Journal of Risk Model Validation **9**(2), 49–78

Risk journals

Research Paper

Commodity value-at-risk modeling: comparing RiskMetrics, historic simulation and quantile regression

Marie Steen,[1] Sjur Westgaard[2] and Ole Gjølberg[1]

[1]School of Economics and Business, Norwegian University of Life Sciences, PO Box 5003, 1432 Ås, Norway; emails: marie.steen@nmbu.no, ole.gjolberg@nmbu.no

[2]Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology, Alfred Getz vei 3, 7491 Trondheim, Norway; email: sjur.westgaard@iot.ntnu.no

(Received October 21, 2014; revised February 13, 2015; accepted February 13, 2015)

ABSTRACT

Commodities constitute a nonhomogeneous asset class. Return distributions differ widely across different commodities, both in terms of tail fatness and skewness. These are features that we need to take into account when modeling risk. In this paper, we outline the return characteristics of nineteen different commodity futures during the period 1992–2013. We then evaluate the performance of two standard risk modeling approaches, ie, RiskMetrics and historical simulation, against a quantile regression (QR) approach. Our findings strongly support the conclusion that QR outperforms these standard approaches in predicting value-at-risk for most commodities.

Keywords: quantile regression; value-at-risk; commodity prices; risk management; volatility; return distribution.

Corresponding author: M. Steen

Print ISSN 1753-9579 | Online ISSN 1753-9587

1 INTRODUCTION AND LITERATURE REVIEW

Over the last ten to fifteen years, there has been a growing interest in commodities as an asset class for investments. Several textbooks providing descriptions of the different commodity markets have been published. Excellent references are Dunsby *et al* (2008), Geman (2005), Gregoriou *et al* (2011) and Fabozzi *et al* (2008). These books provide valuable insights into the economics of commodity markets, the pricing and hedging of commodity derivatives, the valuation of commodity-producing companies, commodity-risk modeling and the performance of pure commodity funds (eg, exchange traded funds) and funds that include commodities together with traditional financial assets.

Commodities have risk characteristics that are different from those of financial assets such as stocks, bonds and currencies. Commodities are a nonhomogeneous asset class, and risk and return may differ substantially across commodities. Each commodity is driven by specific supply-and-demand conditions. While stock prices are discounted expected cashflows well into the future, the pricing of commodities will often be driven by short-term variations in supply. Major agricultural commodities are harvested once a year, and their pricing is therefore largely driven by weather conditions during the growing season. Extreme weather events that are likely to have an impact on harvesting yields may cause prices to move dramatically from one day or one week to the next. While some commodities (eg, electric power and cut flowers) are highly perishable and nonstorable, most commodities can be stored. This contributes to price stabilization. The storage costs, however, differ widely across commodities, and the effect of variations in storage volumes on prices and risk is typically nonlinear (see, for example, Pindyck 1994; Gorton *et al* 2013). As demonstrated in the seminal paper of Deaton and Laroque (1992), the rational expectations competitive storage model explains skewness, autocorrelation and the existence of violent explosions in prices. Beyond the theoretical arguments for nonlinearities and jumps in commodity prices due to changes in storage volumes, the way storage information is disseminated may be the cause of violent price changes. While commodities are priced continuously, storage statistics are published less frequently, typically once a month (or once a week for oil products). Unexpected changes in storage volumes may thus have dramatic effects on prices on announcement days.

The complex commodity-pricing relationship is manifest in differences in return distributions, as measured by volatility, skewness, kurtosis and empirical quantiles. Some commodity distributions have very high volatility, while others display less volatility. Some commodity distributions are skewed to the left and some to the right, in both cases generating an asymmetrical tail risk. Others have low/high kurtosis and, hence, low/high tail risk. Commodity return distributions also change over time because of changing market regimes, changes in commodity-specific business cycles, weather conditions, etc. It is therefore not obvious that standard risk models applied to stock, bond and foreign exchange markets, such as RiskMetrics and historical simulation, can be applied to commodities.

One problem with existing standard risk models such as RiskMetrics and historical simulation is that the former does not necessarily capture the correct return distribution conditional on the changing volatility. The latter has the opposite problem in that it captures the empirical return distribution but does not make it conditional on volatility. More advanced generalized autoregressive conditional heteroscedasticity (GARCH) models with different error distributions and conditional autoregressive value-at-risk (CAViaR) models, in which the quantiles are modeled as an autoregressive process, typically improve the fit; however, they are only used by market participants to a limited extent because of the complexity of estimating these models. A robust alternative that we propose in this paper is quantile regression (QR), with returns as the dependent variable and conditional volatility as the independent variable. This model may capture the complex relationship between the distribution of returns and volatility differing across commodities. The model is applied by running an exponentially weighted moving average volatility model (similar to RiskMetrics) before running a linear QR model using this volatility as an input. We analyze nineteen different commodity futures and the GSCI commodity index using daily observations from August 3, 1992 to November 11, 2013. The econometric results demonstrate that the QR applied to commodity returns with volatility as the explanatory variable is able to capture the complex distributions of commodity returns and predict value-at-risk (VaR) in sample better than RiskMetrics or historical simulation.

Few studies have applied QR to commodity market prices. One such study, however, is presented by Kuralbayeva and Malone (2012). They applied QR to model extreme commodity prices using a wide range of economic and financial factors as explanatory variables. Taking this approach, they describe: (1) nonlinear sensitivities to the fundamentals and (2) the full distribution of commodity returns (including both tails). Kuralbayeva and Malone found that these models explain more variation in extreme than median price innovations and that global financial and demand factors account for a greater proportion than commodity-specific factors such as basis and open interest. This is taken as evidence of the financialization of commodity markets during the period 2000–2009 via the increasing covariation of extreme commodity price changes within the US equity market.

While Kuralbayeva and Malone (2012) allowed for a thorough investigation of how different factors influence the tails of commodity price distributions, our paper has a narrower focus. We investigate how different models are suited for VaR estimates of commodity returns by comparing the standard RiskMetrics and historical simulation models with a QR approach using volatility as the only input. Our sample is also different as we include observations after 2009 through November 2013, and our selection of commodities is somewhat larger.

Several papers have investigated the risk-and-return characteristics of commodities in comparison with stock and bond markets. An early study by Bodie and Rosansky (1980) compared commodity returns with stock returns over the period 1950–76 using twenty-three different commodities. It concluded that the mean return for a commodity benchmark portfolio had about the same return as stocks during this period. Commodity futures tended to do well in years that stocks were doing badly and vice versa. Switching from a 100/0 stock/commodity portfolio to a 60/40 portfolio, Bodie and Rosansky found that investors could reduce their risk by one-third without sacrificing any returns. They also found that commodity futures provide a good inflation hedge. An additional finding was that individual commodities (unless they are physically related to each other) have a very low correlation. Bodie and Rosansky's findings are supported by Greer (2000), who studied the nature of commodity index returns rather than future positions. By analyzing data from 1970 to 1999, Greer found commodity returns and volatilities to be similar to those of equities. He also found commodity returns to be negatively correlated to stocks and bonds but positively correlated to inflation. The study also found a low correlation between different commodities. Gorton and Rouwenhorst (2004) investigated monthly commodity futures returns between 1954 and 2004. They found commodity risk premiums to be essentially the same as those of equities, while commodity returns during this period had been negatively correlated with equity and bond returns. In addition, they found a positive correlation between commodities and inflation. Erb and Harvey (2006) presented quite different results, concluding that commodity futures contracts have had annualized returns not significantly different from 0. Commodity returns might have equity-like returns if we focus on those having positive roll-over returns. Historic positive roll-over returns, however, cannot be expected to occur. Likewise, Kat (2006) and Kat and Oomen (2007a) did not find a significant risk premium in commodities using daily data for 142 different commodities (including different trading locations for the same commodity) from January 1965 to February 2005. They found that commodity futures returns vary substantially over business cycles, and the shape of the forward curve has a major impact on commodity returns. They also found that commodities can make an attractive diversifier, even without a proper risk premium. Hence, the case for investing in commodities together with stocks and bonds appears remarkably robust.

