Simple Estimators for the GARCH(1,1) Model[1]

Todd Prono[2]
Kogod School of Business, American University

This Version: June 2014

Abstract

I propose closed-form estimators for the GARCH(1,1) model that are based on second-order covariances. The ability to obtain closed-form estimates derives from skewness in the sequence being modeled, which permits separate identification and estimation of the ARCH and GARCH effect. The proposed estimators can be close competitors to the QMLE. I demonstrate consistency and asymptotic normality of these estimators using properties of weak dependence and martingale limit theory. I also demonstrate conditions under which an iterative GLS estimator reliant on these closed-form estimates as starting values shares the same asymptotic distribution with the QMLE. This asymptotic equivalence is achievable under only slightly stronger than third moment existence for the data series being modeled, which represents a substantial departure from the moment existence criteria generally required for OLS- and TSLS-style estimators of GARCH processes. The estimators are studied in Monte Carlo experiments and applied to (i) intraday EUR/USD returns, (ii) weekly equity return innovations, and (iii) risk factors spanning the investment opportunity set. The MC and empirical results are benchmarked against the QMLE.

Keywords: GARCH, GMM, closed-form, many moments, skewness. JEL codes: C13, C22, C53.

[1] I owe thanks to Celso Brunetti, Stephen Kane, Dennis Kristensen, David Rapach, David Reiffen, Aaron Smith, and two anonymous referees, as well as seminar participants at New York Camp Econometrics VI, the 2011 NBER-NSF Time Series Conference, the 2013 Midwest Econometrics Group meeting, the Federal Reserve Board, the Commodity Futures Trading Commission, the Office of the Comptroller of the Currency, and American University. Any errors are mine.

[2] Corresponding Author: Todd Prono, Department of Finance and Real Estate, Kogod School of Business at American University, 4400 Massachusetts Ave, N.W., Washington, DC 20016. (202) 885-2745, tprono@american.edu.

1. Introduction

Because of the pioneering work by Mandelbrot (1963), financial returns are known to display non-normalities like skewness and leptokurtosis. When estimating stochastic volatility processes, generally, and GARCH processes, specifically, these stylized facts are typically treated as nuisance parameters, their existence complicating (asymptotic) efficiency and, at times, even consistency (see Newey and Steigerwald, 1997). In contrast to this typical treatment, I show in this paper that skewness forms the basis for entirely closed-form estimators of the GARCH(1,1) model. Specifically, it is the presence of skewness that identifies these closed-form estimators and enables consistency and asymptotic normality (AN) to each follow in a straightforward fashion.

Despite a plethora of alternative volatility models, the standard GARCH(1,1) model of Bollerslev (1986) remains the workhorse of conditional heteroskedasticity (CH) modeling in financial economics. Explaining its popularity, in part, are the findings of Hansen and Lunde (2005), who demonstrate that nothing outperforms the GARCH(1,1) in modeling exchange rate volatility. The most common estimator for the GARCH model is the QMLE. Properties of this estimator are well-studied (see, e.g., Weiss, 1986 and Lumsdaine, 1995, 1996 for treatment of the strong GARCH case, and Lee and Hansen, 1994 and Escanciano, 2009 for the semi-strong case). However, GMM estimation is also possible. Bollerslev and Wooldridge (1992) state, for instance, that the "results of Chamberlain (1982), Hansen (1982), White (1982), and Cragg (1983) can be extended to produce an instrumental variables estimator asymptotically more efficient than QMLE under nonnormality" (p. 5-6). Skoglund (2001) studies this result in detail for the strong GARCH(1,1) case. A drawback of extending this result to the semi-strong case is the need for preliminary (non)parametric estimates of both the third and fourth conditional moments, a need that renders estimation nontrivial.

Kristensen and Linton (2006) proposed a closed-form estimator for the GARCH(1,1) model based on second-order autocorrelations. Being closed form, this estimator is free from the need to choose (arbitrary) starting values and rely on numerical optimization techniques. It also does not require explicit treatment of the third and fourth conditional moments. In this paper, I both propose closed-form estimators for the GARCH(1,1) model that are based on second-order covariances and offer a generalization of the Kristensen and Linton estimator. These covariance-based estimators rely on skewness in the distribution of the sequence being modeled to separately identify and estimate the ARCH and GARCH effect. As a consequence of this reliance, these estimators are consistent given slightly stronger than third moment existence and asymptotically normal given slightly stronger than sixth-moment existence. While these moment-existence criteria are strong relative to the QMLE case, they are mild relative to the second-order autocorrelation case, which requires slightly stronger than fourth-moment existence and eighth-moment existence, respectively. As a result, the covariance-based estimators I propose admit heavier-tailed processes than the autocorrelation-based estimators.

Closed-form estimators for the GARCH(1,1) model are not only interesting in their own right, they are also useful in providing consistent starting values for more complicated estimators like the QMLE and fully-efficient GMM. As an additional example of this usefulness, I show that an iterative GLS estimator shares the same asymptotic distribution with the QMLE in the ARCH(1) case and approximately the same asymptotic distribution in the GARCH(1,1) case, when the GARCH parameter is (relatively) small. This asymptotic result requires only slightly stronger than third moment existence of the time-series under consideration. In other words, even if the data displays heavy tails for which the fourth moment does not appear to be well defined, statistical inference can still proceed exactly as under the familiar QMLE case.

The estimators I propose are studied in Monte Carlo experiments, where it is found that

1. At low levels of GARCH persistence, the closed-form estimators are close competitors to the QMLE (in terms of efficiency).[3]

2. For the closed-form estimators, as the ARCH effect increases in size, the efficiency of the estimated GARCH effect improves (sometimes, markedly so).

3. At low levels of GARCH persistence, the iterative GLS estimator matches the performance of the QMLE, with the same performance comparison holding true across all

[3] Here, GARCH persitence is defined as the sum of the ARCH and GARCH terms. Low GARCH persistence refers to cases where that sum is well inside of one.

levels of ARCH persistence.

4. Even when the moment existence criteria underlying the asymptotic properties of these estimators are violated, the estimators remain consistent (in finite samples).

A complete review of these Monte Carlo experiments is found in this paper's Supplemental Appendix.

The estimators I propose are also applied to several financial return series; specifically, intraday FX spot and futures returns, equity return innovations, and risk factors spanning the investment opportunity set in a version of Merton's (1973) Intertemporal CAPM (ICAPM). The FX spot and futures returns applications involve very large T. The closed-form estimators produce point estimates quite comparable to those of the QMLE in a fraction of the required computing time. As a consequence, these estimators appear well suited to market microstructure applications, since "news arrivals and the resolution of their informational impact are intimately related to the dynamics of the return volatility process" (p. 116, Anderson and Bollerslev, 1997), and (very) high sampling frequencies (with the accompanying large T) tend to be the norm in these studies. Moreover, high-frequency applications of the time-series momentum trading strategies investigated by Moskowitz, Ooi, and Pedersen (2012) that determine individual position-size within the portfolio as being inversely proportional to the asset's ex ante volatility (conveniently estimated by univariate GARCH models) are made feasible with the aid of this paper's estimators, as are high-frequency options trading strategies based on the Heston and Nandi (2000) closed-form GARCH option valuation formula.

2. Identification

For the sequence $\{Y_t\}_{t\in\mathbb{Z}}$, let $\mathcal{F}_t$ be the associated σ-algebra where $\mathcal{F}_{t-1} \subseteq \mathcal{F}_t \subseteq \cdots \subseteq \mathcal{F}$. The model is

$$E\left[Y_t \mid \mathcal{F}_{t-1}\right] = 0, \qquad E\left[Y_t^2 \mid \mathcal{F}_{t-1}\right] = h_t, \tag{1}$$

where

$$h_t = \omega_0 + \alpha_0 Y_{t-1}^2 + \beta_0 h_{t-1}. \tag{2}$$

In what follows, ω_0 denotes the true value, ω any one of a set of possible values, and $\widehat{\omega}$ an estimate (with parallel definitions holding for all other parameter values). (1) and (2) characterize a semi-strong GARCH process (see Drost and Nijman, 1993, Definition 2). This characterization nests the strong GARCH case, where

$$\epsilon_t = \frac{Y_t}{h_t^{1/2}} \sim i.i.d.\ D\left(0,1\right) \tag{3}$$

for some known distribution D.

Let $X_t \equiv Y_t^2$. From (2) it follows that

$$\begin{aligned} X_t &= h_t + W_t \\ &= \omega_0 + \phi_0 X_{t-1} + W_t - \beta_0 W_{t-1}, \end{aligned} \tag{4}$$

where $\{W_t\}$ is a martingale difference sequence (MDS), and $\phi_0 = \alpha_0 + \beta_0$.

ASSUMPTION A1: $\sigma_0^2 \equiv E\left[Y_t^2\right] = \frac{\omega_0}{1-(\alpha_0+\beta_0)}$, *where* $0 < \sigma_0^2 < \infty$, $\alpha_0 > 0$, *and* $\beta_0 \geq 0$.

Under A1, h_t is strictly positive, and Y_t is covariance stationary (see Theorem 1 of Bollerslev, 1986). Let $\widetilde{X}_t \equiv X_t - \sigma_0^2$. From (4) it follows that

$$\begin{aligned} \widetilde{X}_t &= \phi_0 \widetilde{X}_{t-1} + W_t - \beta_0 W_{t-1} \\ &= \sum_{i=0}^{\infty} \varphi_{i,0} W_{t-i}, \end{aligned} \tag{5}$$

where $\varphi_{0,0} = 1$ and $\varphi_{i,0} = \alpha_0 \phi_0^{i-1}$ for $i = 1, 2, \ldots$.

ASSUMPTION A2: $\gamma_0 \equiv E\left[Y_t^3\right]$, *where* $0 < |\gamma_0| < \infty$.

