Option-implied risk measures: An empirical examination on the S&P500 index

Giovanni Barone-Adesi * **Chiara Legnazzi** †‡

Carlo Sala §

May 29, 2019

Abstract

The forward-looking nature of option market data allows one to derive economically-based and model-free risk measures. This article proposes an extensive analysis of the performances of option-implied VaR and CVaR, and compare them with classical risk measures for the S&P500 Index. Delivering good results both at short and long time horizons, the proposed option-implied risk metrics emerge as a convenient alternative to the existing risk measures.

Keywords: Option Prices, Option-Implied VaR and CVaR, Long and Short-term Risk Measures, S&P 500 Index.

JEL classification: G13, G32, D81

---

[*]Swiss Finance Institute at Università della Svizzera Italiana (USI), Institute of Finance, Via G. Buffi 13, CH-6900 Lugano, Switzerland.
Tel: +41 58 666 4753 E-mail: `giovanni.baroneadesi@usi.ch`

[†]Swiss Finance Institute at Università della Svizzera Italiana (USI), Institute of Finance, Via G. Buffi 13, CH-6900 Lugano, Switzerland.
Tel: +41 78 653 77 18 E-mail: `chiara.legnazzi@usi.ch`

[‡]We are grateful for the support of the Swiss Finance Institute (SFI) and the Swiss National Foundation (SNF) grant 153135.

[§]Department of Financial Management and Control, ESADE Business School, Ramon LLull University, Avenida de Torreblanca 59, 08172 Sant Cugat, Barcelona, Spain.
Tel: +34 932 806 162 E-mail: `carlo.sala@esade.edu`

1 Introduction

"...risk management models generally covered only the past two decades, a period of euphoria. Had instead the models been fitted more appropriately to historic periods of stress, capital requirements would have been much higher and the financial world would be in far better shape today."

Alan Greenspan, 1987-2006 Chairman of the US Federal Reserve (FED)

Thanks to their ease of application and to different empirical and theoretical convenient features,[1] the Value at Risk (VaR) and the Conditional Value at Risk (CVaR) are to date the most widely used risk measures in finance. Summarizing the market risk in a single value, both risk measures assess the maximal loss at a specified probability level over a fixed time horizon. The risk management literature proposes a wide variety of approaches to estimate the VaR and CVaR. To date, most of the existing methodologies infer their estimates from past returns, using more or less sophisticated statistical procedures (henceforth: "statistically-based" models). Unfortunately, the use of past returns leads to backward-looking final results which, especially during crashes, may poorly predict the dynamics of financial markets. This is opposed to option market data. Reflecting the investors' expectations over different time horizons, option prices are by construction forward-looking, and so the estimation deriving from them.
Using option market data as input and proposing both a nonparametric and a parametric approach, this paper tests if the option-based (henceforth: "option-implied") VaR and CVaR can deliver accurate risk forecasts, and compare their performances with those of the relative statistically-based risk measures. Performing well both at short and long time horizons and at different risk levels, our results show that option-implied risk measures can be considered as valid alternatives to the classical statistically-based risk metrics.

From a theoretical viewpoint, most of results coming from the existing statistically-based models might be prone to economic and econometric biases linked to the estimation of the future profit and loss (P&L) distribution. First, even though at different extents, all statistically-based risk models involve errors in the modeling of the underlying asset volatility, thus being exposed to the "risk

[1] Refer to Duffie and Pan (1997), Dowd (1999), Artzner et al. (1999), Penza and Bansal (2000) and Jorion (2007) for an overview on the topic.

that the risk will change" (see Engle (2009) and Engle (2011)).

Second, as anticipated, statistically-based risk models use historical data as inputs, thus making the resulting estimates backward-looking by construction. More formally, at each trading day only one stock return is observable. Being the observations just a single value (a scalar), it is not possible to infer a density out of it. To tackle this problem, a common approach is to enlarge the amount of data used in the estimation process by adding historical returns to the current value. Unfortunately, when piling up historical returns to estimate the future P&L distribution, the last values become almost uninformative. The obtained P&L distribution cannot be sensitive to the current market situation; on the contrary, it is strongly dependent to the time window and amount of data used in the estimation process. As consequence, the resulting statistically-based risk measures are biased by construction.

Third, it is well-known that the statistical properties of market prices depend on the general market outlook (e.g.: Danielsson (2002) and Brunnermeier and Pedersen (2009) and references therein). In stable periods the heterogeneity among investors' behavior increases, whereas during high volatility periods agents' actions become more similar (e.g.: bank runs and flights to safety). As a consequence, the statistical analysis conducted in stable periods can rarely be helpful during highly volatile ones, and viceversa.

Fourth, the time horizon of the estimates is a crucial element in risk management and the use of historical data makes the estimation of long-time horizon risk measures challenging. From an econometric viewpoint, statistically-based risk metrics often need to rely on a proper estimation of the market volatility. Brownlees et al. (2011) show that several ARCH models perform well at short time horizons, while results are often unsatisfactory at medium and long time horizons. To date, statistically-based risk measures at time horizons longer than one day are obtained with more or less complex numerical transformations.[2] Unfortunately, how to economically justify most of these transformations is still an open question. All these points have great practical relevance, as short-time horizon estimates might not be of great help if one has to liquidate a large (and possibly illiquid) financial position.

Option market data can possibly overcome the aforementioned problems. First, compared to

[2]A very common approach is to estimate a one day risk forecast and extend it to a longer time horizon by multiplying it times the square root of the time horizon desired.

historical returns, option market data have a superior informative power, especially as concerns the modeling of the underlying asset volatility.[3] Second, the implied moments of option market data reflect investors' expectations, thus being a naturally forward-looking indicator of agents' beliefs. Third, at each observation date, option market data have a matrix structure, the so called option cross-section, which reflects the investors' future beliefs across different strikes and times-to-maturity. Both dimensions, strike prices and times-to-maturity, are important in risk management. Through the different strike prices, it is possible to infer the desired future P&L distribution, without the need of additional historical data. Through the different times-to-maturity, it is possible to naturally extract option-implied risk forecasts with a time horizon equal to the time-to-maturity of the options. All these properties make the daily option cross-section an interesting input to use to infer the future P&L distribution.

Surprisingly, despite the increasing attention of academics and practitioners on the predictive content of derivative securities, the use of option market data in risk management is still widely unexplored. To the best of our knowledge, only Äit-Sahalia and A.W.Lo (2000), Bali et al. (2011), Samit (2012), Mitra (2015) and Huggenberger et al. (2018) and reference therein relate the topic.[4]

After presenting the option-implied methodology developed by Barone-Adesi (2016), this paper proposes an extensive empirical analysis on the performances of the option-implied risk measures and it compares them with the statistically-based ones. Several backtesting results show that the option-implied VaR and CVaR estimates are accurate in the short (weekly) and long (monthly) term. On a relative level, the option-implied methodology outperforms the statistically-based VaR and CVaR estimated with an asymmetric GJR-GARCH model with nonparametric innovations estimated with the filtered historical simulated (FHS) approach of Barone-Adesi et al. (1999) (henceforth denoted as GJR-GARCH-FHS). The choice of the volatility model stems from Engle and Rosenberg (2002), Barone-Adesi et al. (2008), Brownlees et al. (2011) and **?** who show that, for the S&P 500, the asymmetric GJR-GARCH model of Glosten et al. (1993) is the best performer

[3]From Merton (1973) this is a large and still very active stream of literature.

[4]Less directly related but still worth mentioning are Bali et al. (2011), Xing et al. (2010), and Yan (2011) that show how option-implied skewness or jump risk measures have significant explanatory power in pricing the cross-section of asset returns.

among a big family of different discrete volatility models.[5] The use of FHS innovations is to better capture the nonparametric features of financial markets.

Technicalities aside, the main difference between the two approaches (option-implied and statistically-based) is the type of information contained into the input data. Our results show that inferring the risk measures from option market data has both theoretical and practical advantages. Results could be of interest for regulators, central banks and single companies. Regulators and central banks can derive risk estimates of large companies without knowing the precise composition of their portfolio. Large companies can compare the option-implied estimates with those delivered by their internal models.

The remainder of the article is as follows: Section 2 reviews the concepts of VaR and CVaR. Section 3 introduces and derives the option-implied VaR and CVaR, showing how to link the risk measures presented in the previous section to the option market data. In this section we present both the parametric and nonparametric approach to derive the option-implied VaR and CVaR. Section 4 describes the dataset. Section 5 presents the empirical results of the nonparametric and parametric option-implied VaR and CVaR. Using stock market data, the same risk metrics are also estimated using the GJR-GARCH-FHS model. Section 6 analyses and compares the backtesting results for all the aforementioned risk estimates. Section 7 concludes.

2 VaR and CVaR

This section quickly recalls the definitions and the main theoretical properties of the VaR and CVaR. Defined as the quantile of the projected P&L distribution, the $\mathrm{VaR}^{\alpha}_{t,T}$ is the maximum theoretical loss forecasted at current time t, associated with a generic portfolio value, $S_{t,T}$, at a given time horizon, T$<\infty$, and at a given risk level $(1-\alpha)$, where $\alpha \in [0,1]$ determines the chosen risk level. Formally, for $t < T < \infty$, the $VaR^{\alpha}_{t,T}$ is defined as the α-quantile of the return distribution:

$$\mathrm{VaR}^{\alpha}_{t,T} = \int_{-\infty}^{K} f(S_{t,T}) dS_{t,T} = \alpha \tag{1}$$

[5] We also performed the same analysis on many other models e.g.: the variance-covariance model, the historical model, the expected weighted moving average model and many different ARCH specifications. Results confirm the findings of the cited literature. Being the best performer among all models, and for ease of space, we only propose the GJR-GARCH-FHS estimates. All other results are available upon request.

where $f(S_{t,T})$ is the probability distribution function of the *future* portfolio values based on its current value. Because of its forward-looking nature, the pricing probability distribution of the *future* portfolio values, $f(S_{t,T})$, is the most challenging element to estimate. Statistically-based risk measures usually extract this future distribution from past data, thus making the final risk measure backward-looking by construction.

Although empirically easy to implement and practically appealing, the lack of sub-additivity and the impossibility to quantify the risk over the chosen quantile make the VaR an incomplete and unreliable risk measure. The lack of sub-additivity[6] implies that the VaR is not even weakly coherent[7] and discourages diversification.

Proposed by Artzner et al. (1999) and supported by the Basel Committee, the CVaR overcomes most of the aforementioned VaR limitations. The $\mathrm{CVaR}^{\alpha}_{t,T}$ is formally defined as the time t conditional expected loss determined in Equation (1) at level α:

$$\mathrm{CVaR}^{\alpha}_{t,T} = \frac{1}{\alpha}\int_{-\infty}^{K} L(S_{t,T}) f(S_{t,T}) dS_{t,T} \qquad L(S_{t,T}) = S_t - S_{t,T} \tag{2}$$

where $L(S_{t,T})$ defines the predicted loss function. It is worth noticing that, as for the VaR, also the CVaR requires the forecast of $f(S_{t,T})$. While the CVaR accounts for losses exceeding the VaR and is coherent, the lack of elicitability (Gneiting (2011) and Ziegel (2014)) makes the CVaR inferior to the VaR when performing a backtesting analysis.[8] Directly linked to the historical nature of the estimates, the general concept of elicitability allows one to perform meaningful comparisons among models based on the backtesting results. While the VaR is a quantile and is thus naturally elicitable, the CVaR is not.

[6]The VaR is sub-additive only if all marginals of the joint distribution are elliptical. This condition coincides with the assumption underneath the Markowitz's variance-minimizing portfolio, which makes the VaR calculations not even needed once that the variance is known (Markowitz (1952)).
[7]A risk measure is weakly coherent if it lacks of the positive homogeneity (Carr et al. (2001)).
[8]According to Acerbi and Székely (2014) the CVaR is only backtestable if the P&L distribution is known.

3 Linking VaR and CVaR to the option market

This section shows how to derive the option-implied VaR and the CVaR. As a main advantage, inferring the risk measures from the option cross-section leads to economically-grounded and naturally forward-looking risk forecasts.

As defined in Equation (1), the $\text{VaR}_{t,T}^{\alpha}$ is a quantile, a single numeric value determined at a specific threshold over the cumulative P&L distribution. In the context of an Arrow-Debreu economy (Arrow and Debreu (1954)) and following Breeden and Litzenberger (1978), the first derivative of a European Put price, $p_{t,T} = e^{-r_t(T-t)} \int_0^K (K - S_T) f(S_{t,T}) dS_{t,T}$, over its strike price, K, is:

$$\frac{dp_{t,T}}{dK} = \frac{d[e^{-r_t(T-t)} \int_0^K (K - S_T) f(S_{t,T}) dS_{t,T}]}{dK} \tag{3}$$

$$= e^{-r_t(T-t)} \int_0^K f(S_T) dS_{t,T} \tag{4}$$

$$= e^{-r_t(T-t)} F(K) \tag{5}$$

$$= e^{-r_t(T-t)} \alpha \tag{6}$$

where r_t is the current market risk-free rate such that $e^{-r_t(T-t)}$ discounts the future expected payoff at today t, and α is the risk level (i.e. 1 - confidence level), as defined in Section 2. With no loss of generality, the lower bound of the integral has been set to 0, reflecting the economy of an investor holding a portfolio with limited liability (i.e.: where the maximum loss is equal to 100%). It follows that for each date t, the *option-implied* $\text{VaR}_{t,T}^{\alpha}$ is the difference between the initial portfolio value, S_t, and the strike price of a European Put option at level α:

$$\text{VaR}_{t,T}^{\alpha} = S_t - K_{t,T}^{\alpha} \tag{7}$$

where $K_{t,T}^{\alpha}$ is the strike price of the option contract traded at time t, with maturity T, at level α. Being $K_{t,T}^{\alpha}$ fully option-based and not requiring any *past* data, the proposed option-implied VaR is naturally *forward-looking*.