Some academics, market participants and policy makers have been quick to associate the strong inflows into commodity investments such as exchange traded funds with the recent commodity price spikes of 2007–8 and 2009–11. The argument is that commodities have become more volatile and more correlated with stocks and bonds, hence weakening the diversification benefits from commodities. There are, however, many authors who disagree that there has been a "finalization" of commodities. Using recent data, Brooks and Prokopczuk (2013) observed that each commodity has very distinct dynamics, and it is inappropriate to treat each commodity as a single asset class. In addition, they found that many commodities are still useful for diversifying stock and bond portfolios. This is also supported by Miffre (2011), especially when investors follow long–short trading strategies rather than hold long-only portfolios. Miffre (2011) finds no support for the hypothesis that speculators have destabilized commodity prices by increasing volatility or comovements between commodity prices and those of financial assets. Support for no systematic change in volatility for commodities is found in Steen and Gjølberg (2014a,b). In the latter paper, they found that skewness and kurtosis have not changed significantly either. In another paper, Steen and Gjøberg (2013) found no support for commodities having turned into "one" asset.

Although these studies provide key insights, they only investigate risk characteristics (eg, distributional properties of returns over time) across commodities to a limited extent. There are, however, a set of studies analyzing commodity risk characteristics in more detail. Kroner *et al* (1995) analyzed how to forecast commodity price volatility using implied volatilities from options written on seven agricultural and metal futures. Applying GARCH models to data from the period January 1987–November 1990, they provided good forecasts for almost all the commodities in the study. However, we could criticize the authors for basing their conclusions on a rather short sample period. More importantly, only under certain assumptions of parametric distribution of returns could we directly link volatility to VaR.

Giot and Laurent (2003) presented one of the first attempts at modeling VaR in commodity markets. They investigated a series of agricultural, metal and energy commodities (both spot and futures prices) using daily data from 1987 to 2002. The last five years were used for out-of-sample testing of several VaR models including RiskMetrics and several ARCH/GARCH models, in which the conditional volatility was assumed to follow an autoregressive process.[1] Tests for both long and short positions were conducted, evaluating both sides of the distribution. Giot and Laurent found that ARCH/GARCH models with skewed t-error distributions perform best. A more recent and extended study was performed by Füss *et al* (2010). The study included CAViaR models in which the quantiles were assumed to follow an autoregressive process (see Engle and Manganelli (2004) for more details). They used daily in-sample data from 1991 to 2004 and out-of-sample data from 2004 to 2006 covering energy, agricultural and metal markets. In addition, to verify that advanced GARCH

---

[1] See Engle (1982) for ARCH models, Bollerslev (1986) for GARCH models with normally distributed errors and Lambert and Laurent (2001) for ARCH and GARCH models with skewed t-distributed errors.

models perform well, they also concluded that CAViaR models return very good out-of-sample VaR performances.[2]

The results above indicate that simpler models such as RiskMetrics and GARCH models assuming normal error distributions do not perform well. To improve the risk forecasts, GARCH with skewed t or CAViaR models could be applied. The problem with these models is that they are difficult to implement, as they require nonlinear optimization procedures. As a consequence, these models have only been used by market participants to a limited extent. The goal of our paper is to investigate updated data for commodity futures by applying a model that (1) provides good VaR predictions and (2) can be estimated using easily available information. This has been the motivation for building a QR model with returns as the dependent variable and conditional volatility as the independent variable. We will describe and evaluate this model later.

The rest of this paper is organized as follows. In Section 2, we describe the data set we use and the descriptive commodity return statistics. In Section 3, we describe the various VaR models (RiskMetrics, historical simulation and QR). We also describe backtesting procedures for VaR (the Kupiec test and the Christoffersen test). Section 4 backtests the different VaR models for various quantiles. Section 5 concludes and discusses the implications of our results for commodity-price risk management. We also propose ideas for further research based on the insights from this study.

2 DATA AND DESCRIPTIVE STATISTICS

Figure 1 on page 56 graphs the development of commodity prices from August 1992 to November 2013, totaling 5392 trading days. The data covers front month futures prices from the CME Group. The commodities are (see Table 1 on page 58 for abbreviations): crude oil, heating oil, natural gas, cotton, corn, wheat, soybeans, rough rice, soybean meal, soybean oil, lean hogs, feeder cattle, copper, gold, silver, cocoa, coffee, orange juice and sugar. We also include the Goldman Sachs Commodity Index (GSCI). The prices were downloaded from Quandl (see www.quandl.com for more information).

When constructing the daily return series for commodities, it is crucial to adjust for the roll-over returns, that is, the jumps that are typically generated when a front month contract is rolled over to the next front contract. These roll effects can easily give a distorted picture of volatility, in particular when it comes to commodities with seasonal production and consumption patterns. In this paper, we have simply deleted the return at each roll date. This correction is time consuming since different

[2] In the papers listed in the paragraph above and in this paper, we focus on univariate VaR modeling for commodities. Papers such as those by Kat and Oomen (2007b) and Börger (2009) focus on multivariate commodity analysis and applications to risk management that suggest a direction for further research not conducted here.

commodity contracts have different roll-over dates. We delete rows of observations where one or more roll-overs occur. In our data set, this adjustment reduces the number of returns from 5392 to 4470, creating a continuous series of corrected returns that can be compared and analyzed across all commodities.