THEOREM 1. *Let Assumptions A1 and A2 hold for the model of (1) and (2). Then*

$$E\left[\widetilde{X}_t Y_{t-1}\right] = \alpha_0 E\left[Y_t^3\right], \tag{6}$$

and

$$E\left[\widetilde{X}_t Y_{t-(k+1)}\right] = \phi_0 E\left[\widetilde{X}_t Y_{t-k}\right],\ k \geq 1. \tag{7}$$

Proof. All proofs are given in the Supplemental Appendix. ■

Theorem 1 relates the covariance between $\widetilde{X}_t$ and Y_{t-k} to γ_0 (see equation 9 in the Supplemental Appendix). Lemma 1 of Guo and Phillips (2001) establishes an analogous result for the ARCH(p) model. In contrast to Guo and Phillips, Theorem 1 is central to establishing identification of a moments-based estimator for (1) and (2) because it provides (6), which uniquely determines α_0. Separation of α_0 from β_0 in (6) is the direct consequence of a nonzero third moment. Skewness, therefore, is the key identifying assumption for the estimators I propose. Newey and Steigerwald (1997) explore the effects of skewness on the identification of CH models using the QMLE and show that given skewness, there exist conditions under which the standard QMLE is not identified. In contrast, I develop estimators that are not identified without such skewness.

ASSUMPTION A3: $\kappa_0 \equiv E\left[W_t^2\right]$, *where* $0 < \kappa_0 < \infty$.

In the case of (3), A3 is equivalent to $E\left[\left(\beta_0 + \alpha_0 \epsilon_t^2\right)^2\right] < 1$, which grants Y_t to have a finite fourth moment (see Carrasco and Chen, 2002, Corollary 6) and so strengthens A1.[4]

Given (5) and Lemma 1 in the Supplemental Appendix,

$$E\left[\widetilde{X}_t \widetilde{X}_{t-1}\right] = \left[\frac{\left(1-\phi_0\beta_0\right)\left(\phi_0-\beta_0\right)}{1-\phi_0^2}\right]\kappa_0. \tag{8}$$

Given stationarity of $E\left[\widetilde{X}_t \widetilde{X}_{t-1}\right]$ in (8), it can be shown that multiplying both sides of (5) by $\widetilde{X}_{t-k}$, for $k \geq 1$ and taking expectations produces

$$E\left[\widetilde{X}_t \widetilde{X}_{t-k}\right] = \phi_0^{k-1} E\left[\widetilde{X}_t \widetilde{X}_{t-1}\right],$$

which is analogous to (9) in the Supplemental Appendix and implies that

$$E\left[\widetilde{X}_t \widetilde{X}_{t-(k+1)}\right] = \phi_0 E\left[\widetilde{X}_t \widetilde{X}_{t-k}\right]. \tag{9}$$

The result in (9) immediately above compliments (7), provided, of course, that A3 holds.

[4] A3, of course, strengthens A1 regardless of whether (3) holds. In the absence of (3), however, the way in which it does so is generally unknown, owing to possible dependence in the fourth moment of ϵ_t.

Notice that (9) identifies ϕ_0, but this result is not sufficient for identifying β_0 in (8), given the presence of κ_0. As a result, autocovariances of $\widetilde{X}_t$ are not sufficient for identifying the GARCH(1,1) model.

Let $\rho(k) = \frac{E[\widetilde{X}_t \widetilde{X}_{t-k}]}{E[\widetilde{X}_t^2]}$ for $k \geq 1$. From (8) and Lemma 1 in the Supplemental Appendix,

$$\rho(1) = \frac{(1 - \phi_0 \beta_0)(\phi_0 - \beta_0)}{1 + \beta_0^2 - 2\phi_0 \beta_0}. \tag{10}$$

Let $b_0 = \frac{\phi_0^2 + 1 - 2\rho(1)\phi_0}{\phi_0 - \rho(1)}$, and assume that $\beta_0 > 0$. Then, the unique solution to (10) is $\beta_0 = \frac{b_0 - \sqrt{b_0^2 - 4}}{2}$ (see Kristensen and Linton, 2006). As a result, autocorrelations of $\widetilde{X}_t$ are sufficient for identifying the GARCH(1,1) model.[5] (9) and (10), therefore, are an analog to (6) and (7). Basing identification on the latter pair of moment equations, however, has the advantage of not requiring A3. Moreover, identification based on (6) and (7) (and, potentially, 9) nests that ARCH(1) case, while (9) and (10) does not.

3. Estimation

3.1. Notation

Let $S_T = \{Y_t\}_{t=1}^T$, $Z_{1,t-2} = \left[Y_{t-2}, \cdots, Y_{t-(k+1)}\right]'$ and $Z_{2,t-2} = \left[Y_{t-2}^2 - \sigma^2, \cdots, Y_{t-(k+1)}^2 - \sigma^2\right]'$ for $k = 1, \ldots, K-1$, $Z_{t-2} = \begin{pmatrix} Z_{1,t-2} \\ Z_{2,t-2} \end{pmatrix}$, $\theta = (\sigma^2, \alpha, \beta)'$, and $\vartheta = (\sigma^2, \phi)'$. In addition, let $\{C_m : m \geq 1\}$ and $\{c_r\}_{r=1}^m$ denote sequences of constants that may take different values in different places.

Consider the following vector valued functions

$$g_{1,t}(S_T;\ \theta) = \left(Y_t^2 - \sigma^2\right) Y_{t-1} - \alpha Y_t^3, \tag{11}$$

$$g_{2,t}(S_T;\ \theta) = \left(Y_t^2 - \sigma^2\right)\left(Z_{1,t-2} - \phi Z_{1,t-1}\right),$$

$$g_{3,t}(S_T;\ \theta) = \left(Y_t^2 - \sigma^2\right)\left(Z_{2,t-2} - \phi Z_{2,t-1}\right),$$

[5] $\rho(k+1) = \phi_0 \rho(k)$, of course, comes from (9).

and the following definitions

$$
\begin{aligned}
g_{i,t}\left(S_T;\ \theta\right) &= g_{i,t}\left(\theta\right),\ i=1,2,3, \\
g_t\left(\theta\right) &= \left[g_{n,t}\left(\theta\right)\right],\ n=2,\ldots,N,\ 2\leq N\leq 3, \\
\widehat{g}\left(\theta\right) &= T^{-1}\sum_{t=k+1}^{T} g_t\left(\theta\right), \qquad \overline{g}\left(\theta\right)=E\left[g_t\left(\theta\right)\right], \\
\widehat{S}_{\theta_i}\left(\theta\right) &= \frac{\partial\widehat{g}\left(\theta\right)}{\partial\theta_i}, \qquad S_{\theta_i}\left(\theta\right)=E\left[\frac{\partial g_t\left(\theta\right)}{\partial\theta_i}\right], \\
\widehat{S}_{\theta_{ij}}\left(\theta\right) &= \frac{\partial\widehat{g}\left(\theta\right)}{\partial\theta_i\partial\theta_j}, \qquad S_{\theta_{ij}}\left(\theta\right)=E\left[\frac{\partial g_t\left(\theta\right)}{\partial\theta_i\partial\theta_j}\right],
\end{aligned}
$$

3.2. Consistency and Asymptotic Normality

Consider the multi-step estimator

$$
\widehat{\sigma}^2=T^{-1}\sum_t Y_t^2, \qquad \widehat{\alpha}=\frac{\sum_t\left(Y_t^2-\widehat{\sigma}^2\right)Y_{t-1}}{\sum_t Y_t^3}, \tag{12}
$$

and

$$
\widehat{\beta}=\arg\min_{\beta\in B}\ \widehat{g}\left(\widehat{\sigma}^2,\widehat{\alpha},\beta\right)'\Lambda_T\widehat{g}\left(\widehat{\sigma}^2,\widehat{\alpha},\beta\right), \tag{13}
$$

where Λ_T is positive semi-definite and of dimension $(N-1)(k-1)\times(N-1)(k-1)$.[6] (12) and (13) represent a variance targeting estimator (VTE).[7] (12) is the finite sample analog of (6), while (13) considers finite sample versions of (7) alone when $N=2$ and (7) and (9) stacked together when $N=3$.

If Λ_T is not dependent on β (as is the case for the familiar two-step GMM estimator of Hansen, 1982), the solution to (13) is

$$
\widehat{\beta}=\frac{\left(\sum_t\left(Y_t^2-\widehat{\sigma}^2\right)\widehat{Z}_{t-1}\right)'\Lambda_T\left(\sum_t\left(Y_t^2-\widehat{\sigma}^2\right)\left(\widehat{Z}_{t-2}-\widehat{\alpha}\widehat{Z}_{t-1}\right)\right)}{\left(\sum_t\left(Y_t^2-\widehat{\sigma}^2\right)\widehat{Z}_{t-1}\right)'\Lambda_T\left(\sum_t\left(Y_t^2-\widehat{\sigma}^2\right)\widehat{Z}_{t-1}\right)}. \tag{14}
$$

[6]From this multi-step estimator, $\widehat{\omega}$ can be retrieved as $\widehat{\omega}=\widehat{\sigma}^2\left(1-\widehat{\phi}\right)$.

[7]See Engle and Mezrich (1996) as well as Francq, Horath, and Zakoïan (2011) for QMLE analogs.

In this case, (12) and (14) describe an entirely closed-form estimator for the GARCH(1,1) model.

Through its dependence on K, (13) involves (potentially) many moment conditions. While higher values of K should increase asymptotic efficiency, they can also lead to (substantial) finite sample bias. Solutions to this problem include jackknife two-step GMM estimation and the CUE of Hansen, Heaton, and Yaron (1996).[8] The former results in a JIVE2-like version of (14) (see Angrist, Imbens, and Kreuger, 1999), thus preserving the multi-step estimator's closed form. The latter renders (13) no longer closed form; however, since β_0 is a scalar with bounded support $[0,\ 1)$ by assumption, (13) can be solved simply using a grid search.