The use of Put options links the analysis to the left tail of the distribution. The use of Call options leads to the same result but for right tail of the distribution thus identifying the upside risk. A right tail VaR is a useful risk measure for an investor holding a short position in the underlying asset. The option-implied VaR for a short portfolio is:

$$\text{VaR}^{\alpha}_{t,T} = K^{\alpha}_{t,T} - S_t \tag{8}$$

where $K^{\alpha}_{t,T}$ is now the strike price identified by α of an European Call option[9].

Following the same intuition, it is possible to derive the *option-implied* CVaR. Starting from the definition of a compounded Put option:

$$e^{r_t(T-t)} p_{t,T} = \int_0^K (K - S_T) f(S_{t,T}) dS_{t,T} \tag{9}$$

and setting the current stock value of the loss function as $S_t = K + S_t - K$, it follow from Equations (1), (2) and (9) that the *option-implied* $\text{CVaR}^{\alpha}_{t,T}$ based on the left tail of the distribution (downside risk) is:

$$\text{CVaR}^{\alpha}_{t,T} = \text{VaR}^{\alpha}_{t,T} + e^{r_t(T-t)} \frac{p^{\alpha}_{t,T}}{\alpha_{t,T}} \tag{10}$$

where $p^{\alpha}_{t,T}$ is the current price of the compounded Put option identified by the risk level α. Just relying on the option market data, the option-implied $\text{CVaR}^{\alpha}_{t,T}$ equals the sum of the corresponding $\text{VaR}^{\alpha}_{t,T}$ and an additional term, which equals the compounded Put price divided by the probability of that negative scenario, i.e. alpha. As for the VaR, the upside risk (right tail) option-implied CVaR can be derived from Call prices and is defined as:

$$\text{CVaR}^{\alpha}_{t,T} = \text{VaR}^{\alpha}_{t,T} + e^{r_t(T-t)} \frac{c^{\alpha}_{t,T}}{\alpha_{t,T}} \tag{11}$$

where $c^{\alpha}_{t,T}$ is the current price of the compounded Call option identified by the risk level α. It is worth noticing that all quantities in Equations (7), (8), (10) and (11) are readily available in the market. The only parameter to estimate is the threshold $\alpha_{t,T}$, which is also needed to determine $K^{\alpha}_{t,T}$, the value of the strike price at the desired risk level. In general, $\alpha_{t,T}$ can be estimated either i) nonparametrically, i.e.: using more or less sophisticated numerical differentiation approaches, or ii) parametrically, i.e.: by specifying an option pricing model. In this paper $\alpha_{t,T}$ is estimated

[9] In addition to the stock market, the knowledge of the upside risk is also of interest in the commodities market, where sellers (buyers) are mostly affected by the downside (upside) risk (Barone-Adesi et al. (2019))

i) nonparametrically, using the interpolation method of Barone-Adesi and Elliot (2007) and, ii) parametrically, under a Black and Scholes economy. While the nonparametric method provides more accurate results, the parametric one is less data intensive.

For the nonparametric approach, the time series of $\alpha_{t,T}$ is extracted from the daily cross-section of European Put/Call options. Free of any distributional assumption, the estimates based on this methodology are denoted as "model-free" (MF) option-implied risk measures. Following Barone-Adesi and Elliot (2007) we discretize equations (3)-(6) and define the MF $\alpha_{t,T}$ as the average of the first order condition of three contiguous option prices (e.g.: Put), with maturities equal to the desired investment horizon T:

$$\alpha_{t,T} = e^{r_t(T-t)}\left[\frac{1}{2}\left(\frac{dp_{t,T}^{\mathrm{UP}}}{dK^{\mathrm{UP}}} + \frac{dp_{t,T}^{\mathrm{DOWN}}}{dK^{\mathrm{DOWN}}}\right)\right] \tag{12}$$

$$= e^{r_t(T-t)}\left[\frac{1}{2}\left(\frac{p_3 - p_2}{K_3 - K_2} + \frac{p_2 - p_1}{K_2 - K_1}\right)\right] \tag{13}$$

where $p_3 > p_2 > p_1$ and $K_3 > K_2 > K_1$. Equation (13) holds if $K_3 - K_2 = K_2 - K_1$. In presence of non equidistant option prices, Equation (13) can be easily generalized by performing a weighted average with the weights equal to the distance among the strike prices. Lastly, as defined in Equation (6), the obtained value is corrected for the daily market risk-free rate, $r_t(T-t)$.

Similarly, in the case of a short portfolio $\alpha_{t,T}$ can be either derived using Call options:

$$\alpha_{t,T} = e^{r_t(T-t)}\left[\frac{1}{2}\left(\frac{dc_{t,T}^{\mathrm{UP}}}{dK^{\mathrm{UP}}} + \frac{dc_{t,T}^{\mathrm{DOWN}}}{dK^{\mathrm{DOWN}}}\right)\right] \tag{14}$$

$$= e^{r_t(T-t)}\left[\frac{1}{2}\left(\frac{c_1 - c_2}{K_1 - K_2} + \frac{c_2 - c_1}{K_2 - K_1}\right)\right] \tag{15}$$

where $c_1 > c_2 > c_3$ and $K_1 > K_2 > K_3$.

Two features make the proposed MF approach numerically efficient and accurate. First, it eliminates the first-order error in the Taylor expansion of the derivative. Second, it eliminates the first-order error due to the implied volatility changing across strike prices. As a drawback, the derivation of $\alpha_{t,T}$ might be a problem for non liquid assets. Given the satisfactory level of liquidity of the option market data used in this paper (European Put and Call options written on the S&P500 Index) we do not need to implement any extrapolation/smoothing technique, and we leave the dataset almost untouched. In case of necessity, it is indeed possible to adopt one of the many numerical techniques present in the literature to extrapolate/smooth the option pricing curve and have a more populated final dataset.

Differently from the nonparametric MF approach, the option-implied parametric methodology implicitly assumes that the observed option prices are generated from a model. In this paper we use the Black and Scholes (B&S) pricing model and denote the estimates based on this methodology as the "Black and Scholes" (B&S) option-implied risk measures. Following Black and Scholes (1973), $\alpha_{t,T}$ is the probability that the today option contract will expire in-the-money at expiration T. For example, for a European Put option:

$$\begin{aligned} \alpha_{t,T} &= Pr(S_T < K^{\alpha}_{t,T}) \\ &= N(-d_2) \\ &= 1 - N\left[\frac{\ln(\frac{S_t}{K^{\alpha}_{t,T}}) + (r_t - \frac{\sigma_t^2}{2})(T-t)}{\sigma_t\sqrt{(T-t)}}\right] \end{aligned} \tag{16}$$

The values of $\alpha_{t,T}$ obtained numerically in equation (13) are compared with the ones derived in equation (16) at the same confidence level. As presented in the empirical section, the nonparametric nature of the MF approach delivers more reliable estimates, while the B&S is a valid alternative for low liquidity periods of the option market. Using Call options, the probability that the today option contract will expire in-the-money at expiration T is $Pr(S_T > K^{\alpha}_{t,T}) = N(d_2)$.

The proposed option-implied MF and B&S risk estimates are obtained from a small number of option prices, a characteristic which makes the resulting estimates more robust with respect to those ones based on the VIX and SKEW indexes, which instead require the existence of an infinite amount of Put and Call options. Requiring an infinite amount of options traded, both the VIX and the SKEW indexes may be more prone to numerical errors (e.g. truncation errors).
It is worth noticing that the risk-neutral measure can be inferred from option market data using different econometric approaches (see Bahara (1997) for a comprehensive summary). Unfortunately, aside from the methodologies that use or extend the nonparametric approach of Breeden and Litzenberger (1978), most of the alternatives either rely on parametric assumptions, or require possibly

approximative (more or less complex) numerical simulations to infer the entire risk-neutral density. For example, simulated densities might not integrate to one and/or need to rely on a parametric extreme value theory approach to complete and model the tails of the distributions (Figlewski (2008)), thus imposing a distribution in the CVaR computations. These drawbacks may potentially bias the CVaR estimates. Bliss and Panigirtzoglou (2004)(2004) estimate the risk-neutral measure after a second order interpolation and extrapolation of the implied volatility. The cubic smoothing spline interpolation uses the options vega as a weight. Indeed, this approach puts most of weights at-the-money thus under-performing in the tails of the distribution and possibly impacting on the relative VaR and CVaR computations.

The nonparametric approach nature of $\alpha_{t,T}$, along with the no need of historical data for the estimation of the option-implied VaR and CVaR have interesting economic and econometric features. Economically, the MF option-implied VaR and CVaR are not simulated but inferred from current market data. This makes our option-implied risk measure similar to the Economic measures proposed by Äit-Sahalia and A.W.Lo (2000), that the authors consider economically stronger than the Statistical measure inferred from historical data. Econometrically, being free of any simulation and not relying on any parametric assumption, the option-implied risk measures can better capture the nonparametric nature of the finance markets.

As a disadvantage, the proposed option-implied risk measures depend on the option market liquidity. Therefore they can only be used for assets with traded and liquid options.

In conclusion: while risk estimates should provide economically consistent and econometrically robust results at different time horizons, this is hardly achievable with most of the statistically-based risk models, while it becomes possible with the proposed option-implied MF approach.

4 Dataset

Supported by the evidence[10] that a large class of investors holds a market portfolio as a form of investment, the chosen underlying asset is the S&P 500 Index. The methodology presented in

[10] As reported by different surveys, a big fraction of U.S. investors relies on indexing policies for investment. Analysing 60 years of market data, Bogle (2005) shows how, since 2000, index funds account for roughly one-third of equity fund cash inflows and represent about one-seventh of the total amount of equity fund assets.

Section 3 is tested on the European Put and Call options (SPX) traded on the S&P 500 Index with maturity of 7 and 30 days. Options with maturity of 7 and 30 days[11] are used to estimate the weekly and monthly risk metrics, respectively. To propose a broad empirical analysis, the risk levels used for the weekly time horizon are 1%, 2.5% and 5%, while for the monthly time horizon are 5%, 10% and 15%. To lighten the notation in the paper, the subscript (t, T) of the VaR and CVaR estimates indicates a time interval equal to 7 or 30 days for the weekly and monthly estimates, respectively. It will be made clear whether we will refer to estimations with a short (7 days - weekly) or a long (30 days - monthly) time horizon.

Our analysis is thus performed on two different datasets, one composed by weekly options and another one by monthly options. Starting from January 1, 2012 to August 31, 2015 the sample period of the weekly estimations consists of 159 weekly observations. Introduced by the CBOE in October 2005, it is only from the beginning of 2012 that the liquidity of the S&P 500 weekly options is high enough to implement the option-implied methodologies presented in the previous section. Accounting for more then 40% of the total volume of options traded on a daily basis in the US market[12], weekly options are now an interesting tool to better understand the behavior of markets and investors.

Starting from January 1, 2005 to August 31, 2015 the sample period of the monthly estimations consists of 128 weekly observations. The monthly analysis allows us to include into the analysis the financial and credit crisis. Given the monthly expiration convention, the presented risk measures refer to the third Thursday of each single month. All contracts have strike price intervals of 5$ for in-, at- and out-of-the-money strike prices and of 25$ for deeply-out-of-the-money strike prices. To discard possible mispricings, prices violating classical no-arbitrage lower bounds are excluded from the dataset:

$$p_{t,T} \geq \max[K\,\mathrm{e}^{-r_t(T-t)} - S_t\,\mathrm{e}^{-d_t(T-t)}, 0] \tag{17}$$

$$c_{t,T} \geq \max[S_t\,\mathrm{e}^{-d_t(T-t)} - K\,\mathrm{e}^{-r_t(T-t)}, 0] \tag{18}$$

[11]The VIX Index, probably the most famous options-based risk measures, also has a 30 days time horizon and is derived from a portfolio of Put and Call options written on the S&P 500 Index with 30 days to maturity.

[12]The most traded weekly options have Friday, Wednesday and Monday expirations. The great success of the S&P 500 weekly options convinced the CBOE to propose other similar products, e.g.: new even shorter expirations (2/3 days) and new underlying assets (weekly VIX). For more details, see the CBOE website.

where d_t is the continuously compounded dividend at time t. Moreover, following the literature, all quotes with a price lower than 3/8 and those with zero trading volume and/or zero open interest have been removed from the dataset.

From the time series of the daily zero coupons, the curve of the daily risk-free interest rates is obtained by numerical interpolation. For each time-to-maturity $(T-t)$, the previous $(T-t)_-$ and next $(T-t)_+$ period zero-coupon whose values straddle the time $(T-t)$ are linearly interpolated. As it concerns the dividend, the time series of S&P 500 Index dividend yield from January 1, 2012 to August 31, 2015 (weekly) and from January 1, 2005 to August 31, 2015 (monthly) are used to compute the ex-dividend spot Index level. All presented data are from OptionMetrics.

To validate the performances of the option-implied VaR and CVaR, these risk measures are compared with the statistically-based VaR and CVaR derived using a GJR-GARCH-FHS model and going back 10 years of data to estimate the GARCH parameters (details in Section 5). For the GARCH estimation, daily prices on the S&P500 Index are obtained from CRSP. We use data from January 1, 1995 to August 31, 2015.

