

TI 2015-027/III
Tinbergen Institute Discussion Paper



In-Sample Bounds for Time-Varying Parameters of Observation Driven Models

Francisco Blasques
Siem Jan Koopman
Katarzyna Lasak
André Lucas

*Faculty of Economics and Business Administration, VU University Amsterdam, and Tinbergen
Institute, the Netherlands.*

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: <http://www.dsf.nl/>

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

IN-SAMPLE BOUNDS FOR TIME-VARYING PARAMETERS OF OBSERVATION DRIVEN MODELS

Francisco Blasques^a Siem Jan Koopman^{a,b}
Katarzyna Łasak^a and André Lucas^a

^(a) VU University Amsterdam and Tinbergen Institute

^(b) CREATES, Aarhus University

February 21, 2015

Abstract

We study the performance of two analytical methods and one simulation method for computing in-sample confidence bounds for time-varying parameters. These in-sample bounds are designed to reflect parameter uncertainty in the associated filter. They are applicable to the complete class of observation driven models and are valid for a wide range of estimation procedures. A Monte Carlo study is conducted for time-varying parameter models such as generalized autoregressive conditional heteroskedasticity and autoregressive conditional duration models. Our results show clear differences between the actual coverage provided by our three methods of computing in-sample bounds. The analytical methods may be less reliable than the simulation method, their coverage performance is sufficiently adequate to provide a reasonable impression of the parameter uncertainty that is embedded in the time-varying parameter path. We illustrate our findings in a volatility analysis for monthly Standard & Poor's 500 index returns.

Keywords: autoregressive conditional duration; delta-method; generalized autoregressive conditional heteroskedasticity; score driven models; time-varying mean.

Acknowledgments: Blasques and Lucas thank the Dutch National Science Foundation (NWO; grant VICI453-09-005) for financial support. Koopman acknowledges support from CREATES, the Center for Research in Econometric Analysis of Time Series (DNRF78) at Aarhus University, Denmark, funded by the Danish National Research Foundation, (DNRF78). We thank the Editor for the invitation and his general support.

1 INTRODUCTION

Over the last decades, time-varying parameter models have become increasingly popular in empirical economics and finance. The fast development of new methods for filtering time-varying parameters in dynamic models with nonlinear and non-Gaussian features has made these models more accessible, flexible and attractive. Starting from the 1960s, time-varying parameters for the mean equation in linear Gaussian models were initially typically handled by the Kalman filter and related methods. For a given linear Gaussian state space model, the Kalman filter can be used to calculate conditional means and variances of unobserved time-varying parameters (or linear functions thereof) in a computationally efficient manner; see, for example, Durbin and Koopman (2012) for a detailed treatment. In this modeling framework, the construction of in-sample and out-of-sample confidence bounds is straightforward and is performed on a routine basis as expressions for the conditional variances of the time-varying parameters are explicitly available. In case of nonlinear and/or non-Gaussian extensions of state space models, the computation of confidence bounds can be somewhat more involved. An example is the stochastic volatility model for which simulation-based methods are typically used for its analysis; see the discussions in Shephard (2005).

Since the 1980s, other model classes for time-varying parameters have been developed. Specifically, models for time-varying conditional variances have received much attention in the empirical economic financial literature. For example, the generalized autoregressive conditional heteroskedasticity ((G)ARCH) model by Engle (1982) and Bollerslev (1986) have led to a range of model formulations for time-varying parameters. In the standard ARCH

and GARCH models, the conditional variance is obtained from filtering past observations through a volatility updating equation. The relative simplicity of GARCH models has spurred their widespread use in both the academic and professional worlds.

In most empirical studies, the estimated volatility from the GARCH model is presented without in-sample bounds that reflect parameter uncertainty in the volatility updating equation. Exact analytical results are not available because the actual estimates are highly nonlinear functions of past observations. As a result, statistical software programs rarely provide bounds for estimates of such time-varying parameters. This argument also applies to other models related to GARCH including the autoregressive conditional duration (ACD) model of Engle and Russell (1998), the multiplicative error model of Engle (2002), the observation driven Poisson count model of Davis, Dunsmuir, and Streett (2003), and the score driven models of Creal et al. (2013). All these models belong to the class of observation driven models as opposed to the class of parameter driven models, which includes the state space model briefly discussed above; see Cox (1981) for a detailed description.

Here we review and analyze the behavior of two simple approximate analytical expressions for in-sample bounds that reflect parameter uncertainty. We further consider a simulation method for the evaluation of these bounds. The three different methods can be readily implemented in software packages. The approximate analytical bounds are much simpler to compute and are not subject to small random fluctuations originating from simulation error. The in-sample bounds can be used as long as the updating equation is differentiable and the estimator of the static parameters is asymptotically normally

distributed. We investigate in detail the coverage probabilities of each of these different approaches over a range of different time-varying parameter models and confidence levels. We find that simulation based methods are the most reliable, but that the more advanced approximate analytical bound also performs well in many settings, and is much easier to compute.

The presented in-sample bounds ignore other sources of uncertainty, apart from parameter uncertainty, such as measurement uncertainty and model uncertainty. Measurement uncertainty typically arises from problems in data collection or from difficulties in observing the data of interest. In certain applications, measurement uncertainty can be safely ignored. In other cases, we can treat this uncertainty by, for example, incorporating observational noise in the model. This effectively turns the problem of measurement uncertainty back into a problem of the uncertainties in model specification and parameter estimation. Model uncertainty is virtually always ignored when producing confidence bounds as it is simply too difficult to integrate it into the design of bounds. Researchers will usually work under the assumption of correct specification, and apply a battery of specification tests in order to judge how reasonable that assumption is. In certain cases, however, model uncertainty is less problematic. For example, we can ignore model uncertainty when the statistical model is sufficiently general to contain the data generating process with high probability. We can refer to the method of sieves for estimation in infinite dimensional spaces of unbounded model complexity; see Grenander (1981) and the review of Chen (2007).

To provide evidence of how effective the different methods are in computing the in-sample bounds, we present the results of a Monte Carlo study in

which we compute in-sample bounds for generated time series from GARCH, score driven, ACD and time-varying mean (local level) models. The results for these models reveal that the actual coverage of the analytical bounds from our preferred method is close to the nominal coverage obtained by simulation. An empirical application for the GARCH model applied to a time series of monthly log-returns from the Standard & Poor's 500 index reveal the practical importance of these bounds. We also show that the choice of method for computing in-sample bounds is empirically relevant, and that our preferred analytical bounds provide a good approximation to the more computationally intensive simulation based bounds.

This paper is organized as follows. Section 2 introduces the class of observation driven models. Section 3 introduces different bounds for the time-varying parameter. Section 4 analyzes the relative performance of these bounds in a Monte Carlo study. Section 5 presents our empirical findings for the Standard & Poor's 500 monthly returns. Section 6 concludes.