In a slight change of notation, let $g_t(\vartheta) = \left[g_{3,t}(\vartheta)\right]$ so that $\widehat{g}(\vartheta)$ only contains sample second-order autocovariances of Y_t. Then consider the alternative closed-form estimator

$$\begin{aligned}\widehat{\phi} &= \arg\min_{\phi\in\Phi} \widehat{g}\left(\widehat{\sigma}^2, \phi\right)' \Lambda_T \widehat{g}\left(\widehat{\sigma}^2, \phi\right) \\ &= \frac{\left(\sum_t \left(Y_t^2 - \widehat{\sigma}^2\right)\widehat{Z}_{2,t-1}\right)' \Lambda_T \left(\sum_t \left(Y_t^2 - \widehat{\sigma}^2\right)\widehat{Z}_{2,t-2}\right)}{\left(\sum_t \left(Y_t^2 - \widehat{\sigma}^2\right)\widehat{Z}_{2,t-1}\right)' \Lambda_T \left(\sum_t \left(Y_t^2 - \widehat{\sigma}^2\right)\widehat{Z}_{2,t-1}\right)},\end{aligned} \tag{15}$$

and

$$\widehat{b} = \frac{\widehat{\phi}^2 + 1 - 2\widehat{\rho}(1)\widehat{\phi}}{\widehat{\phi} - \widehat{\rho}(1)}, \quad \widehat{\beta} = \frac{\widehat{b} - \sqrt{\widehat{b}^2 - 4}}{2}, \quad \widehat{\alpha} = \widehat{\phi} - \widehat{\beta}, \tag{16}$$

where Λ_T, in this case, is dimension $(k-1)\times(k-1)$. (15) and (16) generalize the estimator proposed by Kristensen and Linton (2006).[9] Note that when $N = 3$, (14) nests (15). As a consequence, the same discussion involving (potential) many moments bias in (13) and its possible solutions also applies to (15).

ASSUMPTION A4: *$\theta_0 \in \Theta \subseteq \Re^3$ is in the interior of Θ, a compact parameter space. For any $\theta \in \Theta$, $\partial \leq \omega \leq W$, $\partial \leq \alpha \leq 1-\partial$, $0 \leq \beta \leq 1-\partial$, and $\alpha+\beta < 1$ for some*

[8] See Donald and Newey (2000) for the link between jackknife GMM estimation and the CUE. See Newey and Windmeijer (2009) for a discussion of how the CUE is robust to many moments bias.

[9] In Kristensen and Linton (2006), the estimator for ϕ_0 is $\overline{\phi} = \frac{\widehat{Cov}\left(Y_t^2, Y_{t-2}^2\right)}{\widehat{Cov}\left(Y_t^2, Y_{t-1}^2\right)}$.

constant $\partial > 0$, where ∂ and W are given a priori.[10]

ASSUMPTION A5: *Let $\widetilde{U}_{t,k} \equiv \widetilde{X}_t Y_{t-k} - E\left[\widetilde{X}_t Y_{t-k}\right]$ and $\widetilde{U}_{t,0} \equiv Y_t^3 - \gamma_0$. (i) $E\left[|W_t Y_t|\right] < \infty$. (ii) $\exists$ constants $\{C_m : m \geq 1\}$ such that $C_m \to 0$ as $m \to \infty$ where $\left|E\left[W_t Y_t \mid F_{t-m}\right] - \gamma_0\right| \leq C_m$. (iii) $\left\{\widetilde{U}_{t,k}\right\}$ and $\left\{\widetilde{U}_{t,0}\right\}$ are uniformly integrable.*

ASSUMPTION A6: *Let $\widetilde{V}_{t,k} \equiv \widetilde{X}_t \widetilde{X}_{t-k} - E\left[\widetilde{X}_t \widetilde{X}_{t-k}\right]$. $\left\{\widetilde{V}_{t,k}\right\}$ is uniformly integrable.*

THEOREM 2. *Consider the estimator in (12) and (13) for the model of (1) and (2). Assume that $\Lambda_T \xrightarrow{p} \Lambda_0$, a positive definite matrix. If $N = 2$, then $\widehat{\theta} \xrightarrow{p} \theta_0$ given Assumptions A1, A2, A4, and A5. If $N = 3$, then $\widehat{\theta} \xrightarrow{p} \theta_0$ given Assumptions A1–A6.*

Consistency is established under Theorem 2 using a weak LLN for L^1 mixingales (see Andrews, 1988). When $N = 2$, consistency requires slightly stronger than third moment existence. When $N = 3$, fourth moment existence is required. In addition, notice that a form of weak dependence is required for the third moment through A5(ii).

ASSUMPTION A7: *Following A6, $\widetilde{V}_{t,0} \equiv \widetilde{X}_t^2 - E\left[\widetilde{X}_t^2\right]$. (i) For $\{C_m : m \geq 1\}$ with the same general property as the sequence in A5(ii), $\left|E\left[W_t^2 \mid F_{t-m}\right] - \kappa_0\right| \leq C_m$. (ii) $\left\{\widetilde{V}_{t,0}\right\}$ is uniformly integrable.*

COROLLARY 1. *Consider the estimator in (15) and (16) for the model of (1) and (2). Let $\widehat{\sigma}^2 = T^{-1}\sum_t Y_t^2$, and assume that $\Lambda_T \xrightarrow{p} \Lambda_0$, a positive definite matrix. Then $\widehat{\theta} \xrightarrow{p} \theta_0$ given Assumptions A1–A7.*

Corollary 1 requires stronger assumptions than Theorem 2 for the same consistency result. For instance, 4th moment existence is necessary under Corollary 1. It is only sufficient under Theorem 2 (when $N = 2$). In addition, A7(i), which is a weak dependence condition for the fourth moment, has no analog in Theorem 2 regardless of the value for N.[11]

ASSUMPTION A8: *Let $\widetilde{S}_t = W_t Y_t - \gamma_0$. (i) $\left\{\widetilde{S}_t\right\}$ is a MDS. (ii) $\lim_{T\to\infty} T^{-1}\sum_{t=1}^{T} E\left[W_t^2 \mid F_{t-1}\right] = \kappa_0$, $\lim_{T\to\infty} T^{-1}\sum_{t=1}^{T} E\left[W_t^3 \mid F_{t-1}\right] = E\left[W_t^3\right]$, and, for any $r \geq 1$, $\lim_{T\to\infty} T^{-1}\sum_{t=1}^{T} W_{t-r} E\left[W_t^2 \mid F_{t-1}\right] =$*

[10] Other assumptions impose additional constraints on Θ through higher moment existence critieria (see, e.g., A3 and, additionally, A8 and A9).

[11] The second moment of W_t translates into the fourth moment of Y_t.

$E\left[W_{t-r}W_t^2\right]$. *In addition, (iii)* $T^{-1}\sum_{t=1}^{T}W_t^2 \xrightarrow{p} \kappa_0$, $T^{-1}\sum_{t=1}^{T}W_t^3 \xrightarrow{p} E\left[W_t^3\right]$, *and, for any* $r \geq 1$, $T^{-1}\sum_{t=1}^{T}W_{t-r}W_t^2 \xrightarrow{p} E\left[W_{t-r}W_t^2\right]$.

ASSUMPTION A9: *For any* $r \geq 1$ *and* $s \geq 1$, *(i)* $T^{-1}\sum_{t=1}^{T}W_{t-r}W_{t-s}h_t \xrightarrow{p} E\left[W_{t-r}W_{t-s}Y_t^2\right]$, *and* $T^{-1}\sum_{t=1}^{T}Y_{t-r}Y_{t-s}E\left[W_t^2 \mid F_{t-1}\right] \xrightarrow{p} E\left[Y_{t-r}Y_{t-s}W_t^2\right]$. *In addition, (ii)* $E\left[\left|W_t^j\right|\right] < \infty$ *for* $j > 3$ *and is uniformly bounded.*

ASSUMPTION A10: $\exists$ *an* X *with* $E\left[|X^j|\right] < \infty$ *for* $3 < j \leq 4$ *such that* $P\left(|W_t| > u\right) \leq cP\left(|X| > u\right)$ *for some* $0 < c < \infty$ *and all* $u \geq 0$.

ASSUMPTION A11: *For any* $r \geq 1$ *and* $s \geq 1$, *(i)* $E\left[W_{t-r}W_{t-s}W_t^2\right]$ *is finite and uniformly bounded, and (ii)* $T^{-1}\sum_{t=1}^{T}W_{t-r}W_{t-s}E\left[W_t^2 \mid F_{t-1}\right] \xrightarrow{p} E\left[W_{t-r}W_{t-s}W_t^2\right]$. *In addition, (iii)* $\sum_{j=0}^{\infty}j^{1/2}\varphi_{i,0}^2 < \infty$.

ASSUMPTION A12: $S_{\theta_3}\left(\theta_0\right)' \Lambda_T \; S_{\theta_3}\left(\theta_0\right) > 0$.

THEOREM 3. *Consider the estimator in (12) and (13) for the model of (1) and (2). Assume that* $\Lambda_T \xrightarrow{p} \Lambda_0$, *a positive definite matrix. Let Assumptions A1, A2, A4, A5,and A8–A10 hold. Then*

$$\sqrt{T}\left(\widehat{\alpha} - \alpha_0\right) \xrightarrow{d} N\left(0, \;\; \gamma_0^{-2}\sum_{s=-(L-1)}^{s=(L-1)}E\left[g_{1,t-s}\left(\theta_0\right)g_{1,t}\left(\theta_0\right)\right]\right). \tag{17}$$

If $N = 2$ *and Assumption A12 holds, then*

$$\sqrt{T}\left(\widehat{\beta} - \beta_0\right) \xrightarrow{d} N\left(0, \;\; \frac{S_{\theta_3}(\theta_0)'\Lambda_0\widetilde{\Omega}(\theta_0)\Lambda_0 S_{\theta_3}(\theta_0)}{H(\theta_0)^2}\right), \tag{18}$$

where

$$\widetilde{g}_t\left(\theta_0\right) = \left\{g_t\left(\theta_0\right) + S_{\theta_2}\left(\theta_0\right)\gamma_0^{-1}g_{1,t}\left(\theta_0\right)\right\}, \tag{19}$$

$\widetilde{\Omega}\left(\theta\right) = \sum_{s=-(L-1)}^{s=(L-1)}E\left[\widetilde{g}_{t-s}\left(\theta\right)\widetilde{g}_t\left(\theta\right)'\right]$ *for* $L \geq 1$, $H\left(\theta_0\right) = S_{\theta_3}\left(\theta_0\right)' \Lambda_0 \; S_{\theta_3}\left(\theta_0\right)$, *and* $S_{\theta_2}\left(\theta_0\right) = S_{\theta_3}\left(\theta_0\right) = E\left[\widetilde{X}_t Z_{t-1}\right]$. *If* $N = 3$, *then (18) holds given Assumptions*

A1–A6, and A8–A12.

