that $\mathrm{E}\sigma_t^2 < 2\mu$, asymptotically (i.e., $t \to \infty$) we have

$$\mathrm{E}\sigma_t^2 \overset{a}{=} \frac{\mathrm{E}[(\sigma_t^2)^2]}{1+\mu-\mu^2-\alpha_1-\alpha_2}. \quad \text{(A11)}$$

In the implementation, we use Equation (A11) as σ_1^2, where $\mathrm{E}[(\sigma_t^2)^2]$ is approximated by the forth moment of the sample, $\frac{1}{D}\sum_{t=1}^{D}(y_t - \frac{1}{D}\sum_{s=1}^{D}y_s)^4$.[40]

REFERENCES

Allen, L., Boudoukh, J. and Saunders, A., 2003, *Understanding Market, Credit, and Operational Risk: The Value at Risk Approach*, Wiley-Blackwell.

Andersen, T.G. and Bollerslev, T., 1998, "Answering the skeptics: Yes, standard volatility models do provide accurate forecasts," *International Economic Review*, 39, 885-905.

Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P., 2003, "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 579-625.

BCBS, 2011, "Messages From the Academic Literature on Risk Measurement for the Trading Book," Working Paper No. 19, Basel Committee on Banking Supervision.

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

Bick, A.,1990, "On Viable Diffusion Price Processes of the Market Portfolio," *The Journal of Finance,* 45, 673-689.

Black, F., 1976, "Studies in Stock Price Volatility Changes," In *Proceedings of the 1976 Business Meeting of the Business and Economic Statistics Section*, 177-181, American Statistical Association.

Bollerslev, T., 1986, "Generalized Autoregressive Conditional Heteroskedasticity," *The Journal of Econometrics,* 31, 307-27.

Bollerslev, T., Engle, R.F. and Wooldridge, J. M., 1988, "A Capital Asset Pricing Model with Time-Varying Covariances," *The Journal of Political Economy*, 96, 116-131.

[40] We only need a rough approximation here because the initial value is not essential for estimation or forecasting due to the large sample size D.

Boudoukh, J., Richardson, M. and Whitelaw, R., 1997, "An Investigation of a Class of Volatility Forecasting Models," *Journal of Derivatives*, 4, 63-71.

Christoffersen, P.F. and Diebold, F.X., 2000, "How Relevant Is Volatility Forecasting For Financial Risk Management?" *The Review of Economics and Statistics*, 82, 12-22.

Christoffersen, P.F., Hahn, J. and Inoue, A., 2001, "Testing and Comparing Value-at-Risk Measures," *Journal of Empirical Finance*, 8, 325-342.

Clogg, C.C., Petkova E. and Haritou, A., 1995, "Statistical Methods for Comparing Regression Coefficients Between Models," *American Journal of Sociology*, 100, 1261-1293.

Corradi, V., 2000, "Reconsidering the Continuous Time Limit of the GARCH(1,1) Process," *Journal of Econometrics*, 96, 145-153.

Cox, J.C. and Ross, S.A., 1976, "The Valuation of Options for Alternative Stochastic Processes," *Journal of Financial Economics*, 3, 145–166.

Cox, J.C., Ingersoll, J.E. and Ross, S.A., 1985a, "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica*, 51, 363-383.

Cox, J.C., Ingersoll, J.E. and Ross, S.A., 1985b, "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385-407.

Crouhy, M., Galai, D. and Mark, R., 2001, *Risk Management*, McGraw-Hill.

Diebold, F.X. and Mariano, R.S. 2002, "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 20, 134-144.

Engle, R.F., 1982, "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987-1007.

Feldman, D., 2007, "Incomplete Information Equilibria Separation Theorems and Other Myths," *Annals of Operations Research*, Special Issue on Financial Modeling,151, 119-149.

Fornari F. and Mele, A., 2006, "Approximating volatility diffusions with CEV-ARCH models," *Journal of Economic Dynamics and Control*, 30, 931–966.

Friedman, M, 1953, *Essays in Positive Economics*, The University of Chicago Press.

Gujarati, D.N., 2003, *Basic Econometrics*, 4th ed., McGraw-Hill.

He, H. and Leland, H., 1993, "On Equilibrium Asset Price Processes," *The Review of Financial Studies*, 6, 593-617.

Jorion, P., 2007, *Value at Risk: The New Benchmark for Managing Financial Risk*, 3rd ed., McGraw-Hill.

J. P. Morgan/Reuters, 1996, *RiskMetrics™ – Technical Document*, 4th ed., Morgan Guaranty Trust Company and Reuters Ltd.

Litterman, R. and Winkelmann, K., 1998, "Estimating Covariance Matrices," Risk Management Series, Goldman, Sachs & Co.

Liptser, R. S. and Shiryaev, A.N., 2001, *Statistics of Random Processes, I. General Theory*, 2nd ed., Springer-Verlag.

Merton, R., 1971, "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," *Journal of Economic Theory*, 3, 373-413.

Merton, R., 1973, "An Intertemporal Capital Asset Pricing Model," *Econometrica*, 41, 867-887.

Mincer, J. and Zarnowitz, V., 1969, "The Evaluation of Economic Forecasts," in Zarnowitz, J. (Ed.), *Economic Forecasts and Expectations*, National Bureau of Economic Research.

Nelson, D.B., 1990, "ARCH Models as Diffusion Approximation," *Journal of Econometrics*, 45, 7-38.

Patton, A.J., 2011, "Volatility Forecast Comparison using Imperfect Volatility Proxies," *The Journal of Econometrics*, 160, 246-256.

Phelan, M.J., 1995, "Probability and Statistics Applied To the Practice of Financial Risk Management: The Case of JP Morgan's RiskMetrics™," Working Paper, Wharton School, the University of Pennsylvania.

Schorfheide, F., Song, D. and Yaron, A, 2013, "Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach," Working Paper, University of Pennsylvania.

Singleton, K.J., 2006, *Empirical Dynamic Asset Pricing: Model Specification and Econometric Assessment*, Princeton University Press.

Wiener, Z., 1999, "Introduction to VaR," *Risk Management and Regulation in Banking*, 47-63, Kluwer Academic Publishers.

Xu, X., 2013, "Modeling Risk of the Multi-Period Market Portfolio: An Equilibrium-Based Approach," PhD Thesis, University of New South Wales, Australia. The link for the thesis is: http://primoa.library.unsw.edu.au/primo_library/libweb/tiles/lrs/unsworks/datastream.jsp?pid=UNSWorks_ALMA11687.

Zahavi, A. and Zahavi, A., 1997, *The Handicap Principle: A missing piece of Darwin 's puzzle*, Oxford University Press, New York.

Zumbach, G., 2007, "The RiskMetrics 2006 Methodology," Technical Report, RiskMetrics Group.