2 OBSERVATION DRIVEN MODEL AND ESTIMATION

In observation driven models, the time-varying parameter is filtered using an updating equation that depends only on past observations. The main focus is on how to let past realizations of the variable of interest affect the current value of the time-varying parameter. This can be achieved by means of different specifications.

Consider a time series model for an observed time series y_1, \dots, y_T given by

$$y_t \sim p_y(y_t | f_t; \boldsymbol{\theta}), \quad t = 1, \dots, T, \quad (1)$$

where density $p_y(\cdot)$ is implied by a model equation for y_t and is a function of the time-varying parameter f_t and the static parameter $\boldsymbol{\theta}$, for example, $y_t = f_t + \varepsilon_t$ for a time-varying mean, or $y_t = \mu + f_t^{1/2}\varepsilon_t$ for a fixed mean and time-varying variance, where $\varepsilon_t \sim \text{N}(0, 1)$. The time-varying parameter is formally defined as a function $f_t := f_t(y^{1:t-1}, f_1; \boldsymbol{\theta})$ that depends on the past observations $y^{1:t-1} := \{y_1, y_2, \dots, y_{t-1}\}$, on some initial value f_1 , and on a static parameter vector $\boldsymbol{\theta}$. The updating function for the time-varying parameter can be expressed in many different ways. For example, we can consider a linear updating equation consisting of lagged values of y_t and f_t . When only considering single lags, we obtain

$$f_{t+1} = \omega + \beta f_t + \alpha s(y_t, f_t; \boldsymbol{\theta}), \quad (2)$$

with initialization f_1 and where $s(y_t, f_t; \boldsymbol{\theta})$ is some, possibly, nonlinear function. The function $s(\cdot)$ can be chosen in a flexible way and is often just a transformation of y_t as we will show in the examples below. The coefficients ω , α and β are part of the parameter vector $\boldsymbol{\theta}$. The recursive nature of the formulation implies that f_{t+1} is a (nonlinear) function of y_1, \dots, y_t, f_1 and $\boldsymbol{\theta}$. Hence, the updating equation (2) is consistent with the definition of f_t , that is $f_t := f_t(y^{1:t-1}, f_1; \boldsymbol{\theta})$. We can also write the updating equation (2) in the general form of a stochastic recurrence equation,

$$f_{t+1} = \phi(y_t, f_t; \boldsymbol{\theta}), \quad (3)$$

with initialization f_1 and recurrence function $\phi(\cdot)$. The choice of the function $s(\cdot)$ in (2) thus defines the type of updating used for f_t .

EXAMPLE 1. *The generalized autoregressive conditional heteroskedasticity model of Engle (1982) and Bollerslev (1986), known as the GARCH model, for a mean-adjusted financial return series y_1, \dots, y_T is a special case of equations (1) and (2) with*

$$y_t = f_t^{1/2} \varepsilon_t, \quad \varepsilon_t \sim NID(0, 1), \quad s(y_t, f_t; \boldsymbol{\theta}) = y_t^2,$$

for $t = 1, \dots, T$, where $NID(0, 1)$ refers to a standard normally independently distributed sequence. The GARCH- t model of Bollerslev (1986) is obtained when we replace the normal by the Student's t distribution.

EXAMPLE 2. *The autoregressive conditional duration model of Engle and Russell (1998), known as the ACD model, for irregularly spaced data is also a special case of equations (1) and (2) with*

$$y_t = f_t \varepsilon_t, \quad \varepsilon_t \sim Exp(1), \quad s(y_t, f_t; \boldsymbol{\theta}) = y_t,$$

for $t = 1, \dots, T$, where $Exp(1)$ is the standard Exponential distribution. Different specifications of this modeling framework for durations and intensities are discussed in Grammig and Maurer (2000).

The score driven models of Creal et al. (2013) and Harvey (2013) also belong to the model class represented by equations (1) and (2) where $p_y(y_t|f_t; \boldsymbol{\theta})$ can be any density with f_t as the time-varying parameter. The specific choice of function $s(\cdot)$ distinguishes the score driven model from other models. It is chosen as the scaled first derivative of the predictive log density function

for y_t , that is

$$s(y_t, f_t; \boldsymbol{\theta}) = S_t \frac{\partial \log p_y(y_t | f_t, y^{1:t-1}; \boldsymbol{\theta})}{\partial f_t},$$

for some scaling function $S_t := S_t(f_t; \boldsymbol{\theta})$. In many models of practical interest, the predictive density $p_y(y_t | f_t, y^{1:t-1}; \boldsymbol{\theta})$ reduces to the conditional observation density $p_y(y_t | f_t; \boldsymbol{\theta})$ since f_t is also a function of $y^{1:t-1}$. The scaling S_t can be simply set to unity or alternatively be set to reflect the local curvature in the log conditional density function at time t .

When we consider the observation densities of Examples 1 and 2 and apply the score driven framework, we obtain equivalent updating functions if we scale the score by the inverse conditional Fisher information matrix; see Creal et al. (2013). Indeed the score driven models include many well-known and popular dynamic models. However, in the case of a Student's t distribution for $p_y(y_t | f_t; \boldsymbol{\theta})$ in equation (1), we obtain the model as discussed in Harvey and Chakravarty (2008) and Creal et al. (2011).¹

EXAMPLE 3. *The Student's t score driven model discussed in Harvey and Chakravarty (2008) and Creal et al. (2011, 2013) has $s(\cdot)$ given by*

$$s(y_t, f_t; \boldsymbol{\theta}) = (1 + 3\lambda^{-1}) \left(\frac{(1 - \lambda^{-1})y_t^2}{1 + \lambda^{-1}y_t^2/f_t} - f_t \right),$$

where λ is the degrees of freedom in the Student's t distribution. This expression uses a scaling function based on the inverse Fisher information matrix; see Creal et al. (2011, 2013) for more details.

The parameter vector $\boldsymbol{\theta}$ is unknown and needs to be estimated to obtain

¹More literature on score driven models with theoretical and empirical developments is provided via the website <http://www.gasmodel.com>.

estimates of the time-varying parameter f_t . The estimation of $\boldsymbol{\theta}$ can be based on the principle of maximum likelihood as we can evaluate the loglikelihood function $\ell(\boldsymbol{\theta}; f_1)$ via the prediction error representation. We have

$$\ell(\boldsymbol{\theta}; f_1) = \sum_{t=1}^T \log p_y(y_t | f_t, y^{1:t-1}; \boldsymbol{\theta}),$$

where values for f_t are evaluated via equation (2) with a specific initial value f_1 . When it is difficult to find a “natural” value for f_1 , it can be included in the parameter vector $\boldsymbol{\theta}$ and estimated simultaneously with the other parameters in $\boldsymbol{\theta}$. We notice that in many, if not most, cases of empirical interest we have $p_y(y_t | f_t, y^{1:t-1}; \boldsymbol{\theta}) = p_y(y_t | f_t; \boldsymbol{\theta})$ such that we can rely simply on the model density when computing the loglikelihood function. Hence maximum likelihood estimation reduces to the basic task of numerically maximizing $\ell(\boldsymbol{\theta}; f_1)$ with respect to $\boldsymbol{\theta}$. This is the standard practice for GARCH and related models.