Given A8(i), A5(ii) is satisfied. A8(i) has the advantage of limiting the moment existence criteria needed for Theorem 3 in the case where $N = 2$. A8(i) may be relaxed to allow time-variation in the third moment of Y_t but at the cost of stricter moment existence criteria (see Theorem 3.15 of Phillips and Solo, 1992). When $N = 3$, this cost is minor given the moment existence requirements of A10 and A11.

Theorem 3 establishes AN for the multi-step estimator by examining the limiting distributions of the second-order covariances $T^{-1}\sum_{t=k+1}^{T} g_{2,t}\left(\widehat{\theta}\right)$ and $T^{-1}\sum_{t=k+1}^{T} g_{3,t}\left(\widehat{\theta}\right)$. Theorem 2 of Hannan and Heyde (1972) establishes AN of the latter given A10 and A11. A8–A10 extend the Hannan and Heyde result to $T^{-1}\sum_{t=k+1}^{T} g_{2,t}\left(\widehat{\theta}\right)$. When $N = 2$, Theorem 3 requires slightly stronger than sixth moment existence for Y_t. When $N = 3$, eighth moment existence is required as it is, generally, when second-order autocovariances are used for inference (see, e.g., Francq and Zakoïan, 2000, Giraitis and Robinson, 2001, and Kristensen and Linton, 2006). Finally, notice that the asymptotic variance of $\widehat{\sigma}^2$ impacts neither the asymptotic variance of $\widehat{\alpha}$ nor $\widehat{\beta}$, meaning that nothing is lost (asymptotically) by plugging $\widehat{\sigma}^2$ into (12) and (13) as opposed to σ_0^2. This result stands in contrast to the QMLE-based VTE studied by Francq, Horath, and Zakoian (2011).

COROLLARY 2. *Consider the estimator in (15) and (16) for the model of (1) and (2). Let $\widehat{\sigma}^2 = T^{-1}\sum_t Y_t^2$, and assume that $\Lambda_T \xrightarrow{p} \Lambda_0$, a positive definite matrix. In addition, let*

$$h_t\left(\rho\left(1\right)\right) = \widetilde{X}_t\widetilde{X}_{t-1} - \rho\left(1\right)\widetilde{X}_t^2,$$

$\widehat{\mu} = \left(\widehat{\phi},\ \widehat{\rho}(1)\right)'$, $\widehat{\beta} = f(\widehat{\mu})$, $\widehat{\alpha} = g(\widehat{\mu})$, *and* $V_\mu = \begin{bmatrix} v_{11} & v_{12} \\ v_{12} & v_{22} \end{bmatrix}$, *where*

$$\begin{aligned} v_{11} &= \frac{S_{\vartheta_2}(\vartheta_0)' \Lambda_0 \Omega(\vartheta_0) \Lambda_0 S_{\vartheta_2}(\vartheta_0)}{H(\vartheta_0)^2}, \\ v_{12} &= \frac{S_{\vartheta_2}(\vartheta_0)' \sum_s E\left[h_{t-s}(\rho(1))\, g_{3,t}(\vartheta_0)'\right]}{E\left[\widetilde{X}_t^2\right] H(\vartheta_0)}, \\ v_{22} &= \frac{\sum_s E\left[h_{t-s}(\rho(1))\, h_t(\rho(1))\right]}{E\left[\widetilde{X}_t^2\right]^2}, \end{aligned}$$

$\Omega(\vartheta_0) = \sum_{s=-(L-1)}^{s=(L-1)} E\left[g_{t-s}(\vartheta_0)\, g_t(\vartheta_0)'\right]$ *for* $L \geq 1$, $H(\vartheta_0) = S_{\vartheta_2}(\vartheta_0)' \Lambda_0 S_{\vartheta_2}(\vartheta_0)$, *and* $S_{\vartheta_2}(\vartheta_0) = E\left[\widetilde{X}_t Z_{2,t-1}\right]$. *Then, given Assumptions A1–A12,*

$$\sqrt{T}\left(\widehat{\phi} - \phi_0\right) \xrightarrow{d} N\left(\ 0,\ \ v_{11}\ \right), \tag{20}$$

and

$$\sqrt{T}\left(\widehat{\beta} - \beta_0\right) \xrightarrow{d} N\left(\ 0,\ \ V_\beta\ \right), \quad \sqrt{T}\left(\widehat{\alpha} - \alpha_0\right) \xrightarrow{d} N\left(\ 0,\ \ V_\alpha\ \right), \tag{21}$$

where

$$\begin{aligned} V_\beta &= \left(1 - \frac{b_0}{\sqrt{b_0^2 - 4}}\right)^2 \left\{c_1(\beta_0)^2 v_{11} + c_1(\beta_0)\, c_2(\beta_0)\, v_{12} + c_2(\beta_0)^2 v_{22}\right\} \\ c_1(\beta_0) &= 1 - \frac{b_0}{2(\phi_0 - \rho(1))} \\ c_2(\beta_0) &= \frac{1 - \phi_0^2}{(\phi_0 - \rho(1))^2} \end{aligned}$$

and

$$V_{\alpha} = \left(1-\frac{b_0}{\sqrt{b_0^2-4}}\right)^2\left\{c_1(\alpha_0)^2 v_{11}+c_1(\alpha_0)c_2(\beta_0)v_{12}+c_2(\beta_0)^2 v_{22}\right\}$$

$$c_1(\alpha_0) = 1+\frac{b_0^2-\left(\sqrt{b_0^2-4}\right)\{2(\phi_0-\rho(1))+b_0\}}{2(\phi_0-\rho(1))\left\{\sqrt{b_0^2-4}-b_0\right\}}$$

This corollary is related to Theorem 2 of Kristensen and Linton (2006). The difference stems from the generalization in (15), which gets reflected in the convergence result in (20).

3.3. Efficiency Issues

Given (18), the minimum asymptotic variance is achieved when $\Lambda_T = \widehat{\widehat{\Omega}}\left(\widehat{\theta}\right)^{-1}$. In this case, consistency of Λ_T follows from Lemma 4.3 of Newey and McFadden (1994) if $\left\{\widetilde{g}_{t-s}(\theta_0)\,\widetilde{g}_t(\theta_0)'\right\}$ satisfies the UWLLN (see Wooldridge, 1990, Definition A.1.).[12] An analogous result holds for (20) when $\Lambda_T = \widehat{\Omega}\left(\widehat{\vartheta}\right)^{-1}$.

Kristensen and Linton (2006) consider a GLS estimator that uses the closed-form estimates from (15) and (16) when $K = 1$ as starting values. Such an estimator reduces the problems associated with numerical optimization dependent on initial values since, in this case, those initial values are consistent. Moreover, this GLS estimator remains broadly comparable to closed-form estimators in terms of simplicity by not requiring the calculation of numerical derivatives. In this section, I consider the same GLS estimator, but generate its (consistent) starting values using either (12) and (14) or (15) and (16). Moreover, I discuss conditions under which this GLS estimator achieves the same asymptotic distribution as the QMLE.

Let $\lambda = (\omega, \alpha, \beta)'$, $S(\lambda) = \frac{\partial Q(\lambda)}{\partial \lambda}$, and $H(\lambda) = \frac{\partial^2 Q(\lambda)}{\partial \lambda \partial \lambda'}$, where $Q(\lambda)$ is the Gaussian (quasi-) log-Likelihood. In addition, let $\widehat{\lambda}^{QMLE} = \arg\max_{\lambda \in \Lambda} \widehat{Q}(\lambda)$,

$$\widehat{\lambda}_{l+1}^{NR} = \widehat{H}\left(\widehat{\lambda}_l^{NR}\right)^{-1}\left[\widehat{H}\left(\widehat{\lambda}_l^{NR}\right)\widehat{\lambda}_l^{NR} - \widehat{S}\left(\widehat{\lambda}_l^{NR}\right)\right], \quad l \geq 1, \tag{22}$$

[12] In demonstrating Lemma 4.3, let $a(z,\theta) = \widetilde{g}_{t-s}(\theta)\,\widetilde{g}_t(\theta)'$, and replace Khintchine's law of large numbers with the aforementioned UWLLN.

and

$$\widehat{\lambda}_{l+1}^{GLS} = \left(\sum_t h_t^{-2} \left(\widehat{\lambda}_l^{GLS} \right) X_{l,t-1} X_{l,t-1}' \right)^{-1} \left(\sum_t h_t^{-2} \left(\widehat{\lambda}_l^{GLS} \right) X_{l,t-1} Y_t^2 \right), \tag{23}$$

where $h_t \left(\widehat{\lambda}_l^{GLS} \right) = \widehat{\omega}_l^{GLS} + \widehat{\alpha}_l^{GLS} Y_{t-1}^2 + \widehat{\beta}_l^{GLS} h_{t-1} \left(\widehat{\lambda}_l^{GLS} \right)$, and $X_{l,t-1} = \left[1,\ Y_{t-1}^2,\ h_{t-1} \left(\widehat{\lambda}_l^{GLS} \right) \right]'$. Finally, let $\|\cdot\|$ be the Euclidean norm.