5 Empirical Analysis

Following the methodology presented in Section 3, the weekly and monthly option-implied MF and B&S[13] CVaR and VaR are derived based on the left and right tails of the distribution. For the entire analysis the risk levels are set at 1%, 2.5% and 5% (5%, 10% and 15%) for the weekly (monthly) estimates.

Table 1 collects summary statistics for the option-implied MF and B&S VaR and CVaR. The table is so organized: CVaR (VaR) results are on top (bottom); each section has on the left (right) side the left (right) tail results estimated with Put (Call) options. Finally, each section has in the upper (lower) part the weekly (monthly) results. As expected, the difference between CVaR and VaR is bigger for the left tail of the distribution, thus showing a pricing difference between Put and Call options. Also, the difference between the MF and B&S approach is visible in the maximum and

[13]While we analyse and present summary statistics also of the option-implied B&S the focus of the paper will primarily be on the nonparametric MF approach. As anticipated in fact, the nonparametric nature of the MF approach provide more reliable data than the parametric B&S methodology, that nonetheless is a valid alternative in case of data shortage.

minimum values, with the nonparametric MF always presenting more dispersed data, as confirmed by the higher standard deviation. The same difference is also again present between the left and the right estimates, with the former being always greater than the latter.

To support our findings and provide an alternative intuition, Table 2 reports similar basic statistical information about the weekly and monthly option-implied $\Delta_{\alpha,t} = \text{CVaR}_{\alpha,t} - \text{VaR}_{\alpha,t}$ for both tails of the distribution. From Equation (10), the quantity $\Delta_{\alpha,t}$ represents the future market beliefs with respect to the negative scenario occurring with α% probability. This quantity is highly sensitive to the future market expectations. In periods of market turmoil, for example, investors are more risk averse and the prices of Put options are likely to increase, thus implying a larger difference between the CVaR and VaR estimates.

As the option-implied methodology strongly relies on the liquidity of the option market, the upper part of Table 3 reports the proportion of missing values for each risk level. Over the sample period under consideration the liquidity of weekly Put option contracts (1^{st} row) is always satisfactory as the proportion of missing estimates is above 1% only for the lowest risk level.

The same analysis is then repeated using Call options, thus assuming a portfolio of short positions. The derivation of the risk measures based on the right tail of the distribution has two interesting properties. First, it provides a benchmark for the estimates obtained using Put options. Second, it can provide nonparametric information on the degree of tail asymmetry in the underlying returns distribution. Based on the right tail of the distribution, both the CVaR and the VaR are decreasing in the risk level and display similar patterns to the corresponding metrics based on the left tail, both on a weekly and monthly basis. However, keeping the risk level α fixed, both the weekly and monthly MF VaR and CVaR based on the right tail are almost always below those obtained from the left tail. The second row of Table 3 and the left lower portion of Table 2 report the proportion of missing data and the summary statistics on $\Delta_{\alpha,t}$ for the weekly and monthly estimates based on the right tail of the distribution, respectively. In almost all cases the number of missing observations is larger compared to the one for the left tail, thus confirming the higher liquidity of out-of-the-money Put options compared to out-of-the-money Call contracts. It is also worth noticing that, for each α level, the average $\Delta_{\alpha,t}$ are much lower than the ones based on the left tail. For α fixed, this discrepancy is driven by the price of the identified option, thus confirming that out-of-the-money Call options are, on average, cheaper than the corresponding out-of-the-money Put options.

The three upper panels of Figure 1 show the parametric MF weekly CVaR (thicker in blue) and VaR (in red) for α equal to 1% (top), 2.5% (middle) and 5% (bottom). As expected, the picture shows a more volatile time series for α at 1%. This is a direct consequence of the price distribution, in particular into the deepest part of the left tail of the distribution. Also, the estimates do not substantially vary across the whole sample period, except for a different behavior at the three risk level for to the so-called China's "Black Monday" (happened in August 2015), which carried a correction of about 10% on the value of S&P 500 Index. Whilst the three risk levels show different behavior under the MF, they all have been able to promptly react and sometimes to anticipate it (at 2.5%). The three upper panels of Figure 2 show the parametric MF monthly CVaR (thicker in blue) and VaR (in red) for α equal to 1% (top), 2.5% (middle) and 5% (bottom). As for the weekly results, the trend among the three risk levels is overall similar but more volatile than the weekly estimates. Spanning over a larger time period, it is above all during the financial and credit crisis (June 2007 to December 2012) that both the MF CVaR and VaR exhibit several pronounced upward spikes across all risk levels, thus reflecting the negative outlook of the US market in those periods. For example, the biggest spike within the recession area represents the pre and post Lehman Brother failure (September 15, 2008). The subsequent jumps correspond to the downward S&P 500 movements of June 2010 (S&P 500 down of \$200 at \$1022.58 from \$1217.28 of two months before) and the period of turmoil from July to November 2011 during which the S&P 500 index dropped from above \$1300 losing more than \$200. At the rightmost part of the figure, also the crisis of Summer 2015 is well identified. As expected, and above all for the first part of the sample, some values are missing. This is a direct consequence of the option market liquidity, that improved substantially over time. Similar but unreported results hold also for the MF CVaR and VaR estimates based on the right tail of the distribution.

To provide an alternative to the more data intensive MF approach, we repeat the same analysis, but under the assumption that the observed option prices are generated by a Black and Scholes pricing model with $\alpha_{t,T}$ derived as defined in Equation (16). Under the B&S approach, $\alpha_{t,T}$ is derived with just a single option price. First, for both European Put and Call options, we back-out the option-implied volatilities by numerical inversion from the daily S&P 500 option market prices:

$$\sigma_t^{IV} = f(S_t, K, r_t, d_t, \tau, p_t/c_t) \tag{19}$$

where the time-to-maturity τ is fixed equal to 7 or 30 days for weekly and monthly options, respectively. Obtaining $\alpha_{t,T}$ through Equation (16), we then find $K^{\alpha}_{t,T}$, that is the strike price corresponding to the desired alpha $P(S_T < K^{\alpha}_{t,T}) = N(-d_2)$. The same holds for the Call options, with $Pr(S_T > K^{\alpha}_{t,T}) = N(d_2)$. Once the time series of K^{α}_t is derived, the weekly and monthly B&S VaR and CVaR are computed following Equations (7) and (10). As for most of parametric approaches, the biggest advantage of this approach is the existence of a closed-formula solution for the value of the threshold $\alpha_{t,T}$. Indeed, although one option is sufficient to determine alpha, the discreteness of quoted option prices in general requires interpolating between two adjacent strike price to compute the strike price at the desired α level.

The lower right part of Table 3 reports the proportion of missing values for the entire B&S sample. As expected, the proportion of missing data across all α specifications is almost zero and always smaller than that obtained under the MF methodology (upper right part of Table 3). The higher stability of the B&S values conflicts with the accuracy of the final results. A reduced smile effect is remarkable into the deepest area of the tails, where the B&S model provides approximative estimates. Among others, these effects can be summarized by the difference of alpha under the two methodologies:

$$\Gamma_t = K_t^{\alpha^{\mathrm{MF}}} - K_t^{\alpha^{\mathrm{BS}}} \tag{20}$$

Table 4 reports summary statistics of Γ_t for both the weekly (upper part) and monthly (lower part) observations. For both time horizons and tails, the behavior of the estimates is overall comparable both in terms of mean and median values. As expected, the min-max range is larger, thus confirming the non-normality of the dataset. Interestingly, also the highest and lowest statistical values show a similar behavior across the sample for both Put and Call options and for both short and long time horizons.

The three lower panels of Figure 1 show the parametric B&S weekly CVaR (thicker in blue) and VaR (in red) for α equal to 1% (top), 2.5% (middle) and 5% (bottom). At $\alpha = 1\%$, both the

B&S CVaR and VaR present a few large spikes, which are direct consequence of the high (possibly irrational) Put prices on those dates. Overall, for both the B&S CVaR and VaR the trends of the time series are more homogeneous than the MF ones. Also, and again differently than the MF approach, the B&S estimates at all risk levels capture very similarly (although with different magnitudes) the China's "Black Monday" . The three lower panels of Figure 2 show the parametric B&S monthly CVaR (thicker in blue) and VaR (in red) for α equal to 5% (top), 10% (middle) and 15% (bottom). Once more, the time series at all risk levels are very similar and the period from June 2007 to December 2012 coincides with the highest VaR and CVaR values. While at different magnitude, it is interesting to see how also the parametric approach is able to forecast the quiet pre-crisis period and anticipate the subsequent turmoil caused by the financial crisis. As expected, the parametric B&S has almost no liquidity problem and the time series are almost complete. Similar but unreported results hold also for the B&S CVaR and VaR estimates based on the right tail of the distribution.

Compared with the MF results, while the mean and median values of the weekly $\Delta_{\alpha,t}$ measures are comparable, the min-max range substantially differs. Keeping α and the compounding effect fixed, the two methodologies work differently into the tails of the distribution. Both for the weekly and monthly estimates, once compared with the nonparametric approach (presented in the right part of Table 2) the parametric approach presents a more "conservative" behavior. At most levels in fact, the extreme values are less dispersed and the central moments are lower than nonparametric results. At all levels, values that differ from the log-normal distribution implied by the B&S approach are evidence in favor of a nonparametric and non-normal distribution of returns, thus justifying the use of a data driven MF approach. For the monthly estimates, as K_t^α grows, the B&S VaR and CVaR estimates decrease and the distance between the two measures, $\Delta_{\alpha,t}$, increases.

Economically, the risk premium is what makes the option-implied and the statistically-based historical risk measures different. It is well-known that the risk premium is unobservable and very difficult to estimate (Damodaran (2016)). Indeed, investors are generally not risk-neutral and the ex-ante quantification of their risk preferences is not a trivial task to perform, as their subjective probabilities are unobservable quantities. It follows that the absence of a reliable methodology to estimate the risk premium is a big issue faced by any risk metric under the physical measure. Acting among the risk-neutral and the physical measures, the pricing kernel is the collector of the investors' risk preferences. These risk preferences become smaller the shorter the time horizon in consideration. By theory in fact, the pricing kernel converges to unity as time to expiration shrinks:

$$S = e^{rT} q \cdot 1 = e^{RT} p \cdot 1 \tag{21}$$

where S is the expected payoff of a contingent claim under the risk-neutral measure, q, and under the physical measure, p, respectively. If the no arbitrage assumption applies, the risk-neutral value grows at the risk-free rate, r, whereas the physical value grows at the risk-adjusted risk-free rate, R. As the time to maturity T approaches zero, both quantities converge to the same state price density, S; conversely, when T increases their divergence increases at the rate $(R - r)T$. The impossibility of directly observing both the physical measure and the risk premium makes the estimation of the pricing kernel rather problematic. If follows that, as long as T is not too big, adopting a model-free risk-neutral approach might be convenient. For example, the proposed option-implied risk measures do not require any numerical implementation or estimation of the risk premium to infer the future P&L. This might be an advantage above all in periods of turmoil, as the risk-neutral risk measures deliver more conservative (with respect to the statistically-based ones) risk metrics. This approach is followed also by Martin (2017), who shows that under the physical measure the probability of observing a crash, $\alpha_p \propto \left[\frac{\partial p_t}{\partial K} - \frac{p_t(K)}{\alpha F_t}\right]$, is always lower than the corresponding probability under the risk-neutral measure, $\alpha_q \propto \frac{\partial p_t}{\partial K}$. As a result, at any risk level the identified option under the physical measure has a higher strike compared to the one identified under the risk-neutral measure, thus making the VaR and CVaR estimates under the physical measure less conservative.

Finally, to test the performances of the option-implied risk measures we compare results of the option-implied VaR with a statistically-based VaR calibrated on an GJR-GARCH-FHS model. The choice of the model is to capture the clustering and the heteroskedastic feature shared by the market returns (GARCH), through an asymmetric autoregressive model (GJR) with empirical innovations (FHS). To capture these features, the GJR-GARCH-FHS model is fitted to the historical log-returns of the S&P 500 index. Starting from January 1, 2005 GARCH parameters are estimated using 10 years of daily observations, thus starting from January 1, 1995 for a total of 2279 daily observations. The Portmanteau test of Ljung and Box (1978) and the Lagrange Multiplier (MA) ARCH test of Engle (1984) present evidence that the proposed model and the number of observations used deliver robust results. Formally, the GJR-GARCH-FHS model is defined as:

$$\log\left(\frac{S_t}{S_{t-1}}\right) = r_t = a + \rho r_{t-1} + \epsilon_t \tag{22}$$

$$\epsilon_t = \sqrt{\sigma_t^2} z_t \tag{23}$$

$$z_t \sim f_t(0,1) \tag{24}$$

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha\epsilon_{t-1}^2 + \gamma\mathbb{I}_{t-1}\epsilon_{t-1}^2 \tag{25}$$

$$\mathbb{I}_{t-1} = \begin{cases} 1, & \text{if } \epsilon_{t-1} < 0, \\ 0 & \text{if } \epsilon_{t-1} \geq 0. \end{cases} \tag{26}$$

where z_t defines the empirical (FHS) innovation while Equation (25) model the returns volatility and Equation (26) accounts for the leverage effect (Black (1976)) through the indicator function. Table 5 collects summary statistics on the estimated physical parameters $\theta = (\omega, \alpha, \beta, \gamma)$ obtained by Pseudo Maximum Likelihood (PML). Then, using the set of estimated GARCH parameters and the empirical innovations, i log-returns are simulated:

$$\widehat{S}_{i,t} = S_t \exp\left(\mu - \frac{\widehat{\sigma}_t^2}{2}\right) dt + \widehat{\sigma}_t^2 dW_t \tag{27}$$

where S_t represents the current value of the S&P 500 Index, the drift input μ accounts for the market risk premium, $\hat{\sigma}_t^2$ is the simulated variance, dt is fixed at one day and dW_t models the randomness through the empirical innovations, z_t. As anticipated above, any estimation of the physical measure depends on the risk premium. In this paper, we follow Merton (1980) and much of the literature and we fix the risk premium at 8%.[14] From the estimated i log-returns, the future probability density functions (pdf) and cumulative density functions (cdf) are extracted through a kernel regressions.[15] Results show that, over the sample period in consideration, the distance

[14]The 8% of risk premium proposed by Merton (1980) might be too hight given the different market conditions. As such we repeated the same analysis but with a risk premium of 4% and 6%. Unreported results confirm the robustness of the approach and are available upon request to the authors.
[15]Extra details concerning the estimation in Barone-Adesi et al. (2008) and **?**.

between the statistically-based (GARCH) CVaR and VaR are more stable than the ones obtained with the option-implied approaches.