We write the maximum likelihood estimator of $\boldsymbol{\theta}$ as $\widehat{\boldsymbol{\theta}}_T$. The values of f_t obtained from equation (2) under $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}_T$ are denoted by \widehat{f}_t for $t = 1, \dots, T$. The \widehat{f}_t 's can be regarded as a weighted function of f_1 and $y^{1:t-1}$ with the weights determined by $\widehat{\boldsymbol{\theta}}_T$, for $t = 2, \dots, T$. We can summarize the sources of uncertainty for \widehat{f}_t as follows:

- Parameter uncertainty: we do not know the true parameter vector $\boldsymbol{\theta}$ but we replace it by its maximum likelihood estimate $\widehat{\boldsymbol{\theta}}_T$ to compute \widehat{f}_t ;
- Filtering uncertainty: we consider the updating equation (2) for f_t but we do not know the underlying time-varying parameter process;

- Model uncertainty: we consider $p_y(\cdot)$ in equation (1) but we do not know the true data generation process for y_t .

As discussed in the introduction, in this paper we concentrate on developing bounds that reflect the parameter uncertainty for \hat{f}_t , $t = 1, \dots, T$ while taking the other two sources of uncertainty as given. This is an important step forward in the current literature, where in-sample uncertainty bounds around the time-varying parameter are almost never shown. The next section presents several ways to construct the approximate bounds. Section 4 investigates the coverage properties of these bounds by conducting a Monte Carlo study.

3 ASYMPTOTIC APPROXIMATION OF PARAMETER UNCERTAINTY

Conditional on past information and the initial value f_1 , and based on the maximum likelihood estimate $\hat{\boldsymbol{\theta}}_T$ for $\boldsymbol{\theta}$, we develop confidence bounds for \hat{f}_t that reflect parameter uncertainty about the true value of $\boldsymbol{\theta}$. The bounds are based as usual on the variance of \hat{f}_t , which we denote as $V_t := \text{Var}(\hat{f}_t)$, and reflect the randomness of the estimator $\hat{\boldsymbol{\theta}}_T$. We propose three different methods to obtain approximations for these bounds:

1. simple non-cumulative bounds;
2. cumulative delta-method bounds;
3. simulation based bounds.

The first two methods are analytic and very fast. They approximate the variance of the filter V_t by linearizing the updating function $\phi(\cdot)$ and using the asymptotic variance of $\widehat{\boldsymbol{\theta}}_T$. These analytic methods have the advantage of easy implementation in software packages as they are based on simple calculations. The third method provides probably the most accurate reflection of parameter uncertainty in \widehat{f}_t but it comes with the disadvantage that the computations are subject to randomness by construction. This method also uses the asymptotic variance of $\widehat{\boldsymbol{\theta}}_T$, but does not require the linearization of the updating function $\phi(\cdot)$. Hence this method may be preferable when nonlinearities play an important role in the filter.

The *simple non-cumulative bounds* do not take into account the accumulation of parameter uncertainty in the updating or filtering process for \widehat{f}_t . They rely on simple expressions, but are also somewhat naive and therefore only deliver accurate results for a restrictive class of observation driven models. Specifically, these bounds only become interesting when the autoregressive component is either not present, that is when $\beta = 0$ in equation (2), or when it is sufficiently small. For example, the method is appropriate for the ARCH model of Engle (1982), which is the model of Example 1 with $\beta = 0$.

The *cumulative delta-method bounds* take the accumulation of parameter uncertainty in the updating process into account. These bounds are obtained by a simple application of the delta-method and are relevant for all observation driven models, including GARCH, ACD, and score driven models where the updating equation for \widehat{f}_t features an autoregressive term, that is $\beta \neq 0$.

The *simulation bounds* deal not only with the accumulation of the uncer-

tainty due to the parameter vector $\boldsymbol{\theta}$, but also treat the effects of possible nonlinear functional relations between $\boldsymbol{\theta}$ and \widehat{f}_t . These bounds are most appropriate for cases where we have strong nonlinear expressions in the updating process and the linearization error produced by the delta-method cannot be ignored. In contrast to the first two methods for computing bounds, this method is not fully analytic and requires repeated simulations of the estimated time-varying parameter paths \widehat{f}_t .

All methods for computing bounds rely on the asymptotic normality of the estimator $\widehat{\boldsymbol{\theta}}_T$. The bounds can therefore be computed for any estimator $\widehat{\boldsymbol{\theta}}_T$ as long as it has a normal distribution asymptotically. The Monte Carlo study in Section 4 investigates the extent to which this reliance on the asymptotic distribution of the estimator may be problematic in small samples.

3.1 SIMPLE NON-CUMULATIVE BOUNDS

Our simple non-cumulative bounds can generally be obtained by linearizing the function $\phi(y_t, \widehat{f}_t; \widehat{\boldsymbol{\theta}}_T)$, as defined in equation (3) with $f_t = \widehat{f}_t$ and $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}_T$, around the true parameter vector used for generating y_1, \dots, y_T , that is $\boldsymbol{\theta}_0 \in \Theta$ where Θ is the parameter space. We use \widehat{f}_t to denote the filter that is a function of the point estimate $\widehat{\boldsymbol{\theta}}_T$. In the case of the non-cumulative bounds, we take \widehat{f}_t as given in calculating the approximate variance of \widehat{f}_{t+1} . Our proposed simple bounds then rely on the approximation

$$\widehat{f}_{t+1}(y^{1:t}, \widehat{f}_1; \widehat{\boldsymbol{\theta}}_T) \approx \phi(y_t, \widehat{f}_t; \boldsymbol{\theta}_0) + \sum_{i=1}^q \frac{\partial \phi(y_t, \widehat{f}_t; \boldsymbol{\theta}_0)}{\partial \theta_i} (\widehat{\theta}_{T,i} - \theta_{0,i}), \quad (4)$$

for any f and around some point $\boldsymbol{\theta}$, where $\theta_{0,i}$ is the i th element of the $q \times 1$ true parameter vector $\boldsymbol{\theta}_0$ and, similarly, $\widehat{\theta}_{T,i}$ is the i th element of the $q \times 1$ maximum likelihood estimate vector $\widehat{\boldsymbol{\theta}}_T$, for $i = 1, \dots, q$.