ASSUMPTION A13: $\lim_{l \to \infty} \left\| \widehat{\lambda}_{l+1}^{NR} - \widehat{\lambda}_{l+1}^{GLS} \right\| = 0.$

THEOREM 4. *Consider the estimator in (23) for the model of (1) and (2), where (12) and (14) produce both* $\widehat{\lambda}_1^{GLS}$ *and* $\widehat{\lambda}_1^{NR}$. *Then,*

$$\sqrt{T} \left(\widehat{\lambda}_{l+1}^{GLS} - \lambda_0 \right) \xrightarrow{d} N \left(0,\quad H^{-1} \Sigma H^{-1} \right), \tag{24}$$

where $H = E \left[\widehat{H} \left(\lambda_0 \right) \right]$, *and* $\Sigma = E \left[\widehat{S} \left(\lambda_0 \right) \widehat{S} \left(\lambda_0 \right)' \right]$, *given Assumptions A1, A2, A4, A5, and A13 if* $N = 2$. *If* $N = 3$, *then (24) holds given Assumptions A1–A6, and A13.*

Theorem 4 is closely related to Kristensen and Linton (2006, Theorem 3). In the case where $N = 2$, Theorem 4 shows that (23) shares the same asymptotic distribution with the QMLE (see, e.g., Lee and Hansen 1994, Theorem 2), with this result requiring slightly higher than third moment existence for $\{Y_t\}$. In the case where $N = 3$, slightly higher than fourth moment existence is required, as it also is in Kristensen and Linton (2006, Theorem 3). Both theorems rely upon A13. Given the apparent differences between (22) and (23), it seems important to discover more primitive conditions that satisfy A13. The remainder of this section is dedicated to that purpose. First, however, a corollary to Theorem 4 is stated that enables (15) and (16) to produce $\widehat{\lambda}_1^{GLS}$. This corollary is, essentially, a restatement of Kristensen and Linton (2006, Theorem 3) and, therefore, is not proved.

COROLLARY 3. *Consider the estimator in (23) for the model of (1) and (2), where (15) and (16) produce* $\widehat{\lambda}_1^{GLS}$. *Assume that* $\widehat{\lambda}_1^{NR} \xrightarrow{p} \lambda_0$. *Then, (24) holds given Assumptions A1–A7, and A13.*

Consider the additional notation

$$\begin{aligned}\widetilde{H}(\lambda) &= T^{-1}\sum_t h_t^{-2}(\lambda)\, X_{t-1}X_{t-1}', \qquad \widehat{J}(\lambda) = T^{-1}\sum_t h_t^{-2}(\lambda)\, X_{t-1}Y_t^2, \\ \frac{\partial \widehat{J}(\lambda)}{\partial \lambda} &= T^{-1}\sum_t h_t^{-3}(\lambda)\, X_{t-1}X_{t-1}'Y_t^2, \qquad \widehat{M}(\lambda) = T^{-1}\sum_t h_t^{-2}(\lambda)\, X_{t-1}W_t(\lambda),\end{aligned}$$

where $X_{t-1} = \left[1,\ Y_{t-1}^2\right]'$. If $\beta_0 = 0$ in (2), then (22) simplifies to

$$\widehat{\lambda}_{l+1}^{NR} = \widehat{H}\left(\widehat{\lambda}_l^{NR}\right)^{-1}\left[\widehat{J}\left(\widehat{\lambda}_l^{NR}\right) + \widehat{M}\left(\widehat{\lambda}_l^{NR}\right)\right]. \tag{25}$$

In addition, note that

$$\begin{aligned}H &= T^{-1}\sum_t E\left[h_t^{-2}(\lambda_0)\left\{2E\left[\epsilon_t^2(\lambda_0)\ \mid \mathcal{F}_{t-1}\right] - 1\right\} X_{t-1}X_{t-1}'\right] \\ &= E\left[h_t^{-2}(\lambda_0)\, X_{t-1}X_{t-1}'\right],\end{aligned}$$

where $\epsilon_t^2(\lambda_0) = \frac{Y_t^2}{h_t(\lambda_0)}$, and

$$\begin{aligned}M &= E\left[\widehat{M}(\lambda_0)\right] \\ &= T^{-1}\sum_t E\left[h_t^{-2}(\lambda_0)\, X_{t-1}E\left[W_t(\lambda)\ \mid \mathcal{F}_{t-1}\right]\right] = 0.\end{aligned}$$

ASSUMPTION A14: (i) $\widetilde{H}\left(\widehat{\lambda}\right) \xrightarrow{p} H$. (ii) $\widehat{J}\left(\widehat{\lambda}\right) \xrightarrow{p} J$. (iii) $\frac{\partial \widehat{J}(\lambda)}{\partial \lambda} \xrightarrow{p} \dot{J}$. (iv) $\widehat{M}\left(\widehat{\lambda}\right) \xrightarrow{p} 0$.

In Prono (2015, Lemmas 7–9), the moments underlying H, J, $\dot{J}$, and M are shown to be uniformly bounded $\forall\ \lambda \in \Lambda$. As a consequence, A14(i)–(iv) follows from the Ergodic Theorem.

THEOREM 5. *For the special case of the model in (1) and (2) where $\beta_0 = 0$, consider the estimator in (23) with $h_t\left(\widehat{\lambda}_l^{GLS}\right) = \widehat{\omega}_l^{GLS} + \widehat{\alpha}_l^{GLS}Y_{t-1}^2$ and $X_{l,t-1} = X_{t-1}$. Let (12) produce both $\widehat{\lambda}_1^{GLS}$ and $\widehat{\lambda}_1^{NR}$. Then, for sufficiently large T,*

$$\begin{aligned}\widehat{\lambda}_{l+1}^{NR} - \widehat{\lambda}_{l+1}^{GLS} &= O_p\left(\left\|\widehat{\lambda}_1^{NR} - \widehat{\lambda}_1^{GLS}\right\|^l\right) \\ &= O_p\left(T^{-l}\right)\end{aligned} \tag{26}$$

given Assumptions A1, A2, A4, A5, and A14.

The key to Theorem 5 is $\frac{\partial h_t(\lambda)}{\partial \lambda} = X_{t-1}$, which enables the transition from (22) to (25). Based on Theorem 5, one can hypothesize that (24) offers a descent approximation in cases where $\frac{\partial h_t(\lambda)}{\partial \lambda} \approx X_{t-1}$ and, thus, $\frac{\partial^2 h_t(\lambda)}{\partial \lambda \partial \lambda'} \approx 0$, which is to say that Theorem 4 provides a good approximation to the limiting behavior of $\widehat{\lambda}_{l+1}^{GLS}$ when β_0 is relatively small. This hypothesis is confirmed in the Monte Carlo experiments presented in the Supplemental Appendix.

5. Applications

Let Closed-Form Estimator 2 (CFE2) reference (12) and (14) with $N = 3$.[13] CFE3 then references (15) and (16), while GLS2 references (23) and relies on CFE2 for starting values. Based on results from the Monte Carlo studies (see the Supplemental Appendix), the next two sections apply the CFE2, CFE3, and GLS2 estimators to (i) intraday FX returns and (ii) weekly equity returns and candidate pricing factors for those equity returns. In both applications, CFE2 and CFE3 are optimal in the sense that they rely on the (asymptotic) variance-minimizing weighting matrix. In all cases, that weighting matrix sets $L = 1$. For GLS2, $l = 10$ in all cases.

In practice, the CFEs and GLS estimators may produce estimates that lie outside of $(0, 1]$. Because these estimators are higher-order estimators, the presence of extreme values may lead to such an event, and such values are certainly possible since the financial return series under consideration tend to display heavy tails. To deal with this issue when it arises, the data are trimmed. For $\{Y_t\}_{t=1}^{T}$, the trimmed data is $Y_t^* = Y_t \times I(\epsilon)$, where $I(\epsilon) \equiv I(\widehat{p}(\epsilon) \leq Y_t \leq \widehat{p}(1-\epsilon))$, $\widehat{p}(\epsilon)$ is the $(100 \times \epsilon)$th empirical percentile, and $I(\cdot)$ is the indicator function.

This trimming rule is certainly rather rudimentary. For instance, it is symmetric. Since the data under consideration need to be skewed, an asymmetric trimming rule may be more appropriate. Moreover, a trimming rule that is asymptotically vanishing may also be preferable so as to preserve the limiting behavior of the estimators (see, e.g., Hill and

[13] That is, $\widehat{\beta}$ depends on both second-order covariances (third moment information) and autocovariances (fourth moment information).

Renault 2010). The trimming rule employed here, however, is simple and so is consistent with the theme of the paper. Moreover, as will be seen, this rule seems sufficient, at least, for illustrating the usefulness of the proposed estimators.

5.1 Intraday EUR/USD Returns

This section reports GARCH estimates for EUR/USD spot and nearby futures returns measured at the hourly and 5-minute frequency for spot and 1-minute frequency for nearby futures. FX returns (specifically, the log difference of rates measured at the given frequency and scaled by 100) are analyzed because of the findings in Hansen and Lunde (2005). EUR/USD spot and nearby futures rates are selected because they tend to be the most liquid FX instruments within their respective asset classes as measured by volume.[14] Volatility in intraday financial returns tends to display a U-shaped periodicity that is at odds with the geometric decay implied by standard GARCH models (see, e.g., Anderson and Bollerslev, 1997). As a consequence, all EUR/USD returns are pre-filtered for this periodicity following the approach described in Section 1 of Hecq, Laurent, and Palm (2012) using the specific filter in equation 4.1 with $p_j = 6$ as applied in Anderson and Bollerslev.

Table 1 reports results for hourly EUR/USD spot returns. CFE2, CFE3, and GLS2 are all highly comparable between each other in terms of point estimates and are each broadly comparable to the QMLE under the same criteria. The latter observation is somewhat surprising given the evidenced (high) level of GARCH persistence and the results from the Monte Carlo studies. For $T = 5,000$ and $T = 20,000$, GLS2 estimates are the closest to the QMLE estimates, while CFE3 estimates tend to be the farthest away.[15] At $T = 65,534$, CFE2 estimates are the closest to the QMLE estimates; although, with the exception of $K = 20$, the CFE2 estimates imply a stationary GARCH process. The GLS2 estimates are now the farthest away. CFE2 requires the most trimming, CFE3 the least. At $T = 20,000$, CFE2 and CFE3 can enjoy upwards of 80% reductions in computational times, depending on K, relative to the QMLE. The same general magnitudes of reductions in computational times are also recorded at $T = 65,534$.

[14] Nearby futures rates are from the futures contract nearest to expiration. Spot rates tend to be more liquid than futures rates.

[15] Distance, here, is measured as the sum of absolute deviations for each of the two parameter estimates.