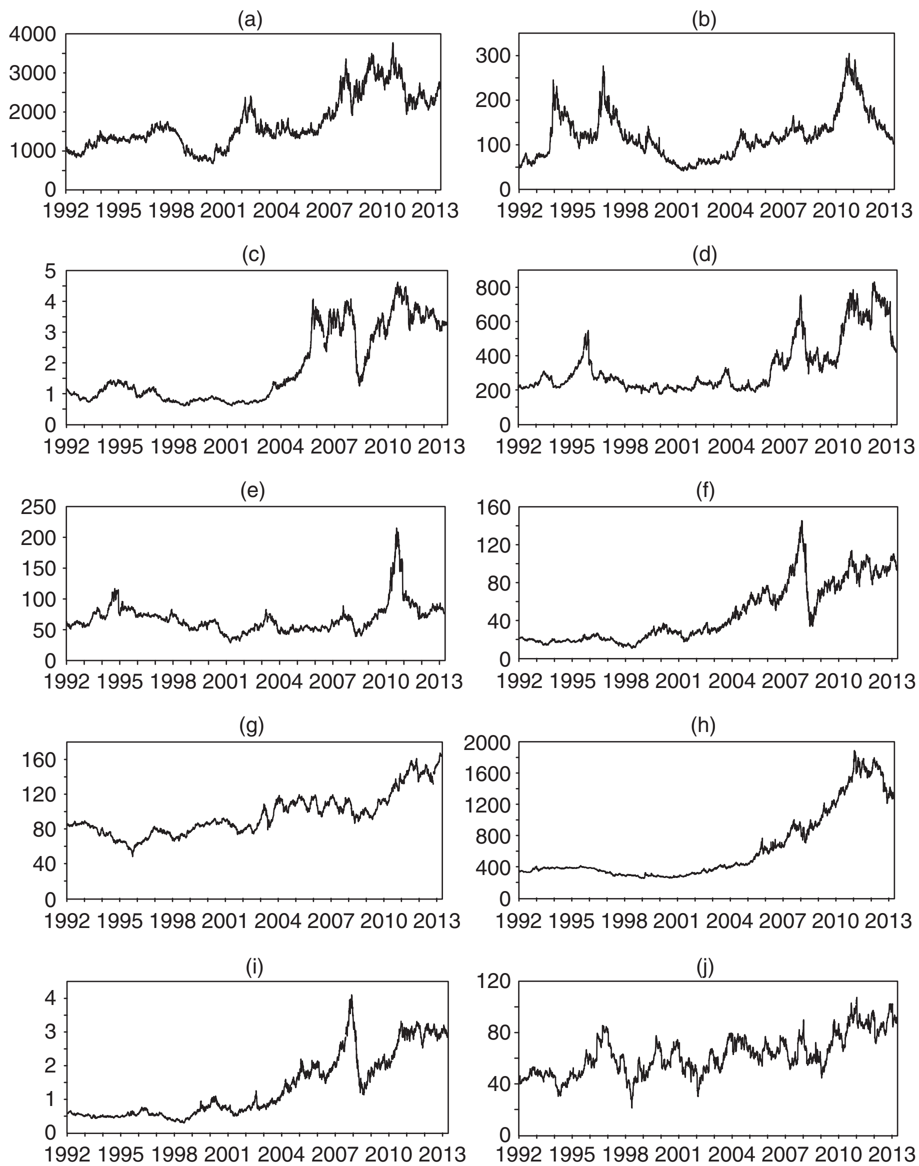
To elaborate on the different daily return characteristics of commodities, we present descriptive statistics for the entire period from 1992 to 2013 in Table 1 on page 58.

Table 1 shows daily values for the mean, standard deviation, minimum, maximum, skewness, kurtosis and the 5% and 95% return quantiles. As the price dynamics of different commodities are different, it comes as no surprise that the return distributions also differ considerably. All mean returns are close to 0, as expected for daily returns. The daily standard deviations, however, differ substantially. As has been found in other studies, the volatility of natural gas prices is the highest (3.65%, or 57.70% on an annual basis after multiplying by the square root of 250 trading days). In contrast, feeder cattle and gold futures are the least risky (0.84% and 1.06% on a daily basis, respectively). Table 1 also reports large differences in the skewness and kurtosis. Strong positive skewness and high excess kurtosis are found for lean hogs (3.99 and 59.62, respectively). Silver, on the other hand, has a negative skewness (−0.583). Soyabean oil has a skewness of close to 0 and an excess kurtosis of 2, making its return distribution close to normal. The different distributional properties make the VaR or quantiles at 5% and 95% different. In addition to the variation in absolute levels, there are asymmetries in the tail distributions. Some distributions have a long left tail (eg, cotton), making the 5% VaR higher in absolute value than the 95% VaR. Others have a long right tail (eg, natural gas), making the 95% VaR higher in absolute value than the 5% VaR. This highlights the need for a risk model that allows for nonnormal distributions to capture such asymmetry.

In this paper, we are particularly interested in modeling risk characteristics over time. A natural question to ask, therefore, is whether distributional properties (specifically the VaR quantiles) change over time due to changing marking conditions. Table 2 on page 59 and Table 3 on page 61 show how the 5% VaR and 95% VaR, respectively, change annually for each commodity.

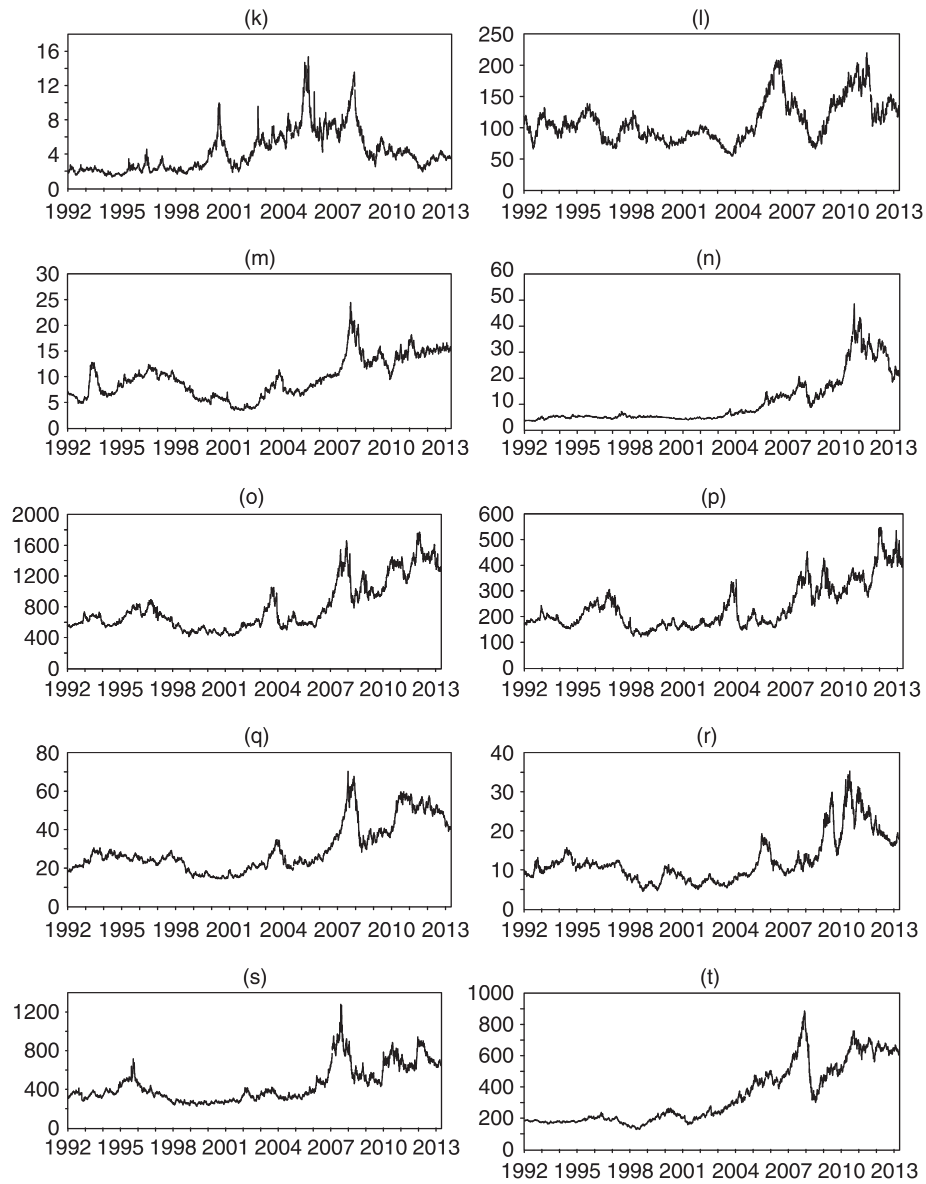
Table 2 and Table 3 reveal that risk varies dynamically over time for most commodities. For example, the 5% and 95% VaRs for crude oil vary in the range [−2.82% to −9.58%] and [2.75% to 13.96%], respectively, from 1993 to 2013. The historical quantiles for many of the agricultural products, on the other hand, are more stable. Many commodities, but not all, faced higher risk for long and short positions during the financial crises. The VaR values for lean hogs and feeder cattle in the years 2007–9 were no higher in absolute value than before the crisis. For example, lean hogs faced the largest risk for the 5% VaR during 1998 (−7.22%), which may be related to the outbreak of mad cow disease. During the financial crisis in 2008, the 5% VaR for lean hogs was only −3.67%. Statements such as "all commodities behaved like