We define the partial derivative of $\phi(y_t, \widehat{f}_t; \boldsymbol{\theta})$ with respect to θ_i , the i th element of $\boldsymbol{\theta}$, as the function

$$\nabla_{i,t} = \nabla_i(y_t, \widehat{f}_t; \boldsymbol{\theta}) := \frac{\partial \phi(y_t, \widehat{f}_t; \boldsymbol{\theta})}{\partial \theta_i}, \quad i = 1, \dots, q. \quad (5)$$

Hence we can write the approximation (4) as

$$\widehat{f}_{t+1} \approx \phi(y_t, \widehat{f}_t; \boldsymbol{\theta}_0) + \sum_{i=1}^q \nabla_{i,t} \times (\widehat{\theta}_{T,i} - \theta_{0,i}),$$

for $t = 1, \dots, T$. The approximate variance of \widehat{f}_{t+1} for given \widehat{f}_t is then obtained by

$$V_{t+1} = \text{Var}(\widehat{f}_{t+1}) \approx \sum_{i=1}^q \nabla_{i,t}^2 \text{Var}(\widehat{\theta}_{T,i}) + 2 \sum_{1 \leq i < j \leq q} \nabla_{i,t} \nabla_{j,t} \text{Cov}(\widehat{\theta}_{T,i}, \widehat{\theta}_{T,j}), \quad (6)$$

where Var and Cov are the variance and covariance operators, respectively. In case the derivative $\nabla_i(y_t, f; \boldsymbol{\theta})$ is a function of $\boldsymbol{\theta}$ and/or f , we evaluate the derivative at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}_T$ and $f = \widehat{f}_t$.

Let the maximum likelihood estimator $\widehat{\boldsymbol{\theta}}_T$ be asymptotically normally distributed as

$$\sqrt{T}(\widehat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0) \xrightarrow{d} N(0, \mathbf{W}), \quad (7)$$

where $\boldsymbol{\theta}_0$ is the true parameter vector and $N(0, \mathbf{W})$ is the multivariate normal distribution with zero mean vector and covariance matrix \mathbf{W} . The standard

notation \xrightarrow{d} is for *convergence in distribution*. We typically take \mathbf{W} as the robust sandwich covariance matrix estimator and evaluate it at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}_T$. We then write explicitly $\mathbf{W} = \widehat{\mathbf{W}}$. It further follows that

$$\text{Var}\left[\sqrt{T}\left(\widehat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\right)\right] \approx \widehat{\mathbf{W}} \quad \Leftrightarrow \quad \text{Var}\left(\widehat{\boldsymbol{\theta}}_T\right) \approx T^{-1}\widehat{\mathbf{W}}.$$

From the asymptotic properties of the maximum likelihood estimator $\widehat{\boldsymbol{\theta}}_T$, we can deduct that the asymptotic variance of \widehat{f}_{t+1} converges asymptotically to V_{t+1} where

$$V_{t+1} \approx \sum_{i=1}^q T^{-1}w_{i,i}\nabla_{i,t}^2 + 2 \sum_{1 \leq i < j \leq q} T^{-1}w_{i,j}\nabla_{i,t}\nabla_{j,t}, \quad (8)$$

and with $w_{i,j}$ as the (i, j) element of matrix \mathbf{W} . In a standard fashion and based on the asymptotic normal distribution for \widehat{f}_{t+1} , we obtain asymptotic 95% in-sample confidence bounds for \widehat{f}_{t+1} as

$$\left[\widehat{f}_{t+1} - 1.96\sqrt{V_{t+1}} \quad , \quad \widehat{f}_{t+1} + 1.96\sqrt{V_{t+1}} \right]. \quad (9)$$

EXAMPLE 4. (*Cont'd from Example 1*) Consider the ARCH model of Engle (1982) which is defined as the GARCH model of Example 1 with $\beta = 0$ in equation (2) and with $\boldsymbol{\theta} = (\omega, \alpha)'$. The resulting updating function for \widehat{f}_{t+1} is given by $\widehat{f}_{t+1} = \widehat{\omega}_T + \widehat{\alpha}_T y_t^2$ where $\widehat{\omega}_T$ and $\widehat{\alpha}_T$ are respectively the first and second elements of the parameter vector $\widehat{\boldsymbol{\theta}}_T$. It follows immediately that

$$\nabla_{1,t} = 1, \quad \nabla_{2,t} = y_t^2, \quad t = 1, \dots, T.$$

The asymptotically correct bounds obtained from our simple non-cumulative approach are then given by (9) with

$$\begin{aligned} V_{t+1} &= \text{Var}(\hat{\omega}_T) + y_t^4 \text{Var}(\hat{\alpha}_T) + 2y_t^2 \text{Cov}(\hat{\omega}_T, \hat{\alpha}_T) \\ &\approx T^{-1}(w_{1,1} + y_t^4 w_{2,2} + 2y_t^2 w_{1,2}). \end{aligned}$$

These expressions are very simple to compute numerically once \mathbf{W} has been replaced by its estimate $\hat{\mathbf{W}}$.

In Section 4 we provide evidence that our simple non-cumulative bounds are not appropriate for observation driven models where the filter has strong autoregressive dynamics. This applies to the GARCH model with $\beta \neq 0$ and in particular to the case where β is far away from zero.

3.2 CUMULATIVE DELTA-METHOD BOUNDS

The previous approximations did not account for the fact that also \hat{f}_t itself depends on the estimator $\hat{\boldsymbol{\theta}}_T$. This is an important defect. In filters with an autoregressive component, the parameter uncertainty accumulates as the filter evolves over time. This occurs because \hat{f}_{t+1} depends on \hat{f}_t , which already contains parameter uncertainty. This accumulation of parameter uncertainty in the filtering process can be analytically tracked by application of the delta-method. It is based on the same linearization (4) of $\phi(\cdot)$, but where we also explicitly account for the dependence of \hat{f}_t on $\hat{\boldsymbol{\theta}}_T$. We redefine $\nabla_{i,t}$ as

$$\nabla_{i,t} = \nabla_i(y_t, f_t; \boldsymbol{\theta}) := \frac{\partial \phi(y_t, f_t; \boldsymbol{\theta})}{\partial f_t} \frac{\partial f_t}{\partial \theta_i} + \frac{\partial \phi(y_t, f_t; \boldsymbol{\theta})}{\partial \theta_i}. \quad (10)$$

We assume that the point estimate $\widehat{\boldsymbol{\theta}}_T$ is consistent and sufficiently close the true parameter $\boldsymbol{\theta}_0$. Hence when calculating $\nabla_{i,t}$ we take f_t as the value of the filter at time t for the point estimate of $\widehat{\boldsymbol{\theta}}_T$, that is \widehat{f}_t . The derivatives $\partial f_t / \partial \theta_i$ can then be computed resursively. For example, for the GARCH model we have

$$\frac{\partial f_{t+1}}{\partial \theta_i} = \left(\frac{\partial \omega}{\partial \theta_i} + \frac{\partial \alpha}{\partial \theta_i} y_t^2 + \frac{\partial \beta}{\partial \theta_i} f_t \right) + \beta \frac{\partial f_t}{\partial \theta_i}, \quad (11)$$

which can be computed in parallel to the recursion for f_t itself.