Table 2 summarizes results for 5-minute EUR/USD spot returns. Across all four estimators, the associated point estimates remain broadly comparable. At $T = 20,000$, the GLS2 estimates are closest to the QMLE estimates, the CFE2 estimates the farthest away. At $T = 40,521$, the CFE3 estimates are the closest to the QMLE estimates, while the CFE2 estimates remain the farthest away. CFE3 continues to require the least trimming, while GLS2 now can require very heavy trimming. At both sample sizes, CFE2 and CFE3 continue to enjoy substantial reductions in computational times relative to the QMLE. These reductions can exceed 90% in the case of CFE3 applied to the largest sample using $K = 10$.

Table 2 also reports results for 1-minute EUR/USD (nearby) futures returns. CFE2 estimates are closest to the QMLE estimates for both sample sizes, while GLS2 estimates are the farthest away. For the CFE2 estimates, no trimming is required at the largest sample size. Even so, however, CFE3 remains the estimator requiring the least amount of trimming, while GLS2 (can) require very heavy trimming. CFE2 and CFE3 enjoy only modest reductions in computational times relative to the QMLE at the smallest sample size. However, at the largest sample size, those reductions, again, appear quite sizable; upwards of 95% when $K = 10$.

Looking across Tables 1–2, the standard errors of CFE3 and GLS2 are fairly comparable. The standard errors of CFE2, on the other hand, are noticeably higher; in particular, for the ARCH effect.[16] The real promise of these estimators seems to be at the (very) high sampling frequencies, where much is gained (computational time wise) from simple-to-implement estimators and not much is lost in terms of point accuracy of these simple estimators relative to the QMLE alternative.

5.2 Equity Returns and Risk Factors

Residuals from Fama and French (1993) three-factor regressions, while still displaying GARCH effects (see Luger, 2013), display different effects than those associated with raw equity returns. Specifically, these residual returns tend to display much higher ARCH effects and much lower GARCH effects, rendering them prime candidates for the estimators

[16] This result, perhaps, should not be too surprising, since the ARCH estimate is simply the ratio of two sample moments.

proposed in this paper (see the Monte Carlo results in the Supplemental Appendix). Moreover, Petkova (2006) tests a version of the ICAPM) where the risk factors are innovations to a VAR that forecasts future investment opportunities. These innovations are also found to display GARCH effects; but, as in the case of certain Fama-French three factor model residuals, GARCH effects characterized by relatively high ARCH parameters and relatively low GARCH parameters. Since these innovations are posited as (risk) pricing factors, estimating GARCH processes for these innovations is one way to empirically estimate and test conditional versions of the ICAPM that include time-varying betas, one which is particularly well-suited to the estimators proposed in this paper, since the theories that support formulations of models like the ICAPM tend to be stated in terms of moments and restrictions on those moments as opposed to distributions.

This section first examines a subset of residuals from applying the Fama-French three factor model to equity portfolios sorted by size and book-to-market (B/M). The equity returns data used in this exercise are measured weekly and cover the period 07/05/1963–12/31/2004.[17] Next, this section examines risk factors in an ICAPM context. The data used for this exercise is also measured weekly and covers the period July 1964 through June 2001.[18] Relative to the other data samples analyzed in this paper, the two samples considered in this section are small.[19]

Table 3 reports results for the Fama-French three factor model residuals. Much higher ARCH effects and much lower GARCH effects are evidenced than those typically encountered when analyzing raw equity returns. For the QMLE estimates, neither the ARCH nor the GARCH estimates are particularly precise (i.e., characterized by small standard errors). In contrast, both the CFE3 and GLS2 estimates are relatively more precise. For instance, estimates from CFE3 and GLS2 are significant across all three residual series, while the QMLE estimates are either only marginally significant or insignificant for two of the three series. Moreover, the ARCH estimate for residual series ϵ_{12} using CFE2, found (generally) to

[17] All equity returns data, including the three Fama-French pricing factors, are downloaded as daily returns from Kenneth French's website and then compounded into weekly returns.

[18] See the notes to Tables 12 and 13 for a description of this data.

[19] In both cases, weekly returns are selected (over daily returns) so as to reduce day-of-the-week and weekend effects as well as the effects of nonsynchronous trading and bid-ask bounce.

deliver relatively imprecise estimates in the FX application, is statistically significant, while its QMLE counterpart is not. In this application, only light trimming is required for CFE2 and GLS2. For residual series ϵ_{12}, no trimming is required for CFE2. For CFE3, some (light) trimming is required.

Finally, Table 4 summarizes results for the ICAPM risk factors. For two of the three factors, QMLE implies IGARCH processes, which challenge the notion of stationary (and finite) factor variances, characteristics that are central to the model being tested. CFE2, CFE3, and GLS2, on the other hand, all imply covariance stationary GARCH processes, some with much less persistence than implied by the QMLE estimates. Applied to these pricing factors, CFE2 now requires the least amount of trimming and produces estimates of the ARCH effect that are significant for both the term and prem factors. CFE3, in contrast to the other applications, now requires the most trimming; which, in the case of the rf factor is very heavy. The CFE2 estimate for the GARCH effect of rf turns out to be negative (although, insignificant). As a consequence, GLS2 applied to the rf series estimates an ARCH(1) model using the CFE2 ARCH effect estimate as its starting value. The ARCH estimate from GLS2 is virtually identical to that obtained using the QMLE to estimate the same ARCH(1) model.

6. Conclusion

The main contribution of this paper is to provide closed-form estimators for the GARCH(1,1) model. Standard $\sqrt{T}$-asymptotics apply to these estimators given moment existence criteria no stronger than those required for comparable moment estimators discussed in the literature. These criteria can even be relaxed somewhat by nature of the fact that identification links to properties of the third as opposed to the fourth moment. These closed-form estimators can be close competitors of the QMLE.

The identification result in this paper can be extended to a GARCH(1,1) model with a leverage effect (see, e.g., Glosten, Jagannathan, and Runkle, 1993). Suppose that $h_t = \omega_0 + \left(\alpha_0 + \alpha_0^- \times 1\left(Y_{t-1} < 0\right)\right) Y_{t-1}^2 + \beta_0 h_{t-1}$. Then (6) can be divided into the set of moment conditions $E\left[\widetilde{Y}_t^2 Y_{t-1}\right] = \left(\alpha_0 + \alpha_0^- \times P\left(Y_t < 0\right)\right) E\left[Y_t^3\right]$, and $E\left[\widetilde{Y}_t^2 Y_{t-1} \times \left(1 - 1\left(Y_{t-1} < 0\right)\right)\right] =$

$\alpha_0 (1 - P(Y_t < 0)) E[Y_t^3]$, which can be used to identify a semi-parametric GMM estimator of the GARCH(1,1) model with a leverage effect. Such an estimator would be applicable to stock returns given the results of Hansen and Lunde (2005).

Finally, the closed-form estimators I propose also apply to multivariate GARCH models. These estimators grant a closed-form solution to Bollerslev's (1990) CCC model and apply to Engle's (2002) DCC model. To the extent that second-order covariances permit a closed-form solution for the parameters governing conditional covariances between (standardized) residuals in the DCC model, a completely closed-form solution would also be possible for this more general multivariate GARCH model.

References

[1] Andersen, T.G. & T. Bollerslev (1997) Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance* 4, 115-158.

[2] Anderson, T.W. & A.M. Walker (1964) On the asymptotic distribution of the autocorrelations of a sample from a linear stochastic process. *The Annals of Mathematical Statistics* 35, 1296-1303.

[3] Andrews, D.W.K. (1988) Laws of large numbers for dependent non-identically distributed random variables. *Econometric Theory* 4, 458-467.

[4] Angrist, J., G. Imbens & A. Kreuger (1999) Jackknife instrumental variables estimation. *Journal of Applied Econometrics* 14, 57-67.

[5] Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307–327.

[6] Bollerslev, T. (1990) Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *Review of Economics and Statistics* 72, 498-505.

[7] Bollerslev, T. & J.M. Wooldridge (1992) Quasi-maximum likelihood estimation and inference in dynamic models with time varying covariances. *Econometric Reviews* 11, 143-172.

[8] Brown, B.M. (1971) Martingale central limit theorems. *The Annals of Mathematical Statistics* 42, 59-66.

[9] Carrasco, M. & X. Chen (2002) Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory* 18, 17-39.

[10] Chamberlain, G. (1982) Multivariate regression models for panel data. *Journal of Econometrics* 18, 5-46.

[11] Cragg, J.G. (1983) More efficient estimation in the presence of heteroskedasticity of unknown form. *Econometrica* 51, 751-764.

[12] Donald, S.G. & W.K. Newey (2000), A jackknife interpretation of the continuous updating estimator. *Economic Letters* 67, 239-243.

[13] Drost, F.C. & T.E. Nijman (1993) Temporal aggregation of GARCH processes. *Econometrica* 61, 909-927.

[14] Engle, R.F. (2002) Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics* 20, 339-350.

[15] Engle, R.F. & J. Mezrich (1996) GARCH for groups. *Risk* 9, 36-40.

[16] Escanciano, J.C. (2009) Quasi-maximum likelihood estimation of semi-strong GARCH models. *Econometric Theory* 25, 561-570.

[17] Fama, E.F. & K. French (1993) Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3-56.

[18] Francq, C., L. Horath & J.M. Zakoïan (2011) Merits and drawbacks of variance targeting in GARCH models. *Journal of Financial Econometrics* 9, 619-656.

[19] Francq, C., & J.M. Zakoïan (2000) Estimating weak GARCH representations. *Econometric Theory* 16, 692-728.

[20] Giraitis, L. & P.M. Robinson (2001) Whittle estimation of ARCH models. *Econometric Theory* 17, 608-631.

[21] Glosten, L.R., R. Jagannathan & D.E. Runkle (1993), On the relation between expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* 48, 1779-1801.

[22] Guo, B. & P.C.B Phillips (2001) Efficient estimation of second moment parameters in ARCH models. Unpublished manuscript.

[23] Hannan, E.J., & C.C. Heyde (1972) On limit theorems for quadratic functions of discrete time series. *The Annals of Mathematical Statistics* 43, 2058-2066.

[24] Hansen, L.P. (1982) Large sample properties of generalized method of moments estimators. *Econometrica* 50, 1029-1054.

[25] Hansen, L.P., J. Heaton & A. Yaron (1996) Finite-sample properties of some alternative GMM estimators. *Journal of Business and Economic Statistics* 14, 262-280.