FIGURE 1 Commodity prices, August 3, 1992–November 12, 2013. [Figure continues on next page.]



(a) Cocoa. (b) Coffee. (c) Copper. (d) Corn. (e) Cotton. (f) Crude oil. (g) Feeder cattle. (h) Gold. (i) Heating oil. (j) Lean hogs. $N = 5.392$ daily observations for front month futures contracts at CME.

FIGURE 1 Continued.



(k) Natural gas. (l) Orange juice. (m) Rough rice. (n) Silver. (o) Soybeans. (p) Soybean meal. (q) Soybean oil. (r) Sugar. (s) Wheat. (t) GSCI. $N = 5.392$ daily observations for front month futures contracts at CME.

TABLE 1 Commodity returns, August 4, 1992–November 12, 2013.

Period 1993–2013		Empirical distribution							
Variable	**Market**	**Mean**	**SD**	**Min**	**Max**	**Skew**	**Kurtosis**	**5% quantile**	**95% quantile**
CL01	Crude oil	0.00	0.02	−0.15	0.18	0.14	4.98	−0.04	0.03
HO01	Heating oil	0.00	0.02	−0.19	0.09	−0.24	3.45	−0.03	0.03
NG01	Natural gas	0.00	0.04	−0.31	0.38	0.69	8.52	−0.05	0.06
CT01	Cotton	0.00	0.02	−0.12	0.12	−0.06	2.73	−0.03	0.03
C01	Corn	0.00	0.02	−0.10	0.09	0.05	2.61	−0.03	0.03
W01	Wheat	0.00	0.02	−0.09	0.08	0.08	2.12	−0.03	0.03
S01	Soybean	0.00	0.02	−0.08	0.23	0.61	13.24	−0.02	0.02
RR01	Rough rice	0.00	0.02	−0.10	0.10	0.12	2.75	−0.03	0.03
SM01	Soybean mean	0.00	0.02	−0.08	0.08	0.00	2.28	−0.02	0.03
BO01	Soybean oil	0.00	0.01	−0.07	0.08	0.19	2.00	−0.02	0.02
LH01	Lean hogs	0.00	0.02	−0.17	0.33	3.99	59.62	−0.03	0.02
FC01	Feeder cattle	0.00	0.01	−0.08	0.07	0.06	9.85	−0.01	0.01
HG01	Copper	0.00	0.02	−0.11	0.12	−0.02	3.66	−0.03	0.03
GC01	Gold	0.00	0.01	−0.09	0.09	0.03	8.44	−0.02	0.02
SI01	Silver	0.00	0.02	−0.18	0.13	−0.58	6.88	−0.03	0.03
CC01	Cocoa	0.00	0.02	−0.10	0.10	0.06	2.11	−0.03	0.03
KC01	Coffee	0.00	0.02	−0.14	0.27	0.86	10.69	−0.04	0.04
OJ01	Orange juice	0.00	0.02	−0.12	0.20	0.39	6.19	−0.03	0.03
SB01	Sugar	0.00	0.02	−0.12	0.12	−0.07	2.21	−0.03	0.04
GSCI	GSCI	0.00	0.01	−0.08	0.08	−0.14	2.79	−0.02	0.02

$N = 4.470$ daily observations, front month futures contracts from CME. Roll-over returns are deleted.

TABLE 2 Empirical 1% VaR (quantile), 1993–2013. [Table continues on next page.]

Market	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Crude oil	–3.5	–4.5	–2.9	–5.4	–3.9	–6.5	–5.2	–5.9	–5.5	–4.7	–6.8
Heating oil	–2.9	–5.5	–2.8	–5.7	–4.3	–4.8	–5.1	–8.5	–7.9	–5.0	–6.3
Natural gas	–8.7	–7.4	–9.4	–12.9	–8.4	–7.4	–7.4	–8.1	–14.4	–6.6	–9.4
Cotton	–3.2	–2.8	–5.5	–3.0	–2.0	–3.6	–4.2	–3.0	–4.9	–4.8	–3.9
Corn	–2.0	–4.1	–1.5	–3.9	–3.5	–2.9	–4.0	–3.7	–2.7	–2.8	–3.2
Wheat	–2.4	–2.7	–3.1	–7.2	–3.0	–2.7	–3.8	–2.8	–2.5	–3.3	–3.7
Soybeans	–3.2	–4.3	–3.2	–3.8	–3.6	–2.3	–3.9	–2.9	–2.8	–2.9	–2.7
Rough rice	–3.3	–4.3	–3.6	–2.4	–2.7	–2.3	–3.4	–3.8	–6.9	–4.5	–3.7
Soybean meal	–2.5	–4.0	–3.4	–3.9	–3.5	–3.6	–4.1	–2.8	–2.6	–3.2	–2.6
Soybean oil	–3.0	–4.1	–2.6	–2.4	–2.4	–2.4	–4.1	–3.0	–3.4	–3.0	–2.5
Lean hogs	–3.1	–4.4	–3.5	–5.3	–2.8	–7.2	–4.6	–3.0	–3.4	–5.6	–3.6
Feeder cattle	–1.4	–1.8	–2.3	–2.6	–1.7	–2.1	–1.3	–1.0	–1.7	–1.9	–5.4
Copper	–4.2	–3.2	–3.9	–7.7	–4.8	–2.9	–4.2	–2.6	–2.5	–2.0	–2.6
Gold	–2.3	–1.6	–1.2	–1.1	–2.1	–1.9	–2.6	–2.0	–1.7	–2.0	–2.7
Silver	–4.6	–3.9	–3.6	–3.4	–3.4	–4.6	–3.8	–2.6	–2.5	–3.5	–2.8
Cocoa	–4.7	–4.0	–2.9	–2.8	–3.1	–2.6	–4.9	–5.1	–4.3	–5.2	–7.1
Coffee	–7.8	–7.6	–6.3	–4.0	–9.2	–6.2	–10.1	–7.3	–5.9	–5.6	–5.0
Orange juice	–5.7	–5.5	–4.6	–4.8	–5.5	–5.3	–6.2	–4.7	–3.2	–2.8	–3.5
Sugar	–5.7	–4.5	–5.2	–3.6	–2.3	–5.0	–6.4	–5.1	–4.7	–5.0	–4.4
GSCI	–1.5	–1.9	–1.3	–2.8	–2.4	–2.6	–2.5	–3.9	–4.4	–2.8	–4.0

All values given in percent.