The variance V_{t+1} of the filtered estimate \widehat{f}_t is again approximated by (8), but with $\nabla_{i,t}$ defined as in equation (10). Also the asymptotic 95% in-sample confidence bounds for \widehat{f}_{t+1} can be based on equation (9), but with V_{t+1} computed as indicated above.

The delta-method bounds are subject to two approximation errors: (i) the linearization of the updating equation; and (ii) the use of the asymptotic distribution of the estimator as an approximate distribution. As we shall see in Section 4, these bounds are accurate when: (i) the the updating equation is sufficiently well approximated by a linear function; and (ii) when the sample is sufficiently large for the distribution of the estimator to be approximately normal. These two conditions are met in many situations of empirical interest.

The following example illustrates the application of the current theory to the case of the GARCH model.

EXAMPLE 5. *(Cont'd from Example 1) Consider the GARCH model of Bollerslev (1986) as given by equation (2) and with $\boldsymbol{\theta} = (\omega, \alpha, \beta)'$. We have*

$y_t = \sqrt{f_t}u_t$ with $u_t \sim NID(0, 1)$ and $f_{t+1} = \omega + \alpha y_t^2 + \beta f_t$ for $t = 1, \dots, T$.

Applying equation (10) to the GARCH updating function, it follows that

$$\nabla_{1,t} = 1 + \beta \frac{\partial f_t}{\partial \omega}, \quad \nabla_{2,t} = y_t^2 + \beta \frac{\partial f_t}{\partial \alpha}, \quad \nabla_{3,t} = f_t + \beta \frac{\partial f_t}{\partial \beta}, \quad t = 1, \dots, T,$$

where the partial derivatives on the right-hand side are defined via the recursion (11). We evaluate f_t and f_{t-1} by \hat{f}_t and \hat{f}_{t-1} , respectively, and replace the elements of $\boldsymbol{\theta}$ by the corresponding elements of $\hat{\boldsymbol{\theta}}_T$. The resulting approximation of the variance of \hat{f}_{t+1} is then given by

$$\begin{aligned} V_{t+1} \approx & \left(1 + \hat{\beta}_T \frac{\partial f_t}{\partial \omega}\right)^2 w_{1,1} + \left(y_t^2 + \hat{\beta}_T \frac{\partial f_t}{\partial \alpha}\right)^2 w_{2,2} + \left(\hat{f}_t + \hat{\beta}_T \frac{\partial f_t}{\partial \beta}\right)^2 w_{3,3} \\ & + 2 \left(1 + \hat{\beta}_T \frac{\partial f_t}{\partial \omega}\right) \left(y_t^2 + \hat{\beta}_T \frac{\partial f_t}{\partial \alpha}\right) w_{1,2} \\ & + 2 \left(1 + \hat{\beta}_T \frac{\partial f_t}{\partial \omega}\right) \left(\hat{f}_t + \hat{\beta}_T \frac{\partial f_t}{\partial \beta}\right) w_{1,3} \\ & + 2 \left(y_t^2 + \hat{\beta}_T \frac{\partial f_t}{\partial \alpha}\right) \left(\hat{f}_t + \hat{\beta}_T \frac{\partial f_t}{\partial \beta}\right) w_{2,3}, \end{aligned}$$

where $w_{i,j}$ denotes element (i, j) of the asymptotic variance matrix \mathbf{W} of $\hat{\boldsymbol{\theta}}_T$. The necessary computations are more involved when compared to those required for the simple non-cumulative method used in Example 4. The bounds, however, are still straightforward to implement and fast to compute.

When the updating equation is highly nonlinear, the error in the above linearizations may become substantial, resulting in inaccurate approximations for V_{t+1} . The in-sample bounds obtained from using the simulation method below avoid the linearization step altogether. The resulting bounds are only subject to the error in approximating the distribution of the esti-

mator by its asymptotic distribution.

3.3 SIMULATION BOUNDS

An alternative approach to obtain in-sample confidence bounds for \widehat{f}_{t+1} is using simulation methods. Since the bounds only reflect the uncertainty in $\widehat{\boldsymbol{\theta}}_T$, one can obtain exact confidence bounds by drawing parameter values $\boldsymbol{\theta}^i$ from the distribution of $\widehat{\boldsymbol{\theta}}_T$ and running the filter from $t = 1$ to $t = T$ for every simulated $\boldsymbol{\theta}^i$. However, since the finite-sample distribution of $\widehat{\boldsymbol{\theta}}_T$ is unknown, we draw parameter values from the approximating asymptotic distribution instead,

$$\widehat{\boldsymbol{\theta}}_T^i \sim N\left(\widehat{\boldsymbol{\theta}}_T, T^{-1}\widehat{\mathbf{W}}\right), \quad i = 1, \dots, M, \quad (12)$$

for some predefined number M . For each i , the sequence $f_1^i, \widehat{f}_2^i, \dots, \widehat{f}_T^i$ can then be determined using equations (2) or (3) with $f_1^i = f_1$. In particular, we consider the filtering recursion for each i , that is

$$\widehat{f}_{t+1}^i = \phi(y_t, \widehat{f}_t^i; \widehat{\boldsymbol{\theta}}_T^i), \quad t = 1, \dots, T.$$

The simulation method is rather different from the two earlier methods. Instead of linearizing the filtering recursion and working with an approximate Gaussian distribution for \widehat{f}_t , we can make use of simulations to obtain more accurate bounds. The uncertainty of $\widehat{\boldsymbol{\theta}}_T$ is characterized by the asymptotic distribution (7). Also note that in the simulation method the bounds do not need to be based on V_{t+1} , but can rather be obtained directly by calculating the appropriate percentiles for each t over the M draws of the filtered paths

\widehat{f}_t^i for $i = 1, \dots, M$.

When nonlinearities play a prominent role in the updating equation $\phi(\cdot)$, the simulation bounds may become more accurate than those based on the linearization methods. However, simulations can be time consuming when the sample size T is large and when a high level of accuracy (large M) is required. Also simulation methods are inevitably subject to simulation error. Hence the simulation bounds may be less attractive for its implementation in software packages.