6 Backtesting

In this section we test the validity of our estimates performing different types of backtests. Backtesting refers to the validation of an estimated risk measure based on realized losses. The chosen backtesting measures for the VaR are: the number of times that an option expires in-the-money, the Binomial test (Bin), the Traffic Light (TL) test, the Proportion Of Failure (POF) test, the Time Until First Failure (TUFF) test, the Conditional Coverage Independence (CCI) test, the Conditional Coverage (CC) test, the Time Between Failure Independence (TBFI) test and the Time Between Failure (TBF) test.

The number of times that an option expires in-the-money is defined as:

$$\sum_{t=1}^{N} I_{[S_T < K_{t,T}^{\alpha}]} \tag{28}$$

$$\sum_{t=1}^{N} I_{[S_T > K_{t,T}^{\alpha}]} \tag{29}$$

where S_T is the price of the underlying asset at the maturity of the option contract (i.e. $T = 7, 30$ days for weekly and monthly estimates, respectively), $K_{t,T}^{\alpha}$ is the strike price of the option at the chosen level α, $I_{[\cdot]}$ is the indicator function, and N is the number of observation dates. Equations (28) and (29) refer to the left and right tail of the distribution, respectively.

The Binomial (Bin) test of Jorion (2011) compares the observed with the expected number of exceedances through the following test statistic:

$$\text{Bin} = \frac{I - Np}{\sqrt{Np(1-p)}} \tag{30}$$

where I represents the number of exceedances observed, p is a positive value between 0 and 1 and is defined as $p = 1 - \alpha$, and N is the total number of observations.

Proposed by the Basel Commitee, the Traffic Light (TL) test compares the number of observed and expected exceedances. Given a number of exceedances, I, this test computes and ranks the probability of observing up to I exceedances. While the Binomial test - as well as all other tests - only gives two mutually exclusive results, this test classifies risk into three regions: red, yellow and green (R, Y and G respectively). The red region is linked to a number of exceedances whose probability goes over 99.9%, the probability of the yellow are goes from 99.9% up to 95% while the green area refers to all values below the yellow area.

The POF and TUFF tests refer to the Proportion Of Failure and the Time Until First Failure tests of Kupiec (1995). Starting from the Binomial test and adding a likelihood ratio test, the POF test suggests that the VaR estimates can be rejected if the likelihood ratio exceeds a given critical value. Asymptotically distributed as a chi square with one degree of freedom, the POF test statistic is defined as:

$$\text{POF} = -2\text{Log}\left(\frac{(1-p)^{N-I}p^{I}}{\left(1-\frac{I}{N}\right)^{N-I}\left(\frac{I}{N}\right)^{I}}\right) \tag{31}$$

Still based on the likelihood ratio - but with a geometric distribution - the TUFF test checks when the first exceedance has occurred through the following statistic:

$$\text{TUFF} = -2\text{Log}\left(\frac{p(1-p)^{d-1}}{\left(\frac{1}{d}\right)\left(1-\frac{1}{d}\right)^{d-1}}\right) \tag{32}$$

where d accounts for the number of days until the first rejection. Also this statistic is asymptotically distributed as a Chi squared distribution with one degree of freedom.

The CCI and CC tests refer to the Conditional Coverage Independence and the Conditional Coverage of Christoffersen (1998). Both tests check if the observed failures are among them dependent between consecutive days. Asymptotically distributed as a χ^2 with one degree of freedom, the CCI test statistic is defined as:

$$\text{CCI} = -2\text{Log}\left(\frac{(1-\pi_2)^{d_{nn}+d_{yn}}\,\pi_2^{d_{ny}+d_{yy}}}{(1-\pi_0)^{d_{nn}}\pi_0^{d_{ny}}(1-\pi_1)^{d_{yn}}\pi_1^{d_{yy}}}\right) \tag{33}$$

where all powers account for the different combinations of d days with no failures, n, or with failures, y, thus producing four possible combinations. The probabilities are so defined:

- π_0: probability of having a time t failure given no $t-1$ failure: $\pi_0 = \frac{d_{ny}}{(d_{nn}+d_{ny})}$
- π_1: probability of having a time t failure given a $t-1$ failure: $\pi_1 = \frac{d_{yy}}{(d_{yn}+d_{yy})}$

- π_2: probability of having a time t failure: $\pi_2 = \frac{(d_{ny}+d_{yy})}{(d_{nn}+d_{ny}+d_{yn}+d_{yy})}$

The CC test is defined as:

$$\text{CC} = \text{POF} + \text{CCI} \tag{34}$$

and is asymptotically distributed as a chi square with two degrees of freedom.

The TBFI and TBF tests refer to the Time Between Failure Independence and the Time Between Failure of Haas (2001). While the TUFF checks only for the first exceedance, thus possibly leaving important information out of the analysis, the TBFI of Haas (2001) considers all exceedances. Asymptotically distributed as a chi square with I degrees of freedom, the TBFI test is defined as:

$$\text{TBFI} = -2\sum_{i=1}^{I} \text{Log}\left(\frac{p(1-p)^{d_i-1}}{\left(\frac{1}{d_i}\right)\left(1-\frac{1}{d_i}\right)^{d_i-1}}\right) \tag{35}$$

As a crucial modification with respect to the TUFF test, d is enriched with the subscript i, which accounts for the number of exceedances between i and $i-1$. As for the CC test, the POF can be used jointly with the TBFI test to define the TBF statistic:

$$\text{TBF} = \text{POF} + \text{TBFI} \tag{36}$$

which is asymptotically distributed as a chi square with $I+1$ degrees of freedom.

As concerns the CVaR, three backtests are performed. First, the average predicted excess loss beyond the VaR based on the left (Equation (37)) and right (Equation (38)) tail of the distribution:

$$L_{exp} = \frac{1}{N}\sum_{t=1}^{N}\left[\text{CVaR}_{t,T}^{\alpha} - \text{VaR}_{t,T}^{\alpha}\right] I_{[S_T<K_{t,T}^{\alpha}]} \tag{37}$$

$$L_{exp} = \frac{1}{N}\sum_{t=1}^{N}\left[\text{CVaR}_{t,T}^{\alpha} - \text{VaR}_{t,T}^{\alpha}\right] I_{[K_{t,T}^{\alpha}>S_T]} \tag{38}$$

is compared with the average realized loss conditional on the observation of a VaR exceedance:

$$L_{real} = \frac{1}{N}\sum_{t=1}^{N}\left[-(S_T-S_t) - \text{VaR}_{t,T}^{\alpha}\right] I_{[S_T<K_{t,T}^{\alpha}]} \tag{39}$$

When computing the expected excess loss, the choice of performing a conditional test has the objective to remove the effect of variation and estimation error in the VaR through time. The mean conditional expected excess loss is compared with the mean conditional realized loss exceeding the VaR by conducting a nonparametric Mann and Whitney (1947) test on the difference between the two means.[16] If the null hypothesis of a zero difference between the two means cannot be rejected, the option-implied CVaR forecasts future losses in a consistent way.

Second, following McNeil et al. (2015), the null of a correctly calibrated CVaR is tested against the alternative of a risk underestimation through the following function:

$$K_t = h_1^{(\alpha)}(\mathrm{VaR}_{t,T}^{\alpha}, \mathrm{CVaR}_{t,T}^{\alpha}, L_t) = \left(\frac{L_t - \mathrm{CVaR}_{t,T}^{\alpha}}{\mathrm{CVaR}_{t,T}^{\alpha}}\right) \cdot I_t \tag{40}$$

where L_t is the realized loss and I_t is the indicator function of a VaR exceedance. Under the null, the expected value of K_t conditional to the information up to time $(t-1)$ is equal to zero, $\mathbb{E}(K_t|\mathcal{F}_{t-1}) = 0$. Under the alternative is strictly positive, $\mathbb{E}(K_t|\mathcal{F}_{t-1}) > 0$, thus suggesting that the CVaR has a negative bias.

Third, since the VaR and the CVaR are jointly elicitable (see Lambert et al. (2008) and Fissler et al. (2015)), the joint calibration function:

$$S_t = h_2^{(\alpha)}(\mathrm{VaR}_{t,T}^{\alpha}, \mathrm{CVaR}_{t,T}^{\alpha}, L_t) = \frac{I_t}{\mathrm{CVaR}_{t,T}^{\alpha}} - \alpha \tag{41}$$

enables to test the null hypothesis of a correct joint calibration of the VaR and CVaR. Rejection of the null implies that the VaR and/or the CVaR underestimate risks. As for K_t, S_t is expected to behave as a martingale difference (i.e. $\mathbb{E}(K_t|\mathcal{F}_{t-1}) = 0$) under the null and to be strictly positive under the alternative (i.e.: $\mathbb{E}(K_t|\mathcal{F}_{t-1}) > 0$).

Table 6 presents basic results for the nine aforementioned VaR tests (Bin, TL, POF, TUFF, CCI, CC, TBFI and TBF) for the entire set of risk measures. Divided in weekly and monthly estimates, and for both tails of the distribution the table reports the acceptance/rejection results for the two option-implied risk measures (MF and BS) and the physical benchmark (FHS). For the option-implied risk measures the left (right) value refers to the (left) right tail of the distribution

[16] In risk management, forecasted losses smaller/larger than the realized ones are both undesirable. The former ones imply a risk underestimation and hence an inadequate capital protection, the latter ones induce an inefficient allocation of capital resources.

computed with Put (Call) options. For all tests: R means rejection and A acceptance. The only exception is the TL test which has three possible outcomes G (green), Y (yellow) and R (red).

As shown in the upper part of Table 6, the weekly MF VaR based on the left tail of the distribution performs very well in all tests and across all levels of risk (1%, 2.5% and 5%). This confirms an accurate risk evaluation of the MF option-implied risk measures for short investment horizons and at multiple risk levels. To further investigate over these results, Tables 7, 8 and 9 present detailed statistics on all proposed tests, for both tails of the distribution and for both short (weekly) and long (monthly) time horizons[17]. Table 7 details on the validity of the traffic light test for the MF option-implied VaR at all levels and for both short and long time horizons. Differently than in the previous table, here L (R) refers to the left (right) tail of the distribution. For the traffic light table, "Probability" is the cumulative probability of observing up to the corresponding number of failures, "Type I" is the probability of observing the corresponding number of failures or more if the model is correct and "Increase" is the eventual increase in the scaling factor. Both at short and long time-horizons the left tail risk measures perform well, confirming and extending the results of Table 6. While the left tail provides reliable risk forecasts, the right one might be suboptimal for the highest α level. At 1% and 2.5% in fact, the traffic light test is red and yellow, thus showing a discrepancy between expected and realized exceedances. A possible explanation for the suboptimal performances is the lower amount (with respect to Put options) of liquid weekly Call options[18]. Regarding the Bin, POF, TUFF and CC tests, Table 8 shows that for the left tail of the distribution and at all levels, the p-values are almost substantially larger that the threshold level of 5%, thus corroborating the statistical soundness of the MF estimates. Results confirm the lower ability of Call options in capturing the risk into the right tail of the distribution (compared with Put options for the left tail of the distribution).

To corroborate the results, Figure 3 presents the time series of left tail MF option-implied VaR plotting, with a black star, the exceedances. Graphically, for α=5% two exceedances occur at close dates at the end of 2014, thus potentially undermining the time independence among the VaR exceedances. Nevertheless, the detailed results of the independence (CCI) test (see first three rows of Table 9) show that the p-values (4^{th} column) are always larger than 5%. As a consequence, the

[17] While omitted, we also performed the analysis on the parametric B&S. All results are available upon request to the authors.

[18] One could rely on the Put-Call parity or use extrapolation techniques to alleviate this problem.

null of time-independent exceedances cannot be rejected at 95% confidence level. As it concerns the CVaR, the upper left part of Table 10 presents the backtesting results on the weekly MF CVaR based on the left tail of the distribution. Being conditional on the occurrence of a VaR exceedance, the p-values of the Mann-Whitney test (ρ_{MW}) and of the two tests proposed by McNeil (ρ_K and ρ_S) are not available at $\alpha = 1\%$, as no VaR exceedances occurred at the lowest risk level (see Figure 3). At $\alpha = 2.5\%$ and 5%, the results of all tests suggest that the MF CVaR estimates provide good forecasts of future risks.