TABLE 2 Continued.

Market	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Crude oil	−5.0	−4.0	−3.8	−4.0	−9.6	−8.7	−4.1	−5.9	−4.1	−2.8
Heating oil	−5.2	−4.5	−3.9	−3.7	−7.0	−6.6	−3.8	−4.7	−2.9	−2.8
Natural gas	−8.1	−6.4	−9.5	−6.9	−7.3	−8.5	−5.1	−4.7	−6.1	−4.4
Cotton	−5.6	−4.8	−3.5	−2.8	−6.3	−4.8	−4.6	−4.7	−5.1	−4.2
Corn	−3.7	−3.3	−3.0	−5.1	−6.6	−4.6	−5.2	−4.6	−4.1	−5.5
Wheat	−3.8	−2.7	−3.9	−3.6	−8.0	−5.0	−6.3	−6.7	−4.2	−3.0
Soybeans	−5.7	−3.9	−2.6	−4.0	−5.7	−6.0	−3.4	−3.1	−3.1	−3.4
Rough rice	−4.4	−3.3	−2.8	−2.9	−5.4	−3.8	−3.6	−3.6	−2.6	−3.0
Soybean meal	−5.8	−3.8	−2.7	−3.9	−5.8	−7.3	−5.0	−3.4	−3.5	−4.1
Lean hogs	−2.6	−2.7	−3.0	−2.5	−3.7	−4.5	−3.4	−3.0	−3.0	−3.9
Feeder cattle	−4.1	−2.5	−2.3	−2.0	−3.0	−2.8	−1.3	−1.8	−1.9	−1.5
Copper	−5.1	−6.2	−6.4	−4.6	−7.0	−5.2	−4.6	−5.5	−2.9	−2.6
Gold	−3.3	−1.8	−3.6	−3.1	−4.3	−3.0	−2.6	−4.4	−3.1	−4.3
Silver	−7.2	−3.0	−9.2	−5.0	−8.1	−4.8	−5.3	−7.6	−5.4	−6.2
Cocoa	−5.7	−3.2	−3.9	−4.6	−6.9	−5.4	−4.2	−4.5	−3.9	−2.7
Coffee	−5.3	−4.6	−3.7	−3.6	−5.6	−3.8	−4.2	−4.2	−4.6	−3.9
Orange juice	−6.9	−4.6	−4.2	−5.4	−6.2	−6.5	−5.0	−4.8	−6.1	−4.7
Sugar	−5.8	−2.9	−6.8	−3.8	−6.1	−3.9	−9.4	−4.7	−4.5	−3.2
GSCI	−3.9	−3.0	−2.9	−2.7	−6.5	−5.3	−3.4	−4.3	−2.6	−2.1

Returns, August 4, 1992–November 12, 2013. $N = 4.470$ daily observations, front month futures contracts from CME. Roll-over returns are deleted. All values given in percent.

TABLE 3 Empirical 99% VaR (quantile), 1993–2013. [Table continues on next page.]

Market	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Crude oil	3.7	4.8	2.9	8.6	3.8	8.0	4.8	5.2	6.1	4.8	5.8
Heating oil	3.7	4.7	2.8	4.3	3.5	6.1	4.6	6.6	6.2	5.3	7.2
Natural gas	7.8	12.3	7.5	12.9	7.9	9.3	8.9	9.6	11.2	9.5	10.1
Cotton	3.3	2.9	3.3	2.7	1.7	2.9	3.2	3.5	5.2	5.5	4.3
Corn	2.9	2.4	2.3	3.3	3.2	3.2	3.6	2.9	3.1	3.7	3.9
Wheat	2.8	2.9	3.8	3.8	3.0	2.9	3.8	2.8	3.0	4.2	4.1
Soybeans	3.0	3.7	1.9	3.3	4.3	2.7	4.3	3.0	2.9	2.8	3.3
Rough rice	6.1	4.6	3.6	2.3	2.7	2.7	3.9	4.3	4.7	5.1	4.8
Soybean meal	3.5	3.0	2.6	2.9	4.1	4.4	4.5	3.4	3.3	3.0	4.5
Soybean oil	4.5	2.6	3.5	2.2	2.9	2.5	4.0	3.4	3.2	3.2	3.3
Lean hogs	4.7	7.6	3.6	2.8	2.8	4.4	6.5	2.9	4.0	6.1	3.4
Feeder cattle	1.2	1.9	1.8	2.9	1.6	2.1	2.1	1.4	1.8	2.0	1.7
Copper	3.2	3.9	3.3	5.4	4.0	3.7	4.1	2.8	2.9	3.3	2.8
Gold	2.1	1.5	1.1	0.9	1.5	2.1	4.0	3.2	2.0	2.2	2.1
Silver	4.8	3.0	4.3	2.7	5.7	4.6	3.7	2.1	2.1	2.5	3.1
Cocoa	4.8	4.5	3.9	2.6	4.0	2.9	5.0	4.2	5.9	4.5	4.6
Coffee	5.6	10.8	4.8	6.2	6.8	6.4	9.8	7.1	8.0	6.8	4.6
Orange juice	5.7	5.0	4.5	4.7	7.2	7.8	5.7	4.4	4.4	2.8	2.8
Sugar	5.9	4.1	4.2	3.9	3.0	5.0	8.6	6.0	4.1	6.4	3.5
GSCI	1.9	2.3	1.3	2.8	2.4	3.0	2.8	3.6	3.6	2.9	3.7

All values given in percent.

TABLE 3 Continued.

Market	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Crude oil	4.9	4.9	4.4	4.8	14.0	10.2	3.9	5.2	4.2	2.7
Heating oil	6.4	7.0	4.0	4.7	7.9	6.9	4.0	3.6	3.5	2.4
Natural gas	13.9	10.2	11.7	8.8	7.6	15.1	5.9	5.6	7.7	4.6
Cotton	6.2	4.5	3.3	3.0	7.2	4.6	4.6	4.4	5.6	3.0
Corn	2.8	3.6	5.0	4.7	5.8	5.2	5.0	6.1	6.5	4.1
Wheat	4.2	3.8	5.8	4.7	7.6	5.0	6.6	6.8	5.7	3.1
Soybeans	4.8	3.5	2.8	2.7	5.6	5.2	3.2	3.7	3.5	3.2
Rough rice	6.3	3.7	4.0	3.7	4.0	3.7	3.8	3.9	2.9	3.0
Soybean meal	4.8	4.7	3.4	3.4	6.0	5.1	3.9	4.0	3.9	4.9
Soybean oil	4.7	4.3	3.2	2.7	5.9	5.5	3.1	3.0	3.1	2.7
Lean hogs	2.9	3.1	4.3	5.1	3.9	5.2	3.9	3.7	3.0	2.7
Feeder cattle	2.2	1.5	1.9	2.2	4.9	2.2	2.0	2.3	2.2	3.2
Copper	4.0	3.2	6.0	4.5	6.7	7.7	4.0	5.5	3.4	2.7
Gold	2.1	2.0	2.7	2.1	5.4	3.3	2.0	2.9	2.5	2.7
Silver	5.0	3.5	6.0	3.1	8.8	5.0	5.2	5.4	4.6	4.6
Cocoa	6.0	5.0	2.8	3.8	5.5	5.3	4.0	3.9	6.3	3.2
Coffee	6.0	5.4	4.7	3.6	4.2	5.0	5.1	4.5	4.7	3.3
Orange juice	5.2	3.5	5.1	5.1	4.9	5.9	3.9	4.1	6.5	5.4
Sugar	4.0	3.5	6.3	5.2	6.4	5.3	6.5	5.3	4.4	3.1
GSCI	3.6	3.8	2.8	3.7	5.8	5.3	3.0	3.4	3.1	1.7

