Economics Letters 157 (2017) 129–132

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

The time-varying GARCH-in-mean model

Gustavo Fruet Dias [1]

CREATES and Aarhus University, Fuglesangs Allé 4, 8210 Aarhus V, Denmark

HIGHLIGHTS

- Estimation of the stochastic time-varying risk premium parameter within the TVGARCH-in-mean models.
- The proposed kernel-based iterative estimator attains good finite sample performance.
- The risk premium parameter is found to be time-varying and highly persistent.

ARTICLE INFO

Article history:
Received 25 April 2017
Received in revised form 2 June 2017
Accepted 4 June 2017
Available online 12 June 2017

JEL classification:
C13
C15
C22
G12

Keywords:
Risk-return tradeoff
Time-varying coefficients
Iterative estimators
GARCH-type models

ABSTRACT

I propose an estimation strategy for the stochastic time-varying risk premium parameter in the context of a time-varying GARCH-in-mean (TVGARCH-in-mean) model. A Monte Carlo study shows that the proposed algorithm has good finite sample properties. Using monthly excess returns on the CRSP index, I document that the risk premium parameter is indeed time-varying and shows high degree of persistence.

1. Introduction

Asset pricing theories suggest that riskier assets should demand higher expected returns. Using Merton's (1973) theoretical framework, the conditional expectation of the market excess returns reads

$$\mathbb{E}\left(r_{t+1}^{m} \mid \mathcal{F}_t\right) - r_t^{f} = \lambda_t Var\left(r_{t+1}^{m} \mid \mathcal{F}_t\right), \tag{1}$$

where r_{t+1}^{m} and r_t^{f} are the returns on the market portfolio and risk-free asset, $\mathcal{F}_t$ is the market-wide information available at time t, and λ_t is the coefficient of relative risk aversion defined as the elasticity of marginal value with respect to wealth. Most studies assume the risk-return trade-off is constant over time and linear in the variance, which is usually associated with the reasons behind mixed empirical evidences when estimating the risk-return trade-off (Linton and Perron (2003), Brandt and Wang, Christensen et al. (2012), among others). To address this issue, I adopt the time-varying GARCH-in-mean (TVGARCH-in-mean) model in the spirit of Anyfantaki and Demos (2016) which allows λ_t to be a time-varying stochastic process and put forward a feasible estimation strategy for λ_t (see references in Anyfantaki and Demos (2016) for variants of the TVGARCH-in-mean models). Specifically, I combine Giraitis et al.'s (2013) time-varying kernel least squares estimator with Linton and Perron's (2003) semiparametric iterative approach to estimate the time-varying risk premium coefficient. A Monte Carlo study shows that the proposed algorithm has good finite sample properties. Using the excess returns of the Center for Research on Security Prices (CRSP) index, I document that the risk premium parameter is indeed time-varying, alternating positive (statistically significant) and nonsignificant values over time.

E-mail address: gdias@econ.au.dk.

[1] I thank George Kapetanios, Cristina Scherrer and an anonymous referee for constructive suggestions. I acknowledge support from CREATES — Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation.

http://dx.doi.org/10.1016/j.econlet.2017.06.005

2. The time-varying GARCH-in-mean

The generic TVGARCH-in-mean(p, q) is defined as:

$$r_t = \lambda_t \sigma_t + \epsilon_t, \tag{2}$$

$$\epsilon_t = \sigma_t \eta_t, \tag{3}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2, \tag{4}$$

$$\epsilon_t^2 = \psi_0 + u_t + \sum_{i=1}^{\infty} \psi_i u_{t-i}, \tag{5}$$

where η_t is an independent and identically distributed (*iid*) zero mean process with unit variance; σ_t is a latent conditional standard deviation; (5) is the $MA(\infty)$ representation of the conditional variance equation; $u_t = \epsilon_t^2 - \sigma_t^2$ is a martingale difference sequence process; $\phi = \left(\omega, \alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q\right)'$ collects the free parameters in (4); and $\psi_i := \varrho_i(\phi)$ $i = 1, 2, \ldots$ are deterministic functions of the elements in ϕ. Similarly as in Giraitis et al. (2013), the time-varying risk premium parameters are assumed to evolve smoothly over time, so that it satisfies a local stability condition in the form of $\sup_{s:\|s-t\| \le h} \|\lambda_t + \lambda_s\|_2 = O_p(h/t)$.