On a monthly basis, Figure 4 depicts the monthly MF VaR exceedances for the left tail of the distribution. Results are supported by Tables 7, 8 and 9 which present detailed statistics on the tests proposed in Table 6 but at a monthly time horizon. Even taking into account the number of missing data and for both tails of the distribution, the proportion of exceedances is always below the corresponding α level (see the first three rows in the 1^{st} column of Table 10). At $\alpha = 5\%, 10\%, 15\%$, the proportion of exceedances equals $1.56\%, 4.68\%$ and 6.25% for the left tail and $2.34\%, 4.58\%, 6.25\%$ for the right one. The proposed VaR risk measures are then conservative forecasts of the potential losses. This result is also reflected in the p-values of the Binomial and POF tests (see the first six rows in the 4^{th} column of Table 8). The discrepancy between the proportion of exceedances and their corresponding theoretical values drives down the p-values of the Binomial and POF tests, whose null hypothesis are rejected when the risk level is 10%. Conversely, at $\alpha = 5\%$ and 15% the difference between the proportion of realized and theoretical exceedances is not statistically significant, and the null of correct coverage cannot be rejected at 95% confidence level. According to all the other tests and across each risk level, the monthly MF VaR correctly predicts risk. The lower left part of Table 10 presents backtesting results on the monthly MF CVaR based on the left tail of the distribution. When α equals 5% and 10%, the null hypothesis cannot be rejected at 99% confidence level. The estimated CVaR thus provide an adequate capital protection. Conversely, at $\alpha = 15\%$ the null is rejected. Being a two-tailed test, the MannWhitney U test does not provide information on whether the CVaR is under/over-estimating the actual losses. At $\alpha = 15\%$, the difference between the sample means of the expected and the realized losses is large and positive (i.e. 25.57$). This result implies that the option-implied CVaR is a conservative risk metric and is overestimating future losses at that confidence level. The lower left part of Table 10 presents backtesting results on the monthly MF CVaR based on the right tail of the distribution. For the right tail of the distribution, and at each risk level, the null hypothesis cannot be rejected at 95% confidence level, thus indicating that the CVaR correctly predicts future losses. When performing the bootstrap hypothesis tests on K_t and S_t, the p-values (see 4^{th} and 5^{th} columns) ρ_K and ρ_S, are always sufficiently high to fail to reject the null hypothesis at 95% confidence level. In other words, for both tails of the distribution there is no evidence that the CVaR alone and the CVaR and VaR jointly are affected by negative biases. The right upper part of Table 10 reports backtesting results of the weekly B&S CVaR based on the left and right tails. At $\alpha = 1\%$ and 2.5%, no exceedances occur; hence, the conditional CVaR backtests deliver no results. At $\alpha = 5\%$, the proportion of the exceedances (0.63%) is much lower than its theoretical value (5%). The expected excess losses are substantially lower than the realized ones, thus suggesting that the B&S CVaR may underestimate risk. Nevertheless, being this conclusion based on just one VaR exceedance, the statistical power is not sufficiently high to assess that in general the weekly B&S CVaR forecasts risks in an inadequate way.

The right lower part of Table 10 reports the same backtesting statistics for the B&S CVaR and for both tails of the distribution at a monthly and weekly investment horizon. For $\alpha = 5\%$ and 10%, the proportion of exceedances (upper part of 6^{th} column) is always lower under the B&S assumption than the MF approach. At all risk level, the proportion of VaR exceedances (middle part of 6^{th} column) under the B&S approach is always larger. The B&S methodology then delivers less conservative risk forecasts. The MannWhitney U test based on the left tail of the distribution always rejects the null hypothesis at 95% confidence level. Compared to the MF methodology and for $\alpha = 15\%$, the difference between the expected and the realized losses is close enough to zero (i.e. 2.40$) to fail to reject the null. Based on the right tail of the distribution, the null of a correctly specified risk metric cannot be rejected at any risk level. As it concerns the McNeil tests, ρ_K and ρ_S (see 9^{th} and 10^{th} columns) are substantially lower than those obtained under the MF methodology. Nevertheless, for both tails of the distribution at 95% confidence level, the p-values suggest that the B&S CVaR and the B&S VaR and CVaR jointly do not underestimate risk.

To conclude the analysis we present backtesting results for the statistically-based measures estimated with the GJR-GARCH-FHS model presented in Section 5 The lower part of Table 6 presents backtesting results for the weekly (top) and monthly (bottom) VaR. Confirming the previous findings, the obtained results suggest that the FHS CVaR alone and the FHS VaR and CVaR jointly do not underestimate risk.. Confirming Brownlees et al. (2011) the statistical-based VaR works well at short time horizon and for almost all tests in consideration. There are in fact no rejections at the weekly horizon. Things change at longer time horizons. Monthly VaR are rejected at 5% for the POF test and at 10% for the TBF. As it concerns the VaR, at $\alpha = 5\%$ and 15%, the proportion of the exceedances (see the 1^{st} column of Table 6) is slightly higher compared to the VaR backtesting results in the left tail under the option-implied methodology, both when using the MF and the B&S approaches (see 1^{st} and 6^{th} columns of Table 10). The VaR risk forecasts are rejected in just two cases (at $\alpha = 5\%$ the null of the POF test is rejected and at $\alpha = 10\%$ the null of the TBF test is rejected).

7 Conclusion

This article presents and compares backtesting results at short and long time horizons for the S&P 500 Index VaR and CVaR estimated under two different methodologies, namely the option-implied and the statistically-based approaches. Option-implied risk measures are naturally forward-looking, economically grounded and react quickly to new market scenarios. Statistically-based risk measures are estimated using a GJR-GARCH-FHS model. At short time horizons, both the option-implied and the statistically-based risk measures perform similarly. At long time horizons, the option-implied risk measures perform well while the statistically-based risk metrics are not always reliable. Option-implied risk measures can thus be seen as possible alternatives to statistically-based risk measures. Option-implied risk measures can be useful to regulators, single companies and investors. Regulators can quantify the risk of a firm without needing to know what is on (and off) its balance sheet. Firms can either gain an overview of their perception among the investors or derive market-based risk metrics to be compared with those obtained by internal models. Finally, investors can evaluate the risk of their existing/potential investments.

8 Bibliography

Acerbi, C. and B. Székely (2014). Backtesting expected shortfall. *Risk*, 76.

Äit-Sahalia, Y. and A.W.Lo (2000). Non parametric risk management and implied risk aversion. *Journal of Econometrics 94*, 9–51.

Arrow, J. and G. Debreu (1954). Existence of an equilibrium for a competitive economy. *Econometrica 22*(3), 265–290.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999). Coherent measures of risks. *Mathematical Finance 9*, 203–228.

Bahara, B. (1997). Implied risk-neutral probability density functions from options prices: Theory and application. *Bank of England ISSN 1368-5562*.

Bali, T., N. Cakici, and F. Chabi-Yo (2011). A generalized measure of riskiness. *Management Science 57*(8), 1406–1423.

Barone-Adesi, G. (2016). Var and cvar implied in option prices. *Journal of Risk and Financial Management 9*(1).

Barone-Adesi, G. and R. Elliot (2007). Cutting the hedge. *Computational Economics 29*(2), 151–158.

Barone-Adesi, G., R. Engle, and L. Mancini (2008). A GARCH option pricing model with Filtered Historical Simulationgarch option pricing model in incomplete markets. *Review of Financial Studies 21*, 1223–1258.

Barone-Adesi, G., K. Giannopoulos, and L. Vosper (1999). Backtesting derivative portfolios with Filtered Historical Simulation (FHS). *Journal of future market 8*(1), 31–58.

Barone-Adesi, G., C. Legnazzi, M. Finta, and C. Sala (2019). Wti crude oil option implied var and cvar: an empirical application. *Journal of Forecasting*.

Black, F. (1976). Studies of stock price volatility changes. *Proceedings of the 1976 Meetings of the American Statistical Association*, 171–181.

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy 81*(3), 637–654.

Bliss, R. and N. Panigirtzoglou (2004). Option-implied risk aversion estimates. *Journal of Finance 59*, 407–446.

Bogle, J. (2005). The mutual fund industry 60 years later: For better or worse? *Financial Analyst Journal 61*, 15–24.

Breeden, D. and R. Litzenberger (1978, October). Prices of state-contingent claims implicit in option prices. *The Journal of Business 51*(4), 621–651.

Brownlees, C., R. Engle, and B. Kelly (2011). A practical guide to volatility forecasting through calm and storm. *The Journal of Risk 14*(2), 3–22.

Brunnermeier, M. and L. H. Pedersen (2009). Market liquidity and funding liquidity. *The Review of Financial Studies 22*(6), 2201–2238.

Carr, P., H. Geman, and D. Madan (2001). Pricing and hedging in incomplete markets. *Journal of Financial Economics 62*, 131–168.

Christoffersen, P. (1998). Evaluating interval forecasts. *International Economic Review 39*, 841–862.

Damodaran, A. (2016). Equity risk premiums: Determinants, estimation and implications- the 2016 edition.

Danielsson, J. (2002). The emperor has no clothes: Limits to risk modelling. *Journal of Banking and Finance 26*(7), 1273–1296.

Dowd, K. (1999). *Beyond Value at Risk: The Science of Risk Management.* Wiley, New York.

Duffie, D. and J. Pan (1997). An overview of value at risk. *Journal of Derivatives 4*(3), 7–49.

Engle, R. (2009). The risk that risk will change. *Journal of Investment Management 7*, 24–28.

Engle, R. (2011). Long-term skewness and systemic risk. *Journal of Financial Econometrics 9*(3), 437–468.

Engle, R. and V. Rosenberg (2002). Empirical pricing kernels. *Review of Financial Studies 64*(341-372).

Engle, R. F. (1984). Wald, likelihood ratio, and lagrange multiplier tests in econometrics. *Handbook of econometrics 2*, 775–826.

Figlewski, S. (2008). *Estimating the Implied Risk Neutral Density for the U.S. Market Portfolio*, Volume Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle (eds. Tim Bollerslev, Jeffrey R. Russell and Mark Watson). Oxford, UK: Oxford University Press.

Fissler, T., J. F. Ziegel, and T. Gneiting (2015). Expected shortfall is jointly elicitable with value at risk-implications for backtesting. *arXiv preprint arXiv:1507.00244*.

Glosten, L., R. Jagannathan, and D. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance 48*(5), 1779–1801.

Gneiting, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association 106*(494), 746–762.

Haas, M. (2001.). New methods in backtesting. *Financial Engineering Research Center Caesar, Bonn,*.

Huggenberger, M., C. Zhang, and T. Zhou (2018). Forward-looking tail risk measures. *Working paper*.

Jorion, P. (2007). *Value at Risk - The New Benchmark for Managing Financial Risk*, Volume 3rd Ed. The McGraw-Hill Companies.

Jorion, P. (2011). *Financial Risk Manager Handbook*. Number 6th Edition. Wiley Finance.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk management models. *Journal of Derivatives 3*, 73–84.

Lambert, N. S., D. M. Pennock, and Y. Shoham (2008). Eliciting properties of probability distributions. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, pp. 129–138. ACM.

Ljung, G. and G. Box (1978). On a measure of lack of fit in time series of models. *Biometrika 65*, 297–303.

Mann, M. B. and D. R. Whitney (1947). On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, 50–60.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance 7*(1), 77–91.

Martin, I. (2017). What is the expected return on the market? *The Quarterly Journal of Economics 132*(1), 367–433.

McNeil, A. J., R. Frey, and P. Embrechts (2015). *Quantitative risk management: Concepts, techniques and tools*. Princeton university press.

Merton, R. (1980). On estimating the expected return on the market. *Journal of Financial Economics 8*, 323–361.

Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of economics and management science*, 141–183.

Mitra, S. (2015). The relationship between conditional value at risk and option prices with a closed form solution. *The European Journal of Finance 21*(5), 400–425.

Penza, P. and V. Bansal (2000). *Measuring Market Risk with Value at Risk*. Wiley, New York.

Samit, A. (2012). Calculating value-at-risk using option implied probability distribution of asset price. *Wilmott Magazine 2012*(59), 56–61.

Xing, Y., X. Zhang, and R. Zhao (2010). What does the individual option volatility smirk tell us about future equity returns? *Journal of Financial Quantitative Analysis 45*(3), 641–662.

Yan, S. (2011). Jump risk, stock returns, and slope of implied volatility smile. *Journal of Financial Economics 99*, 216–233.

Ziegel, J. F. (2014). Coherence and elicitability. *Mathematical Finance 26*(4), 901–918.