Returns, August 4, 1992–November 12, 2013. $N = 4.470$ daily observations, front month futures contracts at CME. Roll-over returns are deleted. All values given in percent.

stocks during the financial crisis" are therefore not very precise. Commodity risk differs across commodities and over time, reflecting differences in supply and demand changes across commodities.

Figure 2 on the next page and Figure 3 on page 65 show the relationship between return and volatility for the various commodities. The volatility is calculated as the exponential weighted moving average using a smoothing parameter of 0.94, which is the value used by RiskMetrics for daily data.

Some commodities (eg, corn) display a symmetrical relationship, while others show asymmetry. For example, when volatility is high, positive returns are higher than negative returns in absolute value for crude oil. The opposite is found for wheat. We sum up the basic empirical facts as follows.

- Distributional properties (volatility, skewness, kurtosis and quantiles) vary substantially across commodities.
- Distributional properties for most commodities vary over time but not in a uniform fashion.
- The relationship between returns and volatility for commodities can be symmetrical or exhibit positive or negative asymmetries.

A proper risk model needs to take these features into account. In Section 3, we discuss the most commonly applied risk models, ie, RiskMetrics and historical simulation, and how to implement alternative models based on QR that take some of the problematic assumptions behind the former two models into account.

3 METHODOLOGY

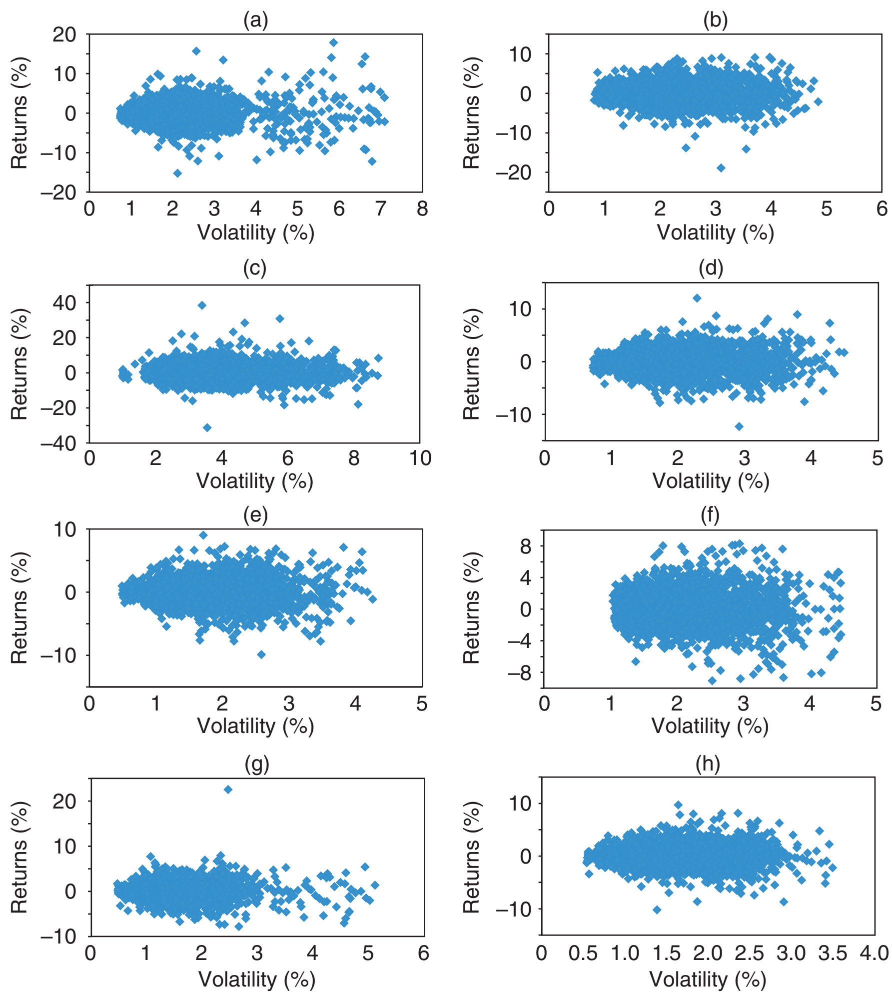
RiskMetrics and historical simulation are the most commonly used models for VaR estimation in financial institutions (see Alexander 2009). We propose QR as an easy-to-implement alternative.[3] First, we briefly outline and explain the estimation techniques used to compute the VaR estimates in sample and test statistics for model evaluation before proceeding to the QR approach.

3.1 Value-at-risk models

The RiskMetrics model is equivalent to an IGARCH (nonstationary conditional volatility) model with normally distributed errors. The conditional volatility from the

[3] Complex GARCH models allowing for different dynamics and error distributions have been shown to perform well in VaR predictions (see Giot and Laurent 2003). Another alternative is the so-called CAViaR model, in which the quantile itself is modeled as an autoregressive process (see Engle and Manganelli 2004; Füss *et al* 2010). As these models are much more difficult to implement because they require nonlinear solvers, we do not consider these models here.

FIGURE 2 Scatterplot of returns, August 4, 1992–November 12, 2013.