Estimating the free parameters in (2) and (4) by maximum-likelihood is not a feasible alternative, as the class of TVGARCH-in-mean(p, q) models involves two unobserved processes: λ_t and ϵ_t. Anyfantaki and Demos (2016) address this issue in the context of the time-varying EGARCH(1,1)-in-mean model. Specifically, their work differs from mine in two ways. First, they parameterize the conditional variance as an EGARCH(1,1) model and, most importantly, λ_t as a stationary AR(1) process. By contrast, λ_t in (2) is assumed to satisfy $\sup_{s:\|s-t\| \le h} \|\lambda_t + \lambda_s\|_2 = O_p(h/t)$, which encompasses the case of the driftless random walk process considered in Chou et al. (1992). Second, while I propose a kernel-based nonparametric method to estimate the time-varying risk premium parameter, Anyfantaki and Demos's (2016) estimation strategy is based on Bayesian methods (Markov chain Monte Carlo (MCMC) likelihood based estimation procedure).

I combine Linton and Perron's (2003) iterative semiparametric estimator with Giraitis et al.'s (2013) kernel-based least squares framework to estimate the free parameters $\theta = \left(\underline{\lambda}, \phi\right)'$, where $\underline{\lambda} = (\lambda_1, \ldots, \lambda_T)'$. This method consists of recursively updating estimates of σ_t and u_t on each iteration, and then computing estimates of $\underline{\lambda}$ and ϕ. To this end, consider moment conditions based on (2) and (5),

$$\mathbb{E}\left[\sigma_t \left(r_t - \lambda_t \sigma_t\right)\right] = 0, \quad \text{for each } t = 1, 2, \ldots, T, \tag{6}$$

$$\mathbb{E}\left[z_t u_t\right] = 0, \quad \text{with } z_t := \frac{\partial \left(\psi_0 + \sum_{i=1}^{\bar{q}} \psi_i u_{t-i}\right)}{\partial \phi}, \tag{7}$$

where (7) is truncated at some lag-order $\bar{q}$ with $\bar{q} > p + q + 1$. Notably, (7) holds because u_t is a martingale difference sequence and z_t is a function of lagged values of u_t. It follows that estimating θ by the standard generalized method of moments (GMM) using the moments defined in (6) and (7) is not operational, as z_t and σ_t are latent variables. Using Linton and Perron's (2003) approach, rewrite (6) and (7) using estimates of σ_t and u_t obtained at some j iteration,

$$\mathbb{E}\left[\sigma_{j,t} \left(r_t - \lambda_{j+1,t} \sigma_{j,t}\right)\right] = 0, \quad \text{for each } t = 1, 2, \ldots, T, \tag{8}$$

$$\mathbb{E}\left[z_{j,t} u_{j+1,t}\right] = 0, \tag{9}$$

where $\sigma_{j,t}$ and $z_{j,t}$ denote the filtered estimates of σ_t and z_t based on $\widehat{\theta}_j$, and $u_{j+1,t} = \epsilon_{j,t}^2 - \psi_{j+1,0} - \sum_{i=1}^{\bar{q}} \psi_{j+1,i} u_{j,t-i}$ with $\epsilon_{j,t}^2 = \left(r_t - \lambda_{j+1,t} \sigma_{j,t}\right)^2$. While the finite sample counterpart of (9) is given by the usual sample mean, computing the sample counterpart of (8) is less obvious. The work of Giraitis et al. (2013) suggests the use of local kernels to construct operational sample counterparts of (8). In turn, a feasible moment condition based on (8) reads

$$K_t^{-1} \sum_{\tau=1}^{T} k_{t,\tau} \sigma_{j,\tau} \left(r_\tau - \widehat{\lambda}_{j+1,t} \sigma_{j,\tau}\right) = 0, \quad \text{for each } t = 1, 2, \ldots, T, \tag{10}$$

where $k_{t,\tau} = K\left((t - \tau)/H\right)$ denotes a kernel function such that $K(x) \ge 0$ for any $x \in \mathbb{R}$ is a continuous bounded function with a bounded first derivative and $\int K(x)dx = 1$; H is the bandwidth parameter satisfying $H = o\left(T/\ln(T)\right)$ as $H \to \infty$; and $K_t = \sum_{\tau=1}^{T} k_{t,\tau}$. Notably, writing the moment conditions as in (10) is consistent with previous studies in the time-varying parameter literature which maximizes kernel weighted log-likelihood functions (see Robinson (1989), Giraitis et al. (2016), among others).