9 Tables

Weekly CVar and VaR estimates										
Model	Left tail					Right tail				
MF CVaR	Mean	Median	Max	Min	St.Dev	Mean	Median	Max	Min	St.Dev
$\alpha = 5\%$	95.38	87.31	411.21	52.26	42.41	52.27	50.22	181.23	34.30	15.91
$\alpha = 2.5\%$	141.07	120.87	755.90	73.71	77.79	60.86	57.34	225.54	36.21	19.63
$\alpha = 1\%$	229.32	202.17	863.21	98.16	105.08	79.20	72.11	711.79	44.21	53.90
MF VaR	Mean	Median	Max	Min	St.Dev	Mean	Median	Max	Min	St.Dev
$\alpha = 5\%$	62.89	56.14	300.89	32.61	33.10	39.81	37.25	126.79	24.30	12.86
$\alpha = 2.5\%$	99.62	84.20	395.89	45.98	73.31	46.24	43.08	171.79	25.21	16.10
$\alpha = 1\%$	193.41	155.74	478.87	65.98	95.84	53.71	51.63	114.11	30.21	13.80
BS CVaR	Mean	Median	Max	Min	St.Dev	Mean	Median	Max	Min	St.Dev
$\alpha = 5\%$	106.20	96.93	469.69	60.68	46.77	84.06	80.39	241.06	56.15	20.30
$\alpha = 2.5\%$	164.46	153.48	630.16	91.73	63.68	83.43	79.52	240.95	56.11	20.13
$\alpha = 1\%$	250.07	239.16	627.50	136.42	71.74	83.91	80.30	236.80	56.38	20.08
BS VaR	Mean	Median	Max	Min	St.Dev	Mean	Median	Max	Min	St.Dev
$\alpha = 5\%$	85.94	77.93	393.21	48.61	39.21	45.49	43.08	156.79	30.21	13.29
$\alpha = 2.5\%$	141.53	132.11	543.21	78.16	55.69	54.90	52.23	171.79	35.21	14.47
$\alpha = 1\%$	223.77	213.48	570.89	121.47	65.33	75.47	72.19	206.79	51.30	17.30
Monthly CVar and VaR estimates										
Model	Left tail					Right tail				
MF CVaR	Mean	Median	Max	Min	St.Dev	Mean	Median	Max	Min	St.Dev
$\alpha = 15\%$	127.40	121.49	297.56	60.61	43.39	85.07	79.01	237.94	46.37	30.16
$\alpha = 10\%$	152.99	146.70	340.61	67.16	53.67	92.55	86.41	259.83	52.64	32.27
$\alpha = 5\%$	210.48	206.31	461.43	96.24	73.70	108.45	102.06	331.31	61.00	39.33
MF VaR	Mean	Median	Max	Min	St.Dev	Mean	Median	Max	Min	St.Dev
$\alpha = 15\%$	67.91	64.66	192.44	26.67	26.96	61.07	55.42	172.56	32.20	22.81
$\alpha = 10\%$	95.94	91.92	253.11	36.96	41.22	68.89	63.81	171.89	34.04	22.69
$\alpha = 5\%$	155.14	145.16	450.26	61.64	67.23	84.00	78.10	217.56	44.25	29.18
BS CVaR	Mean	Median	Max	Min	St.Dev	Mean	Median	Max	Min	St.Dev
$\alpha = 15\%$	130.19	125.22	300.89	56.34	46.42	84.34	77.32	228.76	49.37	28.51
$\alpha = 10\%$	159.77	154.48	358.41	70.43	54.78	96.54	91.35	261.76	56.99	31.35
$\alpha = 5\%$	217.99	213.86	489.25	100.56	68.90	126.46	120.92	333.31	71.37	38.46
BS VaR	Mean	Median	Max	Min	St.Dev	Mean	Median	Max	Min	St.Dev
$\alpha = 15\%$	86.80	83.34	223.11	33.12	34.75	67.93	62.79	167.56	41.88	20.69
$\alpha = 10\%$	117.00	113.36	288.11	48.12	44.05	84.11	79.83	212.56	51.88	25.14
$\alpha = 5\%$	177.63	173.95	407.18	78.12	59.42	117.87	113.01	297.56	67.20	34.50

Table 1: Summary statistics of all option-implied risk measures for both short (weekly) and long (monthly) time horizons and for both tails of the distribution. MF (BS) CVaR refers to the Model Free (Black & Scholes) CVaR computed with nonparametric (parametric) α. The same holds for the VaR. Weekly estimations use weekly options while the monthly estimations use monthly options.

Summary statistics for Δ_α=CVaR$_\alpha$-VaR$_\alpha$.

Weekly Estimations								
Left Tail	Model-Free				Black and Scholes			
CVaR-VaR	Min	Max	Mean	Median	Min	Max	Mean	Median
$\alpha = 1\%$	2.00	122.51	38.13	36.66	14.95	56.61	26.30	25.23
$\alpha = 2.5\%$	3.00	140.00	41.44	38.00	13.57	86.95	22.92	21.44
$\alpha = 5\%$	14.37	123.00	32.49	30.00	12.40	76.48	20.26	18.67
Right Tail	Model-Free				Black and Scholes			
CVaR-VaR	Min	Max	Mean	Median	Min	Max	Mean	Median
$\alpha = 1\%$	7.50	610.02	25.48	20.00	5.03	30.01	8.44	7.97
$\alpha = 2.5\%$	6.66	53.75	14.61	13.75	11.56	69.16	28.53	27.16
$\alpha = 5\%$	6.66	54.44	12.46	11.70	21.26	84.27	38.57	36.54
Monthly Estimations								
Left Tail	Model-Free				Black and Scholes			
CVaR-VaR	Min	Max	Mean	Median	Min	Max	Mean	Median
$\alpha = 5\%$	2.36	150	55.34	49.42	22.44	82.07	40.35	40.24
$\alpha = 10\%$	27.35	142.35	56.08	54.55	22.18	84.99	41.65	41.17
$\alpha = 15\%$	28.51	129.45	59.48	49.43	23.22	88.03	43.14	42.66
Right Tail	Model-Free				Black and Scholes			
CVaR-VaR	Min	Max	Mean	Median	Min	Max	Mean	Median
$\alpha = 5\%$	11.94	114.24	24.78	22.51	10.07	75.80	20.84	19.05
$\alpha = 10\%$	11.53	117.68	23.67	21.36	11.79	83.53	24.01	21.58
$\alpha = 15\%$	12.31	59.94	24.23	23.03	13.22	90.35	26.94	24.48

Table 2: Summary statistics of Δ_α=CVaR$_\alpha$-VaR$_\alpha$ for both short (weekly) and long (monthly) time horizons and for both tails of the distribution. MF(BS) CVaR refers to the Model Free (Black & Scholes) CVaR computed with nonparametric (parametric) α. The same holds for the VaR. Weekly estimations use weekly options while the monthly estimations use monthly options.

Number of missing observations for the weekly and monthly estimates.

	Weekly			Monthly		
Model-Free approach	$\alpha = 1\%$	$\alpha = 2.5\%$	$\alpha = 5\%$	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 15\%$
% missing observations (Left)	1.26	0.00	0.00	6.25	9.38	13.28
% missing observations (Right)	1.88	1.25	18.23	9.38	7.81	12.50
Black and Scholes approach	$\alpha = 1\%$	$\alpha = 2.5\%$	$\alpha = 5\%$	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 15\%$
% missing observations (Left)	0.00	0.00	0.00	0.78	0.00	3.13
% missing observations (Right)	0.00	0.00	2.52	1.56	1.56	4.69

Table 3: Proportion of missing observations for the weekly and monthly estimates extracted with the model free methodology (upper part) and the Black and Scholes approach (lower part) for both the left and right tails of the distribution.

Left and right tail - Summary statistics for $\Gamma_t = K_t^{\alpha^{\mathbf{MF}}} - K_t^{\alpha^{\mathbf{BS}}}$.

Γ_t	Mean	Median	Min	Max
Left tail (W)	-0.0156	-0.0072	-0.2026	0.2225
Right tail (W)	0.0201	0.0113	-0.1790	0.2513
Left tail (M)	-0.0035	-0.0096	-0.3390	0.2430
Right tail (M)	0.0284	0.0229	-0.2027	0.2481

Table 4: Summary statistics for $\Gamma_t = K_t^{\alpha^{\mathrm{MF}}} - K_t^{\alpha^{\mathrm{BS}}}$ where α_t^{MF} represents the time t Model Free estimations and α_t^{BS} the relative Black and Scholes based estimations. The upper/lower part (W)/(M) refers to the weekly/monthly estimations.

2005-2015 S&P 500 Index GJR-GARCH-FHS Parameters

Year	ω	α	β	γ	Persistence
2005	2.43e-06	2.20e-13	0.91	0.15	0.98
2006	1.56e-06	1.16e-13	0.92	0.15	0.99
2007	1.16e-06	3.35e-10	0.93	0.12	0.99
2008	1.24e-06	1.24e-13	0.93	0.12	0.99
2009	1.21e-06	1.37e-13	0.93	0.13	0.99
2010	1.15e-06	1.24e-13	0.93	0.12	0.99
2011	1.25e-06	1.10e-13	0.93	0.12	0.99
2012	1.68e-06	2.45e-10	0.91	0.14	0.98
2013	1.85e-06	9.54e-14	0.91	0.15	0.98
2014	2.08e-06	1.43e-13	0.90	0.17	0.98
2015	2.37e-06	2.45e-13	0.90	0.15	0.98

Table 5: Yearly means of the physical GJR-GARCH-FHS parameters $\theta_t = f(\omega, \alpha, \beta, \gamma)$ calibrated each third Thursday of the month - from January 20, 2005 to December 17, 2015 - on the S&P 500 Index log returns using the Pseudo Maximum Likelihood (PML) approach and 10 years of daily log-returns (up to 1995). The GJR-GARCH-FHS model under the physical measure and the persistence are defined as:

$$\begin{aligned}
\log\left(\frac{S_t}{S_{t-1}}\right) &= r_t = a + \rho r_{t-1} + \epsilon_t \\
\epsilon_t &= \sqrt{\sigma_t^2} z_t \\
\sigma_t^2 &= \omega + \beta\sigma_{t-1}^2 + \alpha\epsilon_{t-1}^2 + \gamma \mathbb{1}[\epsilon_{t-1} < 0]\epsilon_{t-1}^2 \\
\text{Persistence} &= \alpha + \beta + \frac{\gamma}{2}
\end{aligned}$$

Backtesting results

Weekly VaR estimates									
Model	α	Bin	TL	POF	TUFF	CCI	CC	TBFI	TBF
MF	1%	A/R	G/R	A/R	A/A	A/A	A/R	A/R	A/A
MF	2.5%	A/R	G/Y	A/R	A/A	A/A	A/R	A/A	A/A
MF	5%	A/A	G/G	A/A	A/A	A/A	A/A	A/A	A/A
Model	α	Bin	TL	POF	TUFF	CCI	CC	TBFI	TBF
BS	1%	A/A	G/G	A/A	A/A	A/A	A/A	A/A	A/A
BS	2.5%	R/R	G/G	A/R	A/A	A/A	A/R	A/A	A/A
BS	5%	R/R	G/G	A/R	A/A	A/A	R/R	A/A	A/A
Monthly VaR estimates									
Model	α	Bin	TL	POF	TUFF	CCI	CC	TBFI	TBF
MF	5%	A/A	G/G	A/A	A/A	A/A	A/A	A/A	A/A
MF	10%	R/A	G/G	R/A	A/A	A/A	A/A	A/A	A/A
MF	15%	A/A	G/G	A/A	A/A	A/A	A/A	A/A	A/A
Model	α	Bin	TL	POF	TUFF	CCI	CC	TBFI	TBF
BS	5%	A/A	G/G	R/R	A/A	A/A	R/A	R/A	R/R
BS	10%	A/A	G/G	A/A	A/A	A/A	A/A	A/A	A/A
BS	15%	A/A	G/G	A/A	A/A	R/R	R/R	A/A	A/A
Weekly VaR estimates									
Model	α	Bin	TL	POF	TUFF	CCI	CC	TBFI	TBF
FHS	1%	A	G	A	A	A	A	A	A
FHS	2.5%	A	G	A	A	A	A	A	A
FHS	5%	A	G	A	A	A	A	A	A
Monthly VaR estimates									
FHS	5%	A	G	R	A	A	A	A	A
FHS	10%	A	G	A	A	A	A	A	R
FHS	15%	A	G	A	A	A	A	A	A

Table 6: Backtesting results of the Model Free (MF), Black and Scholes (BS) and Filtered Historical Simulation (FHS) VaR computed for the left (right) tail using Put (Call) options. As presented in the test, we only report backtesting results for the left tail FHS. For each section, the upper (lower) part refers to the short (long)-time horizon estimations. On a weekly basis (upper part), the backtesting results are reported for the left tail only; whereas on a monthly basis (lower part), the first value refers to the left tail and the second one to the right one. The first two columns refer to the model in use and the relative confidence level. The third column, Bin, refers to the Binomial test of Jorion (2011). The fourth column, TL, refers to the Traffic Light test proposed by the Basel Commitee (1996). The fifth and sixth columns, POF and TUFF, refer to the Proportion Of Failure and Time Until First Failure tests of Kupiec (1995).
The seventh and eight columns, CCI and CC, refer to the Conditional Coverage Independence and Conditional Coverage of Christoffersen (1998).
The ninth and tenth columns, TBFI and TBF refer to the Time Between Failure Independence and Time Between Failure of Haas (2001). As described in Section 6, aside for the TL that has 3 outcomes (green, yellow and red, or G, Y, R), all other test are binary and produces either acceptances, A, or rejections, R.