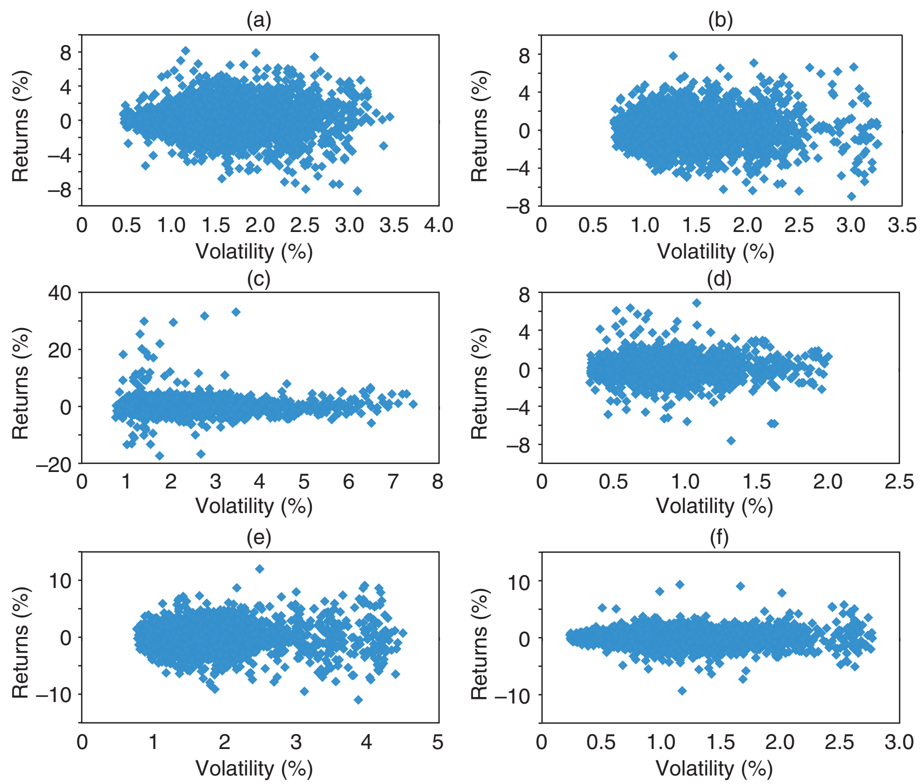
$N = 4.470$ daily observations, front month futures contracts at CME. Roll-over returns are deleted. For ticker codes, see Table 1 on page 58. (a) CL1. (b) H01. (c) NG1. (d) CT1. (e) C1. (f) W1. (g) S1. (h) RR1.

model is similar to an exponentially weighted moving average in which the weighting parameter λ is set equal to 0.94 for daily data. The model is given in (3.1) and (3.2):

$$r_t = \varepsilon_t \sigma_t, \quad \text{where } \varepsilon_t \sim N(0, 1), \tag{3.1}$$

$$\sigma_t^2 = (1 - \lambda) r_{t-1} + \lambda \sigma_{t-1}^2, \tag{3.2}$$

FIGURE 3 Scatterplot of returns, August 4, 1992–November 12, 2013. [Figure continues on next page.]



(a) SM1. (b) BO1. (c) LH1. (d) FC1. (e) HG1. (f) GC1.

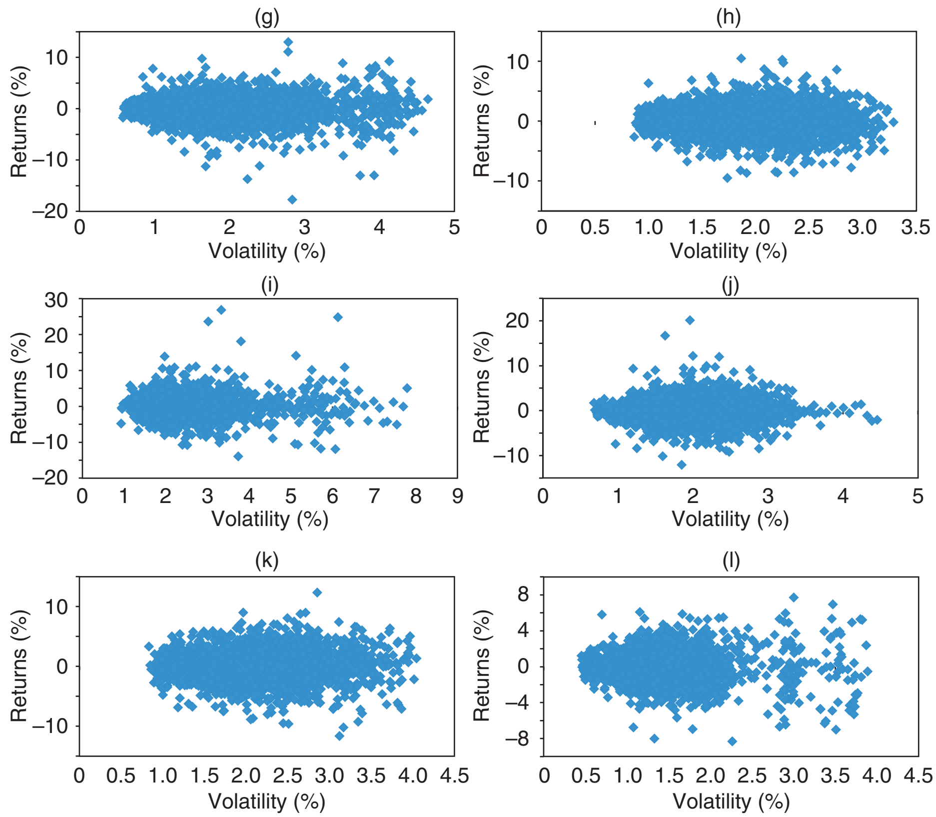
where r_t is the daily (log) return for a given commodity, e_t is the error term and s_t is the volatility of the returns. The weighting parameter is given by λ. Assuming a 0 mean return for daily data, the VaR or quantile of the distribution is given by

$$\mathrm{VaR}_\alpha = Z_\alpha \sigma_t, \tag{3.3}$$

where α is the chosen significance level and Z_α is the quantile in accordance with a normal distribution for this specific significance level. A nice feature of this model is that volatility can be updated dynamically. Financial and commodity return volatility tend to change over time, and this approach captures this property.

A critical assumption in this model is that returns follow a normal distribution. As demonstrated above and in several other papers, this is not typically the case for

FIGURE 3 Continued.



$N = 4.470$ daily observations, front month futures contracts at CME. Roll-over returns are deleted. For ticker codes, see Table 1 on page 58. (g) SI1. (h) CC1. (i) KC1. (j) OJ1. (k) SB1. (l) GSCI.

financial and commodity returns at daily frequency. Numerous studies have concluded that returns are characterized by fat tails. This will lead to an underestimation of risk. In addition, the standard assumption is that the risks associated with long positions (left quantile) and short positions (right quantile) are equal. That is, the risk is symmetrical: $\mathrm{VaR}_\alpha = -\mathrm{VaR}_{1-\alpha}$. However, financial and commodity returns can be negatively or positively skewed, which again will lead to bias in the estimation of VaR.

Historical simulation analysis is a procedure for predicting VaR by simulating or constructing the cumulative distribution function (CDF) of asset returns over time. Based on the constructed CDF, VaR can be found at given confidence levels. By taking the average of many simulations, we obtain not only a VaR estimate but also the distribution around this estimate. Unlike the parametric VaR model just described, historical simulation does not assume a particular distribution of the asset returns. Further, it is relatively easy to implement.

There is a significant shortcoming inherent in historical simulation, as it imposes a restriction on the estimation by assuming that asset returns are independent and identically distributed random variables. From empirical studies, asset returns are found to exhibit volatility clustering. This shortcoming has initiated attempts to develop other nonparametric and semiparametric models. Two such alternatives are historical filtered simulation (not implemented in our study) and QR.

3.2 Quantile regression

VaR is simply a particular quantile of the future return values, conditional on current information. QR, as introduced by Koenker and Bassett (1978), seems an obvious choice for VaR forecasting. Applications in financial risk management can be found in Engle and Granger (1987), Taylor (2008) and Alexander (2009).