EXAMPLE 6. (*Cont'd from Example 1*) Consider the GARCH model with $y_t \sim NID(0, f_t)$ and $f_{t+1} = \omega + \alpha y_t^2 + \beta f_t$, for initial value f_1 . The simulation bounds can be simply obtained by drawing M parameter vectors from the approximate density of the estimator as in (12), we obtain $\widehat{\boldsymbol{\theta}}_T^i = (\widehat{\omega}_T^i, \widetilde{\alpha}_T^i, \widetilde{\beta}_T^i)$ and

$$\widehat{f}_{t+1}^i = \widetilde{\omega}_T^i + \widetilde{\alpha}_T^i y_t^2 + \widetilde{\beta}_T^i \widehat{f}_t^i, \quad \text{for } i = 1, \dots, M.$$

From these simulated paths we compute in-sample bounds for any confidence level.

4 MONTE CARLO STUDY

To verify the performances of the three methods for computing the in-sample bounds of time-varying parameters in observation driven models, we carry out a Monte Carlo study. We consider five different models: (i) the GARCH model based on normally distributed innovations; (ii) GARCH based on Student's $t(5)$ innovations; (iii) the Student's $t(5)$ based score driven volatility model of Creal et al. (2011, 2013) and Harvey (2013); (iv) the autoregres-

sive conditional duration (ACD) model; and (v) a time-varying mean (or local level) model with Gaussian innovations. The GARCH model is introduced in Example 1. In case the density $p_y(y_t|f_t; \boldsymbol{\theta})$ is $NID(0, f_t)$, we obtain Model (i); in case it is Student's $t(5)$ with variance f_t and degrees of freedom λ , we obtain Model (ii). Model (iii) has $S_t(f_t; \boldsymbol{\theta}) = 1 + 3\lambda^{-1}$, and $s(y_t, f_t; \boldsymbol{\theta}) = (1 + \lambda)y_t^2/(\lambda - 2 + y_t^2/f_t)$ with $\lambda > 2$ degrees of freedom. Model (iv) is introduced in Example 2. The local level (LL) model (v) is a basic application of the score driven model of Example 3. We set $p_y(y_t|f_t; \boldsymbol{\theta})$ to the Gaussian density $NID(f_t, \sigma^2)$ and choose $S_t(f_t; \boldsymbol{\theta}) = 1$ and $s(y_t, f_t; \boldsymbol{\theta}) = y_t$. For models (i) and (iv), the parameter vector is given by $\boldsymbol{\theta} = (\omega, \alpha, \beta)'$, for models (ii) and (iii) it is $\boldsymbol{\theta} = (\omega, \alpha, \beta, \lambda)'$ and for model (v) it is $\boldsymbol{\theta} = (\omega, \alpha, \beta, \sigma^2)'$. Further details are given in Table 1. We choose ω and f_1 such that the time-varying parameter process is started at its unconditional mean. We consider two different sample sizes, $T = 500, 1000$, and we generate the time series for three values of $\beta = 0.2, 0.5, 0.8$, with $\omega = 0.05$, $\alpha = 0.1$, $\lambda = 5$, $\sigma^2 = 1$ and $f_1 = 1$.

The aim of this Monte Carlo study is to verify how well the three methods studied in Section 3 approximate the correct in-sample bounds for \widehat{f}_{t+1} . The in-sample bounds can be determined in a Monte Carlo setting by repeating the following steps:

1. Generate a time series y_1, \dots, y_T from one of the models listed in Table 1 for some T .
2. For this generated time series, estimate $\boldsymbol{\theta}$ using the method of maximum likelihood as discussed in Section 2.

Table 1: Models in Monte Carlo study

We present the details of the five observation driven models in our Monte Carlo study. In all cases the updating equation for the time-varying parameter f_t is given by (2), that is $f_{t+1} = \omega + \alpha s_t + \beta f_t$, with $\omega = 0.05$, $\alpha = 0.1$ and for three different values of $\beta = 0.2, 0.5, 0.8$. The remaining details for $p_y(y_t|f_t; \boldsymbol{\theta})$ and $s_t = s(y_t, f_t; \boldsymbol{\theta})$ are given below with a reference to the examples discussed in Section 2. Student's $t(0, f_t, \lambda)$ refers to the Student's t density function with zero mean, variance (rather than scale parameter) f_t and degrees of freedom λ with $\lambda = 5$, $Exp(f_t)$ refers to the Exponential density with intensity parameter f_t , and for model (v) we have set $\sigma^2 = 1$. In all cases, we take $f_1 = 1$.

	Model	$p_y(y_t f_t; \boldsymbol{\theta})$	$s(y_t, f_t; \boldsymbol{\theta})$	Example
(i)	GARCH	$NID(0, f_t)$	y_t^2	1
(ii)	GARCH- t	iid Student's $t(0, f_t, \lambda)$	y_t^2	1
(iii)	GAS- t	iid Student's $t(0, f_t, \lambda)$	$\frac{(1+3\lambda^{-1})(1+\lambda)y_t^2}{\lambda-2+y_t^2/f_t}$	3
(iv)	ACD	iid $Exp(f_t)$	y_t	2
(v)	Local Level (LL)	$NID(f_t, \sigma^2)$	y_t	–

- Calculate the time-varying parameter sequence \widehat{f}_{t+1} as well as its derivative with respect to $\boldsymbol{\theta}$ using (2) and (11) for a given f_1 , based on the generated time series and the estimate $\widehat{\boldsymbol{\theta}}_T$.

These three steps can be repeated N times so that we obtain a set of N paths for \widehat{f}_t . From this set we can empirically determine the coverage of the in-sample bounds for \widehat{f}_t for nominal confidence levels such as 90%, 95% and 99%.

Within this Monte Carlo framework, for each generated time series, we determine the in-sample bounds by the three methods as described in Section 3. We compute the time series average of the number of times the true f_t lies outside the in-sample bound. We take the average of this number over the N replications, and present the results in Table 2.

Table 2: In-sample bounds coverages

For our Monte Carlo study, we present the 90%, 95% and 99% coverages of the in-sample bounds, as computed by the three methods non-cumulative simple, cumulative delta and simulation, for the observation driven models as described in Table 1. In case of the simulation method, we use $M = 1000$ simulations. The Monte Carlo study is based on $N = 1000$ replications.