[26] Hansen, P.R. & A. Lunde (2005) A forecast comparison of volatility models: does anything beat a GARCH(1,1)? *Journal of Applied Econometrics* 20, 873-889.

[27] Hecq, A., S. Laurent & F.C. Palm (2012) Common intraday periodicity. *Journal of Financial Econometrics* 10, 325-353.

[28] Heston, S.L. & S. Nandi (2000), A closed-form GARCH option valuation model. *Review of Financial Studies* 13, 585-625.

[29] Hill, J.B. & E. Renault (2010) Generalized method of moments with tail trimming. unpublished manuscript.

[30] Kristensen, D. & O. Linton (2006) A closed-form estimator for the GARCH(1,1)-model. *Econometric Theory* 22, 323-327.

[31] Lee, S.W & B.E. Hansen (1994) Asymptotic theory for the GARCH(1,1) qausi-maximum likelihood estimator. *Econometric Theory*, 10, 29-52.

[32] Luger, R. (2013) Testing for GARCH effects with quasi-likelihood ratios. Forthcoming in *Journal of Risk*.

[33] Lumsdaine, R.L. (1995) Finite-sample properties of the maximum likelihood estimator in GARCH(1,1) and IGARCH(1,1) models: a monte carlo investigation. *Journal of Business and Economic Statistics* 13, 1-10.

[34] 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.

[35] Mandelbrot, B. (1963) The variation of certain speculative prices. *Journal of Business* 36, 394-419.

[36] Merton, R.C. (1973) An intertemporal capital asset pricing model. *Econometrica* 41, 867-887.

[37] Moskowitz, J., Y.H. Ooi, & L.H. Pedersen (2012), Time series momentum, *Journal of Financial Economics* 104, 228-250.

[38] Newey, W.K. & D. McFadden (1994) Large sample estimation and hypothesis testing. in R.F. Engle and D. McFadded (eds), *Handbook of Econometrics* Vol. 4: Amsterdam North Holland, chapter 36, 2111-2245.

[39] Newey, W.K. & D.G. Steigerwald (1997) Asymptotic bias for quasi-maximum-likelihood estimators in conditional heteroskedasticity models. *Econometrica* 65, 587-599.

[40] Newey, W.K & F. Windmeijer (2009) Generalized method of moments with many weak moment conditions. *Econometrica* 77, 687-719.

[41] Petkova, R. (2006) Do the Fama-French factors proxy for innovations in predictive variables? *Journal of Finance* 61, 581-612.

[42] Prono, T. (2013) Market proxies as factors in linear asset pricing models: still living with the Roll critique. Unpublished manuscript.

[43] Prono, T. (2015) Revisiting ARCH: Simple Alternatives to QMLE. Unpublished manuscript.

[44] Robinson, P.M. (1988) The Stochastic Difference Between Econometric Statistics. *Econometrica* 56, 531-548.

[45] Shao, Y. & M. Zhou (2010) A characterization of multivariate normality through univariate projections. *Journal of Multivariate Analysis* 101, 2637-2640.

[46] Skoglund, J. (2001) A simple efficient GMM estimator of GARCH models. unpublished manuscript.

[47] Weiss, A.A. (1986) Asymptotic theory for ARCH models: estimation and testing. *Econometric Theory* 2, 107-131.

[48] White, H. (1982) Instrumental variables regression with independent observations. *Econometrica* 50, 483-499.

[49] White, H. (1984) *Asymptotic Theory for Econometricians*. Orlando, Fla: Academic Press.

[50] Wooldridge JM. (1990) A unified approach to robust, regression-based specification tests. *Econometric Theory* 6, 17-43.

TABLE 1

				CFE2 $K =$				CFE3 $K =$				GLS2 $K =$				QMLE
Freq.	T	Skew.	Parameter	10	20	30	40	10	20	30	40	10	20	30	40	
hourly	5,000	0.10	α	0.025				0.052	0.049	0.047	0.049	0.068	0.068	0.068	0.068	0.073
		(0.03)		(0.616)				(0.051)	(0.046)	(0.042)	(0.046)	(0.018)	(0.018)	(0.018)	(0.018)	(0.032)
			β	0.491	0.659	0.643	0.524	0.490	0.743	0.826	0.760	0.557	0.557	0.557	0.557	0.504
				(0.628)	(0.478)	(0.490)	(0.587)	(0.061)	(0.052)	(0.045)	(0.048)	(0.078)	(0.078)	(0.078)	(0.078)	(0.328)
			$\alpha+\beta$	0.516	0.684	0.668	0.549	0.542	0.793	0.873	0.809	0.625	0.625	0.625	0.625	0.576
	20,000	0.18	α	0.088				0.111	0.094	0.083	0.086	0.044	0.044	0.044	0.044	0.014
		(0.02)		(0.351)				(0.049)	(0.039)	(0.033)	(0.035)	(0.018)	(0.018)	(0.019)	(0.019)	(0.004)
			β	0.928	0.905	0.908	0.897	0.813	0.863	0.886	0.880	0.946	0.946	0.946	0.946	0.984
				(0.121)	(0.104)	(0.092)	(0.094)	(0.052)	(0.039)	(0.034)	(0.035)	(0.026)	(0.025)	(0.026)	(0.026)	(0.005)
			$\alpha+\beta$	1.015	0.992	0.995	0.984	0.924	0.956	0.969	0.966	0.990	0.990	0.990	0.990	0.999
	65,534	0.11	α	0.032				0.097	0.081	0.070	0.070		0.091	0.091	0.091	0.007
		(0.01)		(0.161)				(0.033)	(0.026)	(0.021)	(0.021)		(0.038)	(0.028)	(0.032)	(0.001)
			β	0.937	0.972	0.965	0.960	0.844	0.885	0.908	0.908		0.899	0.899	0.899	0.992
				(0.054)	(0.033)	(0.029)	(0.028)	(0.037)	(0.026)	(0.022)	(0.021)		(0.047)	(0.034)	(0.039)	(0.001)
			$\alpha+\beta$	0.970	1.004	0.997	0.992	0.941	0.966	0.978	0.978		0.990	0.990	0.990	1.000

TABLE 2

Freq.	T	Skew.	Parameter	CFE2 $K=$ 10	CFE2 20	CFE2 30	CFE2 40	CFE3 $K=$ 10	CFE3 20	CFE3 30	CFE3 40	GLS2 $K=$ 10	GLS2 20	GLS2 30	GLS2 40	QMLE
5-min	20,000	-0.21	α	0.025				0.082	0.085	0.106	0.107	0.063	0.063	0.063	0.063	0.074
		(0.02)		(0.502)				(0.031)	(0.027)	(0.035)	(0.035)	(0.012)	(0.012)	(0.012)	(0.012)	(0.010)
			β	0.788	0.757	0.746	0.766	0.894	0.888	0.843	0.841	0.869	0.870	0.865	0.872	0.879
				(0.265)	(0.281)	(0.287)	(0.273)	(0.035)	(0.028)	(0.035)	(0.035)	(0.093)	(0.093)	(0.094)	(0.093)	(0.018)
			$\alpha+\beta$	0.813	0.782	0.771	0.791	0.976	0.973	0.949	0.948	0.931	0.933	0.928	0.935	0.954
	40,521	0.11	α	0.149				0.150	0.144	0.143	0.140	0.169	0.169	0.058	0.061	0.091
		(0.01)		(0.287)				(0.070)	(0.067)	(0.066)	(0.064)	(0.060)	(0.052)	(0.006)	(0.005)	(0.011)
			β	0.769	0.778	0.711	0.723	0.761	0.777	0.781	0.789	0.711	0.717	0.804	0.890	0.865
				(0.139)	(0.128)	(0.152)	(0.147)	(0.072)	(0.067)	(0.066)	(0.064)	(0.140)	(0.116)	(0.027)	(0.014)	(0.018)
			$\alpha+\beta$	0.918	0.927	0.860	0.872	0.910	0.921	0.923	0.929	0.880	0.886	0.862	0.951	0.955
1-min	20,000	0.67	α	0.258				0.117	0.121	0.112	0.113	0.042	0.042	0.042	0.352	0.251
		(0.02)		(0.411)				(0.079)	(0.070)	(0.064)	(0.064)	(0.011)	(0.021)	(0.018)	(0.046)	(0.093)
			β	0.711	0.700	0.681	0.624	0.851	0.843	0.859	0.857	0.917	0.891	0.896	0.547	0.730
				(0.228)	(0.218)	(0.221)	(0.243)	(0.082)	(0.070)	(0.064)	(0.064)	(0.027)	(0.093)	(0.070)	(0.057)	(0.060)
			$\alpha+\beta$	0.969	0.959	0.939	0.882	0.968	0.965	0.972	0.971	0.959	0.933	0.938	0.898	0.981
	180,000	1.30	α	0.177				0.123	0.115	0.115	0.116	0.358	0.355	0.352	0.351	0.205
		(0.01)		(0.377)				(0.059)	(0.054)	(0.054)	(0.054)	(0.012)	(0.012)	(0.012)	(0.012)	(0.065)
			β	0.772	0.693	0.695	0.701	0.775	0.803	0.804	0.802	0.484	0.478	0.472	0.472	0.720
				(0.176)	(0.170)	(0.155)	(0.152)	(0.060)	(0.054)	(0.054)	(0.054)	(0.023)	(0.024)	(0.025)	(0.025)	(0.079)
			$\alpha+\beta$	0.950	0.871	0.873	0.878	0.898	0.919	0.920	0.918	0.842	0.832	0.824	0.824	0.925