Weekly VaR estimates				
MF VaR	Traffic Light	Probability	Type I	Increase
$\alpha = 5\%$ (L)	Green	0.0981	0.95	0.00
$\alpha = 2.5\%$ (L)	Green	0.2473	0.91	0.00
$\alpha = 1\%$ (L)	Green	0.2112	1.00	0.00
MF VaR	Traffic Light	Probability	Type I	Increase
$\alpha = 5\%$ (R)	Green	0.8872	0.1978	0
$\alpha = 2.5\%$ (R)	Yellow	0.9972	0.0062	0.8652
$\alpha = 1\%$ (R)	Red	1	4.39e-06	1
Monthly VaR estimates				
MF VaR	Traffic Light	Probability	Type I	Increase
$\alpha = 5\%$ (L)	Green	0.0666	0.9815	0
$\alpha = 10\%$ (L)	Green	0.0195	0.9929	0
$\alpha = 15\%$ (L)	Green	0.0603	0.9686	0
MF VaR	Traffic Light	Probability	Type I	Increase
$\alpha = 5\%$ (R)	Green	0.6719	0.4907	0
$\alpha = 10\%$ (R)	Green	0.2859	0.8157	0
$\alpha = 15\%$ (R)	Green	0.0559	0.9711	0

Table 7: Detailed backtesting results of the Traffic Light test for the MF VaR based on the left (L) and right (R) tail of the distribution at weekly and monthly time horizon. Probability is the cumulative probability of observing up to the corresponding number of failures, Type I is the probability of observing the corresponding number of failures or more if the model is correct and Increase is the eventual increase in the scaling factor

Weekly VaR estimates						
	Left tail			Right tail		
MF VaR	Bin	Z Score	P Value	Bin	Z Score	P Value
$\alpha = 5\%$	Accept	-1.4236	0.077281	Accept	1.0301	0.1515
$\alpha = 2.5\%$	Accept	-0.9936	0.1602	Reject	3.1282	0.0008
$\alpha = 1\%$	Accept	-1.2553	0.1047	Reject	6.8214	4.50e-12
MF VaR	POF	Likelihood Ratio	P Value	POF	Likelihood Ratio	P Value
$\alpha = 5\%$	Accept	2.4559	0.11708	Accept	0.9501	0.3297
$\alpha = 2.5\%$	Accept	1.2023	0.27286	Reject	6.8802	0.0087
$\alpha = 1\%$	Accept	3.1357	0.0766	Reject	20.8610	4.93e-06
MF VaR	TUFF	Likelihood Ratio	P Value	TUFF	Likelihood Ratio	P Value
$\alpha = 5\%$	Accept	3.5780	0.0585	Accept	0.0563	0.8124
$\alpha = 2.5\%$	Accept	1.8368	0.1753	Accept	0.0096	0.9219
$\alpha = 1\%$	Accept	–	–	Accept	0.5292	0.4669
MF VaR	CC	Likelihood Ratio	P Value	CC	Likelihood Ratio	P Value
$\alpha = 5\%$	Accept	2.6651	0.2638	Accept	2.3127	0.3146
$\alpha = 2.5\%$	Accept	1.2539	0.5342	Reject	8.2606	0.0161
$\alpha = 1\%$	Accept	3.1357	0.2085	Reject	22.2510	1.47e-05
Monthly VaR estimates						
	Left tail			Right tail		
MF VaR	Bin	Z Score	P Value	Bin	Z Score	P Value
$\alpha = 5\%$	Accept	-1.6189	0.05274	Accept	0.1734	0.4312
$\alpha = 10\%$	Reject	-2.0647	0.01947	Accept	-0.7493	0.2268
$\alpha = 15\%$	Accept	-1.638	0.05071	Accept	-1.6708	0.0474
MF VaR	POF	Likelihood Ratio	P Value	POF	Likelihood Ratio	P Value
$\alpha = 5\%$	Accept	3.4707	0.0624	Accept	0.0294	0.8638
$\alpha = 10\%$	Reject	5.3160	0.0211	Accept	6.6007	0.4338
$\alpha = 15\%$	Accept	3.0313	0.0816	Accept	3.1611	0.0745
MF VaR	TUFF	Likelihood Ratio	P Value	TUFF	Likelihood Ratio	P Value
$\alpha = 5\%$	Accept	1.6679	0.1965	Accept	0.1045	0.7464
$\alpha = 10\%$	Accept	0.2846	0.59367	Accept	0.2074	0.6488
$\alpha = 15\%$	Accept	0.6252	0.4291	Accept	0.6438	0.4223
MF VaR	CC	Likelihood Ratio	P Value	CC	Likelihood Ratio	P Value
$\alpha = 5\%$	Accept	3.5415	0.1702	Accept	1.1023	0.5763
$\alpha = 10\%$	Accept	5.7666	0.0559	Reject	6.0712	0.0480
$\alpha = 15\%$	Accept	4.1933	0.1228	Accept	3.1666	0.2053

Table 8: Detailed backtesting results of the Bin, TUFF, CC an POF tests for the MF VaR based on the left and right tail of the distribution at short (weekly) and long (monthly) time horizon.

Weekly VaR estimates														
	Left tail							Right tail						
MF VaR	CCI	LR	P Value	N00	N10	N01	N11	CCI	LR	P Value	N00	N10	N01	N11
$\alpha = 5\%$	Accept	0.2091	0.6474	149	4	4	0	Accept	1.3626	0.2431	110	9	9	0
$\alpha = 2.5\%$	Accept	0.0516	0.8203	158	2	2	0	Accept	1.3804	0.2401	135	10	10	0
$\alpha = 1\%$	Accept	0	1	155	0	0	0	Accept	1.39	0.2384	134	10	10	0
MF VaR	TBFI	LR	P Value	Min/Max	Q1	Q2	Q3	TBFI	LR	P Value	Min/Max	Q1	Q2	Q3
$\alpha = 5\%$	Accept	6.7574	0.1493	3/83	18	35	60	Accept	9.4066	0.4006	2/44	6	7	19
$\alpha = 2.5\%$	Accept	1.8478	0.39697	36/120	36	78	120	Accept	16.7070	0.0811	3/44	3	10	21
$\alpha = 1\%$	Accept	–	–	–/–	–	–	–	Reject	30.0900	0.0008	3/44	4	10	20
Monthly VaR estimates														
	Left tail							Right tail						
MF VaR	CCI	LR	P Value	N00	N10	N01	N11	CCI	LR	P Value	N00	N10	N01	N11
$\alpha = 5\%$	Accept	0.0708	0.7902	111	2	2	0	Accept	1.0729	0.3003	100	5	5	1
$\alpha = 10\%$	Accept	0.4506	0.5020	106	5	5	0	Reject	5.4705	0.0193	98	6	6	3
$\alpha = 15\%$	Accept	1.1620	0.2811	88	8	8	2	Accept	0.0054	0.9410	88	9	9	1
MF VaR	TBFI	LR	P Value	Min/Max	Q1	Q2	Q3	TBFI	LR	P Value	Min/Max	Q1	Q2	Q3
$\alpha = 5\%$	Accept	1.8394	0.3986	13/57	13	35	57	Accept	10.1490	0.1185	1/27	2	11	17
$\alpha = 10\%$	Accept	0.9089	0.9695	12/18	12	13	16.5	Accept	16.85	0.0511	1/19	1	7	13.5
$\alpha = 15\%$	Accept	12.4840	0.254	1/17	2	4	10	Accept	8.8042	0.5507	10/19	3	6	10

Table 9: Detailed backtesting results of the CCI and TBFI tests for the MF VaR based on the left and right tail of the distribution at short (weekly) and long (monthly) time horizon. N00, N10, N01, N11 are explained in Section 6

Left tail	Weekly Model-Free					Weekly Black and Scholes				
	% I_t	L_{exp}	L_{real}	ρ_K	ρ_S	% I_t	L_{exp}	L_{real}	ρ_K	ρ_S
α=1%	0.00	–	–	–	–	0.00	–	–	–	–
α=2.5%	1.26	26.77	10.89	0.88	1.00	0.00	–	–	–	–
α=5%	2.52	29.61	21.16	0.85	1.00	0.63	19.46	44.11	0.10	1.00
Right tail	Weekly Model-Free					Weekly Black and Scholes				
	% I_t	L_{exp}	L_{real}	ρ_K	ρ_S	% I_t	L_{exp}	L_{real}	ρ_K	ρ_S
α=1%	6.29	23.37	16.54	0.66	0.01	1.26	6.12	29.01	0.13	0.23
α=2.5%	6.29	12.51	24.04	0.08	0.03	5.03	26.33	17.09	0.85	0.12
α=5%	5.66	10.67	21.64	0.03	0.18	6.92	39.91	19.79	0.98	0.41
Left Tail	Monthly Model-Free					Monthly Black and Scholes				
	% I_t	L_{exp}	L_{real}	ρ_K	ρ_S	% I_t	L_{exp}	L_{real}	ρ_K	ρ_S
$\alpha = 5\%$	1.56	43.22	46.26	0.35	1.00	0.78	38.84	46.81	0.11	1.00
$\alpha = 10\%$	4.68	50.90	46.77	0.43	1.00	2.34	44.32	52.38	0.14	1.00
$\alpha = 15\%$	6.25	55.99	30.42	0.92	1.00	7.03	43.75	41.34	0.41	1.00
Right Tail	Monthly Model-Free					Monthly Black and Scholes				
	% I_t	L_{exp}	L_{real}	ρ_K	ρ_S	% I_t	L_{exp}	L_{real}	ρ_K	ρ_S
$\alpha = 5\%$	2.34	30.07	55.96	0.87	1.00	3.91	22.30	26.11	0.12	0.72
$\alpha = 10\%$	4.68	26.92	41.73	0.96	1.00	5.46	24.73	57.14	0.15	1.00
$\alpha = 15\%$	6.25	23.09	37.79	0.88	1.00	7.81	28.93	55.86	0.09	0.96

Table 10: The upper portion of the table reports the backtesting results of the weekly option-implied VaR and CVaR under the MF and B&S methodology based on the left and right tail of the distribution. The lower portion of the table reports the backtesting results of the monthly option-implied VaR and CVaR under the MF and B&S methodology based on the left and right tail of the distribution.

The first column reports the proportion of MF VaR exceedances, I_t, which equals the number of exceedances, computed as in Equations (28)-(29), divided by the total number of observations, respectively.

The second and the third columns report the average expected, L_{exp}, and realized, L_{real}, excess losses conditional on a VaR exceedance expressed in \$ under the MF and B&S approach, respectively. The average expected loss is computed as the difference between the option-implied CVaR and VaR, conditional on a VaR exceedance, $L_{exp} = \frac{1}{N}\sum_{t=1}^{N}\left[\text{CVaR}_{\alpha,t} - \text{VaR}_{\alpha,t}\right] I_{[S_T < K^{\alpha}_{t,T}]}$, where $\mathbb{I}_t$ is the indicator function of a VaR exceedance. The average excess realized loss is defined as $L_{real} = \frac{1}{N}\sum_{t=1}^{N}\left[-(S_T - S_t) - \text{VaR}_{\alpha,t}\right] I_{[S_T < K^{\alpha}_{t,T}]}$, where S_T is the price of the underlying at time T (i.e. one month after time t), S_t is the underlying price at time t and $\mathbb{I}_t$ is the indicator function of a VaR exceedance.

The fourth and fifth columns report the p-values of the bootstrap hypothesis tests on K_t and S_t, ρ_k and ρ_S,respectively. The same statistics are presented in the 6^{th}-10^{th} columns when using the B&S method.

Option-implied MF and B&S CVaR and VaR

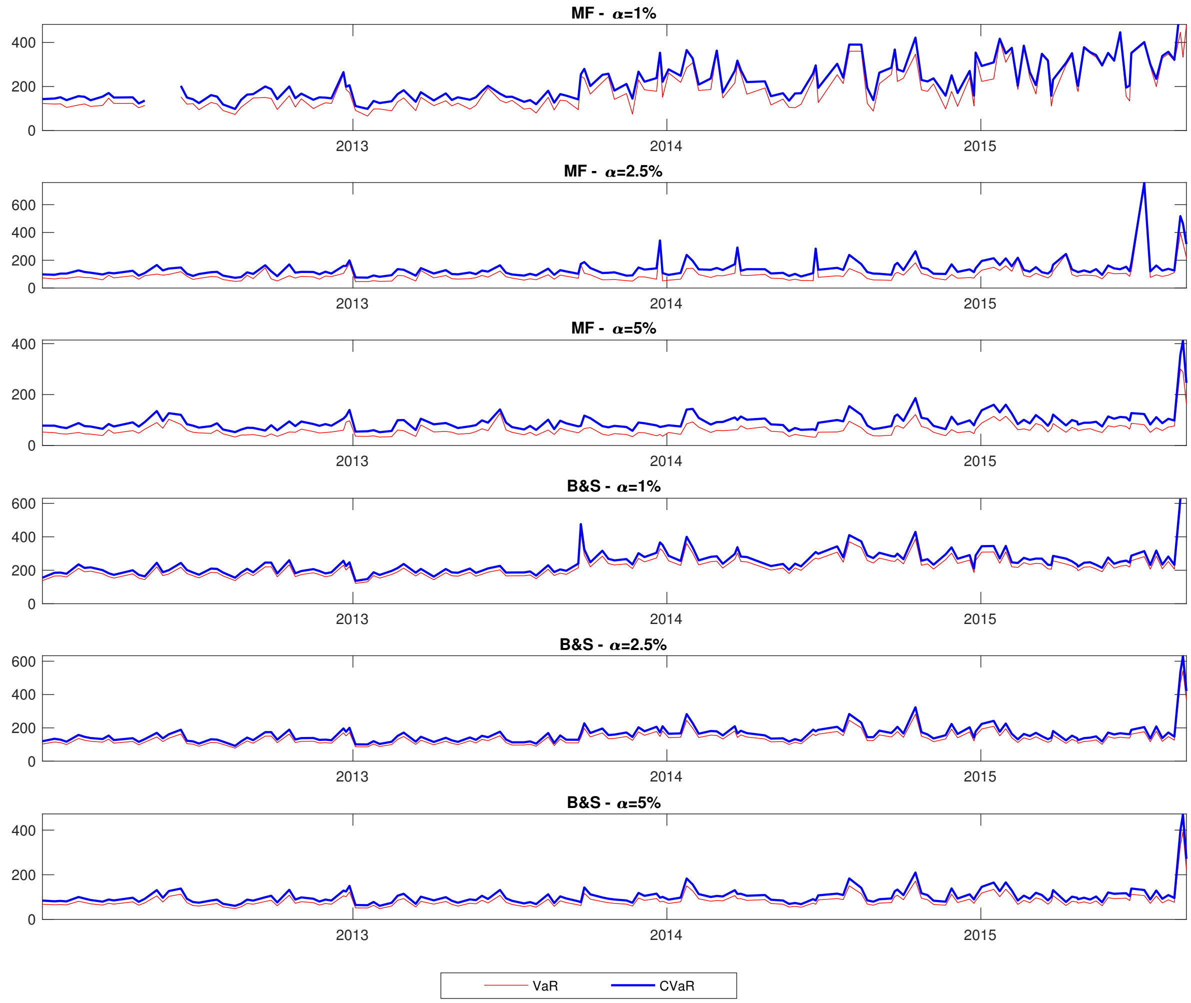

Figure 1: Time series of model-free (MF) and Black and Scholes (B&S) weekly CVaR and VaR. The three upper panels show the model-free (MF) CVaR (thick blue) and VaR (red) at level α=1% (top), 2.5% (central), 5% (bottom). The three lower panels show the Black and Scholes (B&S) CVaR (thick blue) and VaR (red) at level α=1% (top), 2.5% (central), 5% (bottom).