I use the fact that (10) is exactly identified for each t, and hence estimates of λ_t can be obtained independently of ϕ. In turn, estimates of θ are computed iteratively by a two-step procedure. The first step consists of solving (10) for each t, while the second step mimics the work of Linton and Perron (2003) and consists of estimating ϕ using the sample counterpart of (9). In practice, the kernel-based iterative estimator is as follows:

Step 1: Choose starting values $\widehat{\underline{\lambda}}_0$ and $\widehat{\phi}_0$, such that $\widehat{\phi}_0$ satisfies the second-order stationarity conditions of the GARCH(1,1) model. Using $\widehat{\theta}_{0,t} = \left(\widehat{\underline{\lambda}}_0, \widehat{\phi}_0\right)'$, compute recursively $\left\{\sigma_{0,t}^2\right\}_{t=1}^T$, and $\left\{u_{0,t}\right\}_{t=1}^T$ from (2)–(5).

Step 2: Given $\left\{\sigma_{0,t}^2\right\}_{t=1}^T$, calculate

$$\widehat{\lambda}_{1,t} = \left(\sum_{\tau=1}^{T} k_{t,\tau} \sigma_{0,\tau}^2\right)^{-1} \sum_{\tau=1}^{T} k_{t,\tau} \sigma_{0,\tau} r_\tau, \quad \text{for each } t = 1, 2, \ldots, T. \tag{11}$$

Step 3: Solving the sample counterpart of (9) is equivalent to estimate $\widehat{\phi}_1$ by nonlinear least squares. Calculate

$$\widehat{\phi}_1 = \underset{\widehat{\phi}_1}{\text{argmin}} \times \sum_{t=1}^{T} \left\{\left(r_t - \widehat{\lambda}_{1,t} \sigma_{0,t}\right)^2 - \widehat{\psi}_{1,0} - \sum_{i=0}^{\bar{q}} \widehat{\psi}_{1,i} u_{0,t-1-i}\right\}^2. \tag{12}$$

Step 4: Update recursively $\left\{\sigma_{1,t}^2\right\}_{t=1}^T$ and $\left\{u_{1,t}\right\}_{t=1}^T$ based on $\widehat{\theta}_1$.

Repeat steps 2–4 j times until $\widehat{\theta}_j$ converges. Convergence occurs when $\left\|\widehat{\underline{\lambda}}_j - \widehat{\underline{\lambda}}_{j-1}\right\|_2 \le \varepsilon$ and $\left\|\widehat{\phi}_j - \widehat{\phi}_{j-1}\right\|_2 \le \varepsilon$, with ε set to 10^{-5}. Parameters on the jth iteration are given by:

$$\widehat{\lambda}_{j,t} = \left[\sum_{\tau=1}^{T} k_{t,\tau} \sigma_{j-1,\tau}^2\right]^{-1} \sum_{\tau=1}^{T} k_{t,\tau} \sigma_{j-1,\tau} r_\tau, \quad \text{for each } t = 1, 2, \ldots, T, \tag{13}$$

$$\widehat{\phi}_j = \underset{\widehat{\phi}_j}{\text{argmin}} \times \sum_{t=1}^{T} \left[\left[r_t - \widehat{\lambda}_{j,t} \sigma_{j-1,t}\right]^2 - \widehat{\psi}_{j,0} - \sum_{i=0}^{\bar{q}} \widehat{\psi}_{j,i} u_{j-1,t-1-i}\right]^2. \tag{14}$$

Finally, three inputs are still necessary to implement the above algorithm: the kernel function, the bandwidth parameter H, and the truncation lag $\bar{q}$. As in Giraitis et al. (2013), three kernel functions are used: the Epanechnikov, Gaussian, and flat kernels.