TABLE 3

T (weekly)	Residual Return	Skew.	Para.	CFE2 $K=$ 10	CFE2 $K=$ 20	CFE2 $K=$ 30	CFE2 $K=$ 40	CFE3 $K=$ 10	CFE3 $K=$ 20	CFE3 $K=$ 30	CFE3 $K=$ 40	GLS2 $K=$ 10	GLS2 $K=$ 20	GLS2 $K=$ 30	GLS2 $K=$ 40	QMLE
2,166	ϵ_{11}	5.51	α	0.326				0.118	0.110	0.094	0.097	0.095	0.095	0.095	0.095	0.140
		(0.05)		(0.280)				(0.039)	(0.036)	(0.030)	(0.031)	(0.028)	(0.028)	(0.028)	(0.028)	(0.038)
			β	0.433	0.551	0.646	0.584	0.775	0.806	0.859	0.849	0.895	0.895	0.895	0.895	0.668
				(0.208)	(0.177)	(0.143)	(0.153)	(0.039)	(0.033)	(0.025)	(0.025)	(0.033)	(0.033)	(0.033)	(0.033)	(0.072)
			$\alpha+\beta$	0.759	0.877	0.971	0.910	0.893	0.917	0.952	0.946	0.990	0.990	0.990	0.990	0.807
	ϵ_{12}	6.32	α	0.124				0.027	0.027	0.027	0.027	0.374	0.032	0.032	0.032	0.347
		(0.05)		(0.022)				(0.046)	(0.045)	(0.046)	(0.046)	(0.220)	(0.007)	(0.007)	(0.007)	(0.231)
			β	0.103	0.233	0.226	0.215	0.317	0.489	0.370	0.375	0.361	0.958	0.958	0.958	0.598
				(0.050)	(0.037)	(0.031)	(0.030)	(0.049)	(0.051)	(0.048)	(0.047)	(0.710)	(0.009)	(0.009)	(0.009)	(0.215)
			$\alpha+\beta$	0.227	0.357	0.350	0.339	0.344	0.516	0.397	0.402	0.735	0.990	0.990	0.990	0.944
	ϵ_{14}	6.97	α	0.624				0.015	0.015	0.015	0.015	0.643	0.640	0.635	0.632	0.607
		(0.05)		(0.361)				(0.032)	(0.032)	(0.032)	(0.032)	(0.189)	(0.191)	(0.196)	(0.196)	(0.576)
			β	0.323	0.310	0.278	0.248	0.267	0.279	0.269	0.291	0.218	0.218	0.218	0.218	0.332
				(0.327)	(0.315)	(0.298)	(0.276)	(0.044)	(0.053)	(0.055)	(0.044)	(0.075)	(0.082)	(0.087)	(0.086)	(0.507)
			$\alpha+\beta$	0.947	0.935	0.902	0.872	0.282	0.294	0.284	0.306	0.861	0.858	0.852	0.849	0.939

TABLE 4

T (weekly)	Pricing Factor	Skew.	Para.	CFE2 $K=$ 10	CFE2 20	CFE2 30	CFE2 40	CFE3 $K=$ 10	CFE3 20	CFE3 30	CFE3 40	GLS2 $K=$ 10	GLS2 20	GLS2 30	GLS2 40	QMLE
1,929	term	-0.14	α	0.622				0.232	0.202	0.201	0.202	0.363	0.355	0.363	0.363	0.526
		(0.06)		(0.272)				(0.058)	(0.053)	(0.052)	(0.053)	(0.040)	(0.037)	(0.040)	(0.040)	(0.084)
			β	0.287	0.370	0.310	0.319	0.440	0.250	0.265	0.235	0.248	0.235	0.248	0.248	0.515
				(0.215)	(0.176)	(0.161)	(0.136)	(0.029)	(0.053)	(0.052)	(0.053)	(0.188)	(0.183)	(0.190)	(0.189)	(0.068)
			$\alpha+\beta$	0.909	0.991	0.932	0.941	0.672	0.451	0.466	0.438	0.611	0.590	0.611	0.611	1.041
	prem	0.39	α	0.580				0.061	0.062	0.059	0.052	0.395	0.392	0.396	0.393	0.408
		(0.06)		(0.295)				(0.030)	(0.020)	(0.018)	(0.016)	(0.051)	(0.050)	(0.051)	(0.047)	(0.101)
			β	0.272	0.369	0.360	0.371	0.917	0.915	0.921	0.934	0.597	0.600	0.597	0.599	0.665
				(0.200)	(0.136)	(0.104)	(0.085)	(0.041)	(0.023)	(0.019)	(0.016)	(0.063)	(0.061)	(0.063)	(0.058)	(0.073)
			$\alpha+\beta$	0.852	0.948	0.940	0.951	0.978	0.977	0.980	0.986	0.993	0.992	0.993	0.992	1.073
	rf	-0.57	α	0.406				0.170	0.126	0.150	0.146	0.293				0.127
		(0.06)		(0.561)				(0.033)	(0.025)	(0.026)	(0.025)	(0.033)				(0.040)
			β	-0.186	-0.174	-0.140	-0.076	0.778	0.851	0.812	0.819					0.839
				(0.188)	(0.151)	(0.115)	(0.093)	(0.033)	(0.024)	(0.026)	(0.025)					(0.061)
			$\alpha+\beta$	0.219	0.231	0.265	0.329	0.947	0.976	0.962	0.965	0.293				0.965

Notes to Tables 1 and 2: GARCH estimates are reported for EUR/USD spot and nearby futures log returns measured at hourly, 5-minute, and 1-minute frequencies. Hourly EUR/USD spot rates source to fxhistoricaldata.com. 5-minute EUR/USD spot rates, and minute EUR/USD nearby futures rates source to Bloomberg. From these rates, log returns are calculated and scaled by 100. GLS2 estimates use CFE2 estimates as starting values and set $l = 10$. QMLE estimates are based on untrimmed data and serve as benchmarks. At the hourly frequency, data feeding (i) CFE2 are trimmed at $\epsilon = 0.05$ for all sample sizes, (ii) CFE3 are untrimmed for all sample sizes, and (ii) GLS2 are untrimmed for $T = 5,000$, are trimmed at $\epsilon = 0.05$ for $T = 20,000$ and at $\epsilon = 0.01$ for $T = 65,534$. At $T = 65,534$ with $K = 10$, GLS2 did not converge at any of the chosen trimming levels $\epsilon = 0.01, 0.05, 0.10, 0.25$. At the 5-minute frequency, data feeding (i) CFE2 are trimmed at $\epsilon = 0.05$ for $T = 20,000$ and at $\epsilon = 0.01$ for $T = 40,521$, (ii) CFE3 are untrimmed for both sample sizes, (iii) GLS2 is trimmed at $\epsilon = 0.25$ for $T = 20,000$, is untrimmed for $T = 40,521$ using CFE2-generated starting values based on $K = 10, 20$, and is trimmed at $\epsilon = 0.25$ using CFE2-generated starting values based on $K = 30$ and at $\epsilon = 0.01$ using CFE2-generated starting values based on $K = 40$. At the 1-minute frequency, data feeding (i) CFE2 are trimmed at $\epsilon = 0.05$ for $T = 20,000$ and are untrimmed for $T = 180,000$, (ii) CFE3 are untrimmed for both sample sizes, (iii) GLS2 are trimmed at $\epsilon = 0.25$ using CFE2-generated starting values based on $K = 10, 20, 30$ and $\epsilon = 0.01$ using CFE2-generated starting values based on $K = 40$ at $T = 20,000$ and are trimmed at $\epsilon = 0.01$ for $T = 180,000$. Skew. is the skewness of the given return series. Standard errors are given in parentheses. The CFE2 and CFE3 standard error (SE) estimates follow from Theorem 3 and Corollary 2, respectively. The GLS2 and QMLE standard error estimates follow from Theorem 4.

Notes to Tables 3 and 4: GARCH estimates are reported for (i) regression residuals from the Fama and French (1993) three factor model applied to certain equity portfolios constructed from 5×5 sorts on size and book-to-market (B/M), and (ii) risk factors characterizing the investment opportunity set of Merton's (1973) ICAPM. The data supporting (i), including both the aforementioned 5×5 portfolio sorts and the three Fama-French pricing factors, are obtained as daily returns from Kenneth French's website and then compounded into weekly returns. The period covered by these weekly returns is 07/05/1963–12/31/2004. ϵ_{ij} references the residual from regressing the three Fama-French factors on the portfolio return in the ijth position of the 5×5 size-B/M sort. The data supporting (ii) are innovations to a VAR described in Prono(2013) and motivated by Petkova (2006). The individual elements of this VAR are also measured weekly over the period July 1964 through June 2001. The variables term, prem, and rf reference innovations (from the aforementioned VAR) to the yield spread between the 10-year and 1-year Treasury Bond, the yield spread between BAA- and AAA-rated bonds, and the 1-month Treasury Bill yield, respectively. GLS2 estimates use CFE2 estimates as starting values and set $l = 10$. QMLE estimates are based on untrimmed data and serve as benchmarks. For the Fama-French three factor model residuals, data feeding (i) CFE2 are trimmed at $\epsilon = 0.01$ for ϵ_{11} and ϵ_{14} and are untrimmed for ϵ_{12}, (ii) CFE3 are trimmed at $\epsilon = 0.01$ for both ϵ_{11} and ϵ_{12} and are untrimmed for ϵ_{14}, (iii) GLS2 are trimmed at $\epsilon = 0.01$ for ϵ_{11} and ϵ_{12} at $K = 20, 30, 40$ and are untrimmed for ϵ_{12} at $K = 10$ and ϵ_{14}. For the ICAPM risk factors, data feeding (i) CFE2 are trimmed at $\epsilon = 0.01$ for term and are untrimmed for both prem and rf, (ii) CFE3 are trimmed at $\epsilon = 0.05$ for term at $K = 10$ and $\epsilon = 0.10$ at $K = 20, 30, 40$, at $\epsilon = 0.10$ for prem, and at $\epsilon = 0.25$ for rf, (iii) GLS2 are trimmed at $\epsilon = 0.05$ for both term and prem and are untrimmed for rf. The CFE2 estimate for the GARCH effect of rf turns out to be negative (although, insignificant). As a consequence, GLS2 applied to the rf series estimates an ARCH(1) model using the CFE2 ARCH effect estimate as its starting value. The ARCH estimate from GLS2 is virtually identical to that obtained using the QMLE to estimate the same ARCH(1) model. Skew. is the skewness of the given return series. Standard errors are given in parentheses. The CFE2 and CFE3 standard error (SE) estimates follow from Theorem 3 and Corollary 2, respectively. The GLS2 and QMLE standard error estimates follow from Theorem 4.