T	β	GARCH			t -GARCH			t -GAS			ACD			LL		
		90	95	99	90	95	99	90	95	99	90	95	99	90	95	99
<i>Non-cumulative bounds</i>																
500	0.2	76.3	81.9	88.8	74.7	80.1	86.6	74.7	80.2	87.2	76.7	82.1	88.8	79.4	85.0	91.4
500	0.5	67.1	73.4	82.1	62.6	69.0	77.8	63.7	69.8	78.4	68.7	74.7	82.8	69.2	75.6	84.4
500	0.8	46.0	51.8	61.6	40.1	45.4	54.6	43.8	49.4	58.5	50.9	56.8	66.2	46.1	52.0	61.5
1000	0.2	78.9	84.3	91.0	77.7	83.6	89.9	75.8	82.0	89.3	78.1	83.7	90.3	81.1	86.6	92.8
1000	0.5	66.7	73.4	82.7	63.3	69.7	79.2	63.6	69.9	79.1	70.1	76.2	84.5	69.3	76.0	85.1
1000	0.8	46.0	51.9	61.8	38.7	44.0	53.2	43.1	48.6	57.6	51.3	57.4	67.0	46.6	52.6	62.2
<i>Cumulative bounds</i>																
500	0.2	86.3	91.8	97.2	87.1	91.7	96.1	87.8	92.1	96.7	86.7	92.3	97.4	86.2	91.9	97.3
500	0.5	85.4	90.7	96.2	86.9	91.6	96.0	89.3	93.1	96.8	86.1	91.6	96.8	86.5	92.0	97.4
500	0.8	86.4	91.5	96.6	86.2	91.1	95.8	91.6	94.5	97.4	85.7	91.3	96.8	88.6	94.3	98.8
1000	0.2	86.1	91.9	97.2	88.2	93.1	97.5	89.6	94.2	98.0	86.1	91.6	97.2	85.9	91.4	96.8
1000	0.5	85.8	91.6	96.9	86.7	91.8	96.6	90.5	94.3	97.7	86.5	91.9	96.9	85.1	91.0	96.8
1000	0.8	87.7	92.9	97.5	87.4	92.8	97.2	93.3	96.0	98.4	86.5	92.0	97.3	88.5	93.7	98.3
<i>Simulation bounds</i>																
500	0.2	92.1	96.0	99.0	91.8	95.3	98.3	92.7	96.6	99.5	92.6	96.5	99.3	90.8	95.4	98.8
500	0.5	92.2	96.0	99.0	93.1	96.3	98.6	94.5	97.8	99.6	92.7	96.3	99.0	91.0	95.6	99.0
500	0.8	93.0	96.8	99.2	93.1	96.8	99.2	94.5	98.1	99.8	93.3	96.9	99.1	89.2	94.3	98.5
1000	0.2	91.9	96.2	99.1	92.2	96.3	99.2	93.1	97.0	99.3	91.3	95.6	99.0	91.1	95.6	98.9
1000	0.5	92.3	96.5	99.2	92.1	96.1	99.1	93.4	97.1	99.4	92.4	96.4	99.1	90.5	95.3	98.9
1000	0.8	92.0	96.3	99.2	92.1	96.6	99.2	91.9	96.4	99.4	92.0	96.1	99.0	89.1	94.0	98.0

As expected, the simple non-cumulative bounds are not very accurate for the GARCH model with $\beta = 0.8$, while they are somewhat more accurate for the model with $\beta = 0.2$. However, the simple bounds provide poor overall results for the three coverages of 90%, 95% and 99%. As we have indicated in Example 4, the simple method is appropriate for the ARCH model (that is the GARCH model with $\beta = 0.0$). Our Monte Carlo study reveals that the simple method produces rather inaccurate results even when $\beta = 0.2$. These findings are confirmed by the other models and for both sample sizes considered.

The cumulative delta-method bounds are considerably more accurate. The bands are slightly too small, resulting in lower estimated coverage compared to the nominal levels. The differences, however, appear to be tolerable for most practical purposes. The performances for different values of β remain stable, while the performances of the simple non-cumulative bounds deteriorate rapidly if the value for β is increased. The estimated coverage levels for the cumulative delta-method bounds remain close to the corresponding nominal levels for $\beta = 0.2, 0.5, 0.8$. The performance slightly improves if the sample size is increased. The overall performance is very stable across the different models.

The reported coverages for the simulation method are presented in the bottom panel of Table 2. The in-sample bounds from the simulation method are based on $M = 1000$ simulations. The presented coverages show that the simulation method is most accurate overall. Again, the improvements in accuracy compared to the cumulative bounds is clearly noticeable, but appears to be tolerable for empirical applications. This is a relevant finding,

as the cumulative bounds are easier and faster to compute than the simulation based ones. This finding applies to all levels of persistence for the time-varying parameter process and to all different models considered. In our next section, we investigate whether this conclusion also holds up for empirical data.

5 AN EMPIRICAL ILLUSTRATION

To illustrate the use and the appearance of the different in-sample bounds in an empirical study, we consider a monthly time series of Standard & Poor's (S&P) 500 index returns. Our sample covers January 1990 until January 2015. We consider the standard GARCH model of Example 1 for illustrative purposes with non-zero conditional mean μ . The returns for our sample are displayed in Figure 1. The maximum likelihood estimates for the coefficients are given by $\hat{\mu}_T = 0.145$ (0.195), $\hat{\omega}_T = 0.583$ (0.346), $\hat{\alpha}_T = 0.171$ (0.053), and $\hat{\beta}_T = 0.805$ (0.054), with the corresponding standard errors in parentheses, where we set f_1 equal to the sample variance. The filtered conditional variances \hat{f}_t are presented in the second panel of Figure 1.

The standard errors of the filtered \hat{f}_t are computed by our three methods. We present the results over the last 8 years of the sample in Figure 2. We clearly corroborate that the simple non-cumulative bounds are unrealistically small compared to the other two bands. It is interesting to find that the simulation bounds and the cumulative delta bounds can be hardly distinguished from each other in the presented sample. The largest difference can be noted at the height of the financial crisis, where the lower bound of the simulation approach lies substantially below the cumulative lower bound. In all other

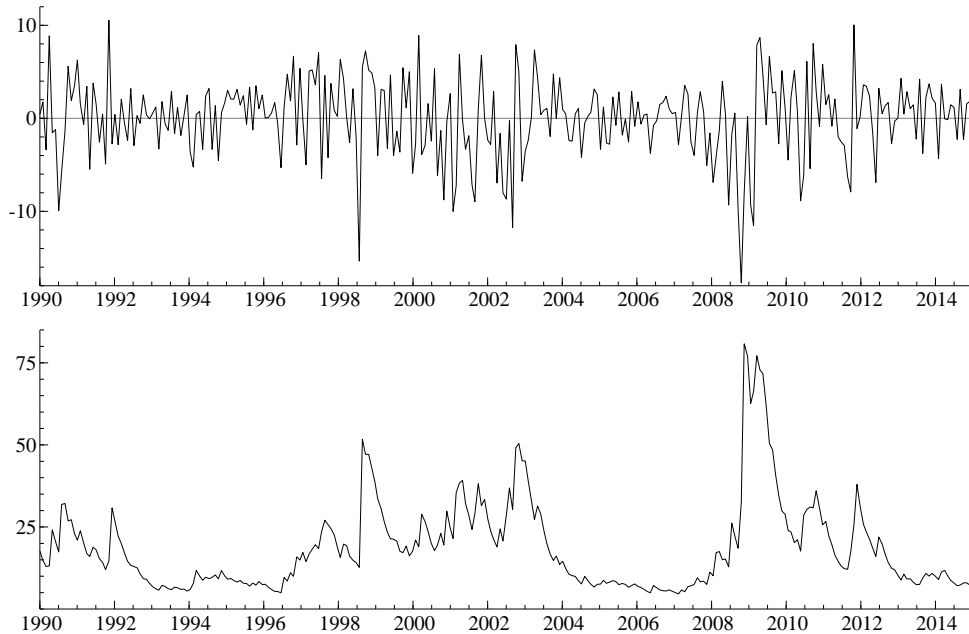


Figure 1: We present the S&P500 monthly returns from January 1990 until January 2015 (top panel) and the conditional volatility estimates obtained from the standard GARCH model (bottom panel).

parts of the sample, the cumulative analytic approximation bounds appear to be highly successful in approximating the bounds of uncertainty around \hat{f}_t , this is in terms of accuracy, computational simplicity and speed.