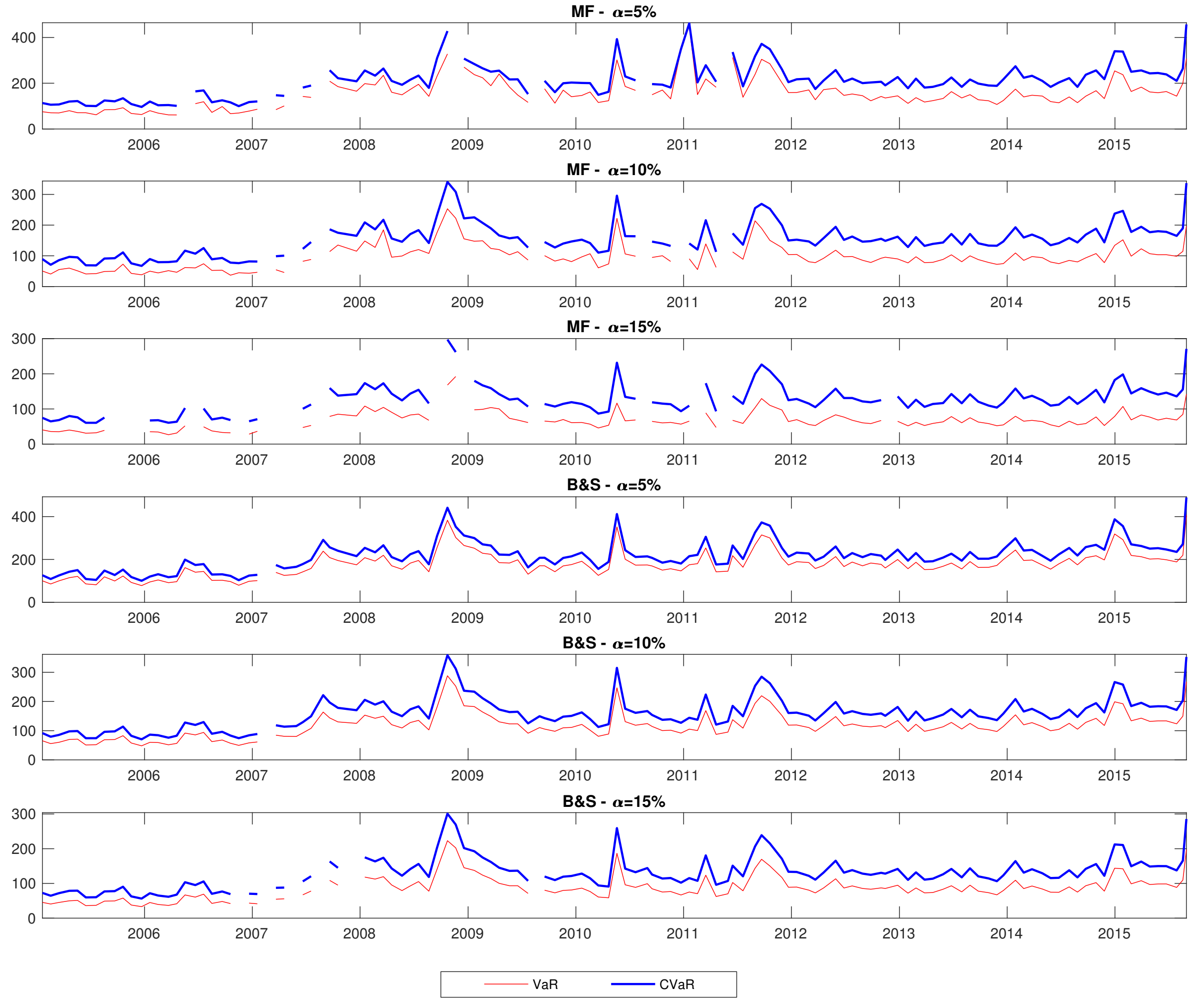

Figure 2: Time series of model-free (MF) and Black and Scholes (B&S) monthly CVaR and VaR. The three upper panels show the model-free (MF) CVaR (thick blue) and VaR (red) at level α=5% (top), 10% (central), 15% (bottom). The three lower panels show the Black and Scholes (B&S) CVaR (thick blue) and VaR (red) at level α=5% (top), 10% (central), 15% (bottom).

Backtesting results of the option-implied MF VaR

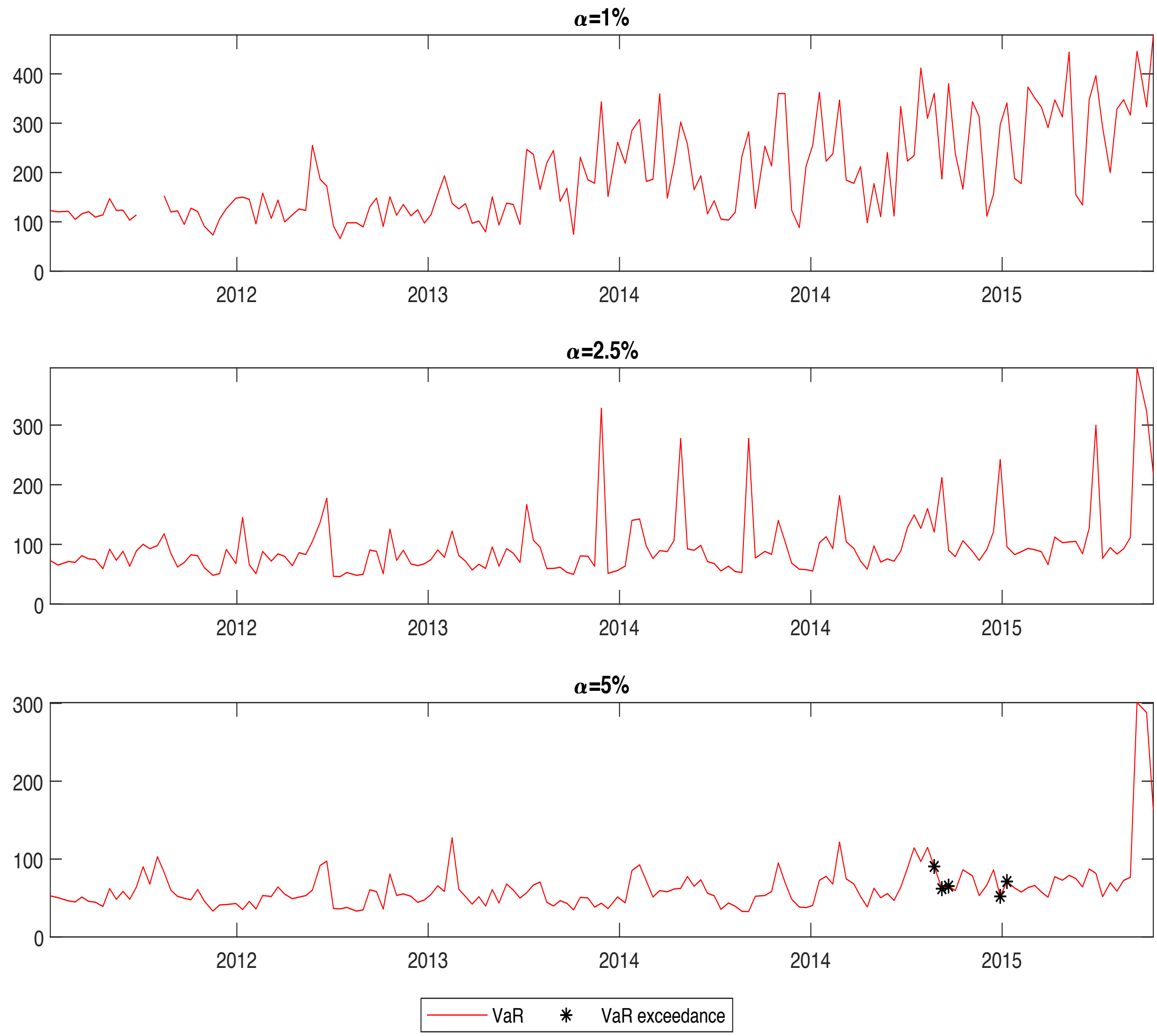

Figure 3: Time series of VaR and relative left tail VaR exceedances for the model-free (MF) weekly VaR at α=1% (top), 2.5% (central), 5% (bottom). VaR exceedances occur whenever the Put option contract at level α expires in-the-money (see Equation (28)) and are plotted with a black star (*).

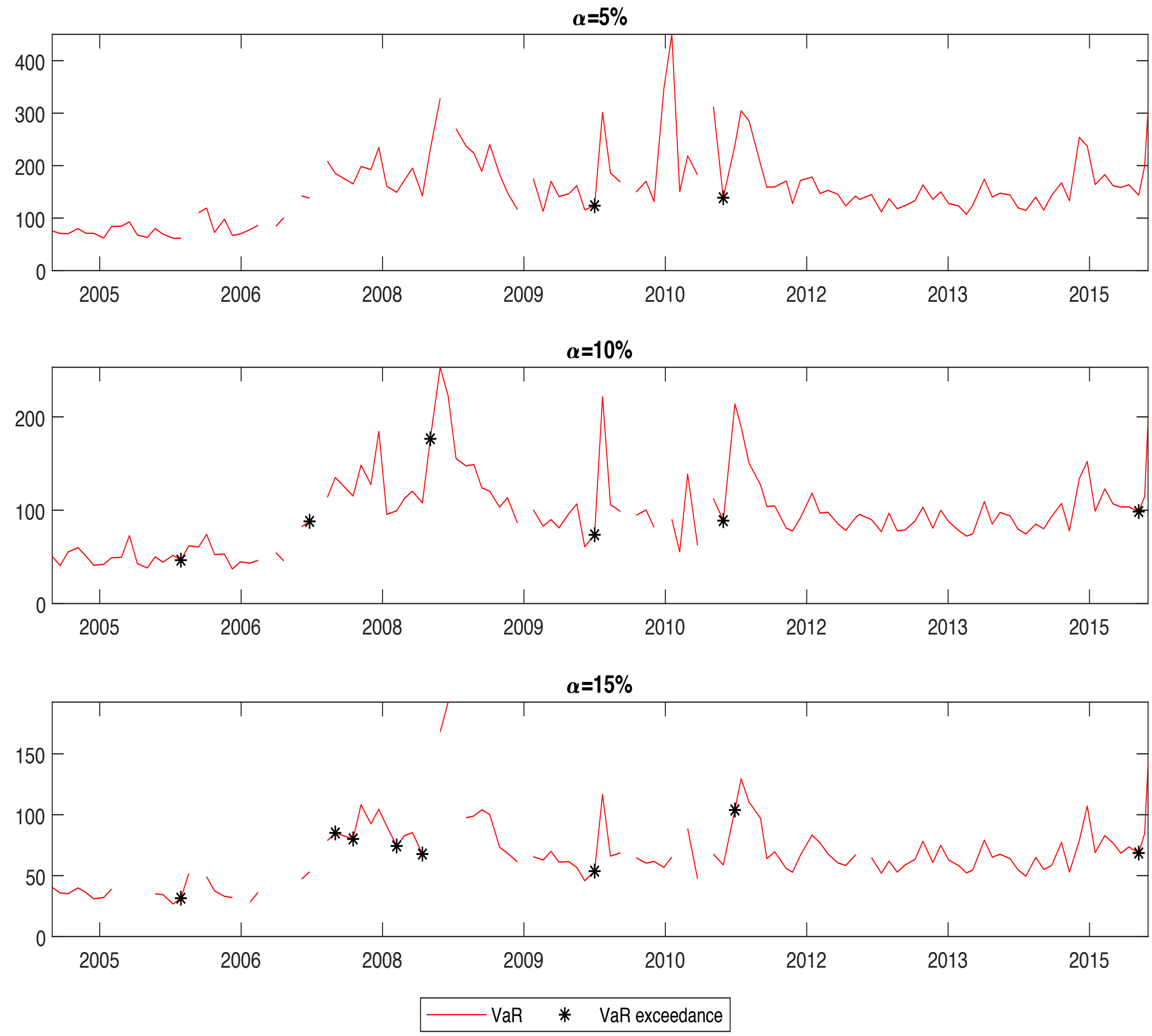

Figure 4: Time series of VaR and relative left tail VaR exceedances for the model-free (MF) monthly VaR at α=5% (top), 10% (central), 15% (bottom). VaR exceedances occur whenever the Put option contract at level α expires in-the-money (see Equation (28)) and are plotted with a black star (*).