6 CONCLUSIONS

We have reviewed three different methods for the computation of in-sample confidence bounds for time-varying parameters in a general class of observation driven models . These bounds reflect the uncertainty due to parameter estimation. In a Monte Carlo study it is shown that the simulation-based

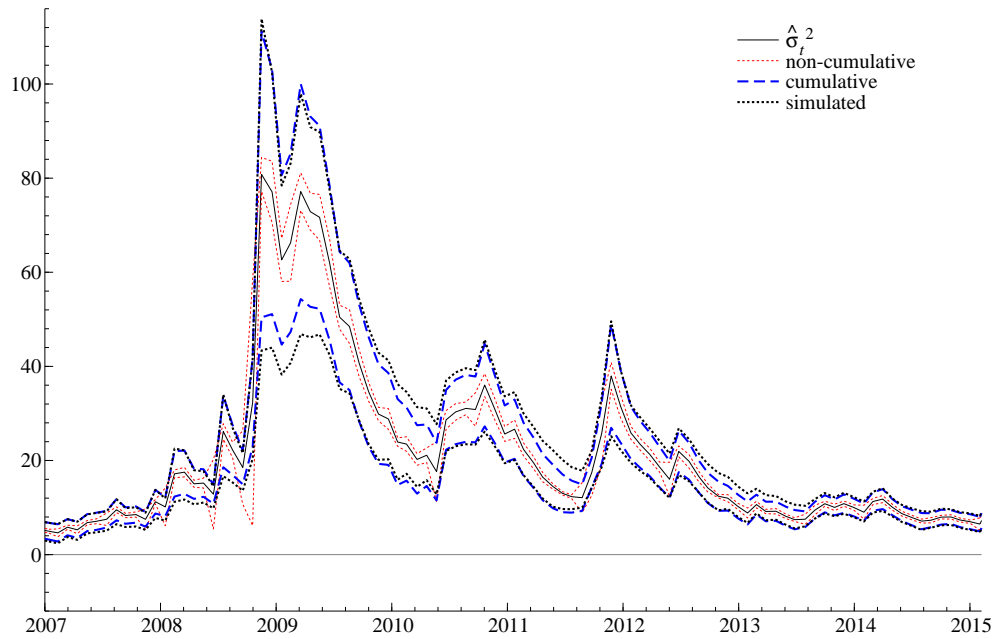


Figure 2: We present conditional volatility estimates and their corresponding in-sample error bounds from the GARCH model applied to the S&P500 index monthly returns from January 1990 until January 2015; we only present these for the last 8 years (96 months), from January 2007. The in-sample bounds are obtained from the non-cumulative simple method (light dotted line), the cumulative delta-method (dashed line) and the simulated method (dark dotted line).

method is most accurate for obtaining the correct coverages. Analytical non-cumulative simple bounds are not sufficiently accurate and should not be used in practice if the time-varying parameter possesses persistent autoregressive dynamics. Interestingly, our analytic approximate bounds based on a simple recursive delta-method approximation appear to reflect parameter uncertainty well, both in a controlled Monte Carlo study and in our empirical illustration. The delta-method is straightforward and fast. Therefore it can easily be implemented in standard software packages. This would provide useful additional information in empirical analyses, where confidence bounds around the time-varying parameters are usually not presented. Although we have illustrated the methods for well-known models, the different approaches for computing in-sample bounds can also be used for the skewed Student's t distribution with time-varying parameters as in Lucas, Schwaab, and Zhang (2014) and for dynamic discrete data models as in Rydberg and Shephard (2003). Future research may concentrate on extending our approach to the computation of out-of-sample bounds that reflect the uncertainty of future innovations for the time-varying factor as well.

REFERENCES

- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31(3), 307–327.
- Chen, X. (2007). Large sample sieve estimation of semi-nonparametric models. In R. F. Engle and D. M. Fadden (Eds.), *Handbook of Econometrics*, Volume 6b, Chapter 76, pp. 5549–5632. Elsevier.

- Cox, D. R. (1981). Statistical analysis of time series: some recent developments. *Scandinavian Journal of Statistics* 8, 93–115.
- Creal, D., S. J. Koopman, and A. Lucas (2011). A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. *Journal of Business and Economic Statistics* 29(4), 552–563.
- Creal, D., S. J. Koopman, and A. Lucas (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics* 28(5), 777–795.
- Davis, R., W. Dunsmuir, and S. Streett (2003). Observation-driven models for Poisson counts. *Biometrika* 90, 777–790.
- Durbin, J. and S. J. Koopman (2012). *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflations. *Econometrica* 50, 987–1008.
- Engle, R. F. (2002). New frontiers for ARCH models. *Journal of Applied Econometrics* 17(5), 425–446.
- Engle, R. F. and J. R. Russell (1998). Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica*, 1127–1162.
- Grammig, J. and K. O. Maurer (2000). Non-monotonic hazard functions and

- the autoregressive conditional duration model. *Econometrics Journal* 3, 16–38.
- Grenander, U. (1981). *Abstract inference*. New York: Wiley.
- Harvey, A. C. (2013). *Dynamic Models for Volatility and Heavy Tails: with Applications to Financial and Economic Time Series*. Econometric Society Monographs. Cambridge: Cambridge University Press.
- Harvey, A. C. and T. Chakravarty (2008). Beta- t (E)GARCH. *University of Cambridge, Faculty of Economics, Working paper CWPE 08340*.
- Lucas, A., B. Schwaab, and X. Zhang (2014). Measuring credit risk in a large banking system: econometric modeling and empirics. *Journal of Business and Economic Statistics* 32, 271–284.
- Rydberg, T. H. and N. Shephard (2003). Dynamics of trade-by-trade price movements: decomposition and models. *Journal of Financial Econometrics* 1(1), 2.
- Shephard, N. (2005). *Stochastic Volatility: Selected Readings*. Oxford: Oxford University Press.