If r_t is the dependent variable and s_t is an independent variable according to the definition under RiskMetrics, the simple linear QR model is given by

$$r_t^q = \alpha^q + \beta^q \sigma_{t-1} + \varepsilon_t^q, \tag{3.4}$$

where ε has an unspecified distribution function. The conditional qth quantile, $0 < q < 1$, is defined as any solution to the minimization problem, as suggested by Koenker and Bassett (1978):

$$\min_{\alpha,\beta} \sum_{t=1}^{T} (q - \mathbf{1}_{r_t \leqslant \alpha + \beta\sigma_{t-1}})(Y_t - (\alpha + \beta\sigma_{t-1})), \tag{3.5}$$

where

$$\mathbf{1}_{r_t \leqslant \alpha + \beta\sigma_{t-1}} = \begin{cases} 1 & \text{if } r_t \leqslant \alpha + \beta\sigma_{t-1}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.6}$$

The least absolute error (the conditional median) is a special case, but the QR method explicitly allows for the modeling of all relevant quantiles of the distribution of the dependent variable. As VaR is a particular conditional quantile of future portfolio returns, the conditional quantile function can be expressed as

$$\mathrm{VaR}_t^q \mid \sigma_{t-1} = \hat{\alpha}_t^q + \hat{\beta}_t^q \sigma_{t-1} + \varepsilon_t^q \mid \sigma_{t-1}. \tag{3.7}$$

A unique set of regression parameters (α, β) can be obtained for each quantile of interest, and the whole return distribution given a value for the conditional volatility can be found.

3.3 Backtesting procedure for value-at-risk

Backtesting refers to testing the accuracy of VaR over a historical period when the true outcome is known. The general approach to backtesting VaR for an asset or portfolio is to record the number of occasions over a historical period on which the actual loss exceeds the model VaR and compare this number with the prespecified level. Each model predicts in-sample VaR at different significance levels for a given day. The predicted one-day VaR (for both long and short positions) is then compared with the observed return for that specific day. If the actual return is lower (higher) than the VaR prediction for a long (short) position, then a violation exists. The violation is often referred to as an "exceedance" or a "hit". For example, a hit occurs when $\mathrm{VaR}_{5\%}$ is predicted to be -2% and the actual return is -3%. Similarly, a hit occurs if $\mathrm{VaR}_{95\%}$ is predicted to be 2.5% and the actual return is 4%. The total number of hits divided by the total number of observations in the sample should be as close as possible to the prespecified VaR number. For example, let the significance level be 5%, assuming 1000 observations in the sample and fifty-three hits. The VaR is $53/1000 = 5.3\%$. The question is this: is 5.3% "close enough" to 5% to classify the VaR model as appropriate? Two test procedures are usually applied to answer this question, ie, the Kupiec test and the Christoffersen test.

The Kupiec (1995) test is a likelihood ratio test designed to reveal whether the model provides the correct unconditional coverage. More precisely, let H_t be an indicator sequence, where H_t takes the value 1 if the observed return, Y_t, is below the predicted VaR quantile, Q_t, at time t:

$$H_t = \begin{cases} 1 & \text{if } r_t \leqslant \mathrm{VaR}_q, \\ 0 & \text{otherwise.} \end{cases} \tag{3.8}$$

Equation (3.8) is true for q less than 50%. For q greater than 50%, we have

$$H_t = \begin{cases} 1 & \text{if } r_t \geqslant \mathrm{VaR}_q, \\ 0 & \text{otherwise.} \end{cases} \tag{3.9}$$

Under the null hypothesis of correct unconditional coverage, the test statistic is

$$\begin{aligned} -2\ln(\mathrm{LR}_{\mathrm{uc}}) = 2 - [&n_0 \ln(1-\pi_{\mathrm{exp}}) + n_1 \ln(\pi_{\mathrm{exp}}) \\ &- n_0 \ln(1-\pi_{\mathrm{obs}}) - n_1 \ln(\pi_{\mathrm{obs}})] \sim \chi_1^2, \end{aligned} \tag{3.10}$$

where n_1 and n_0 are the number of violations and nonviolations, respectively, p_{exp} is the expected proportion of exceedances and $p_{\mathrm{obs}} = n_1/(n_0 + n_1)$ is the observed proportion of exceedances. Under H_0, we have a correctly specified model; hence, we want to keep H_0. Critical levels for the test are 6.63 (1% level), 3.84 (5% level) and 2.71 (10% level).

This test, however, only tests whether the empirical frequency of hits is close to the prespecified level. It does not test whether several quantile exceedances occur in rapid succession or whether they tend to be isolated. This is an important issue because we do not want the model to underpredict or overpredict VaR in certain periods. For example, the unconditional coverage may appear correct numerically, but if all hits occur at the start of the data sample, it will overpredict the VaR at the beginning and underpredict it at the end of the data sample.

Christoffersen (1998) provided a joint test for correct coverage and detecting whether a quantile violation today influences the probability of a violation tomorrow. The test statistic is defined as follows:

$$\begin{aligned}-2\ln(\mathrm{LR}_{\mathrm{cc}}) = 2 - [&n_0 \ln(1-\pi_{\mathrm{exp}}) + n_1 \ln(\pi_{\mathrm{exp}}) - n_0 0 \ln(1-\pi_{01}) \\ &- n_{01}\ln(\pi_{01}) - n_{10}\ln(1-\pi_{10}) - n_{11}\ln(\pi_{11})] \sim \chi^2_2, \quad (3.11)\end{aligned}$$

where n_{ij} represents the number of times an observation with value i is followed by an observation with value j (1 is a hit, 0 is no hit). $P_{01} = n_{01}/(n_{00} + n_{01})$ and $P_{11} = n_{11}/(n_{11} + n_{10})$. Note that the $\mathrm{LR}_{\mathrm{cc}}$ test is only sensitive to one violation immediately followed by another, ignoring all other patterns of clustering. Under H_0, we have a correctly specified model, and, hence, we want to accept H_0. The critical levels for the test are 9.21 (1% level), 5.99 (5% level) and 4.61 (10% level). Values below the critical values indicate that H_0 can be accepted; hence, we have a correctly specified model according to both unconditional and conditional coverage.

4 EMPIRICAL RESULTS

As described above, our data set consists of 4470 daily commodity returns from August 4, 1992 to November 12, 2013 for nineteen commodities and one major commodity index. One-day in-sample VaR estimates are compared with the real observed return values for the various models at the 1%, 5%, 95% and 99% levels for all twenty commodities. After the estimation, Kupiec (1995) and Christoffersen (1998) test statistics are calculated as outlined in Section 3. Under H_0, the model is correctly specified regarding unconditional and conditional coverage. Applying the Kupiec test, H_0 is rejected when the test statistic is larger than 6.63 (1% level), 3.84 (5% level) and 2.71 (10% level). Using the Christoffersen test, the critical values are 9.21 (1% level), 5.99 (5% level) and 4.61 (10% level).

The test statistics are reported in Table 4 on the next page. With four quantiles, three risk models, two backtest methods and twenty commodities, we have 480 test statistics to be evaluated.

In Table 5 on page 74, we aggregate the number of violations of the Kupiec and Christoffersen tests (that is, where we rejected H_0/a proper VaR model specification) at different significance levels over all commodities for the different VaR models.

TABLE 4 The Kupiec and Christoffersen test statistics, August 4, 1992–November 12, 2013. [Table continues on next three pages.]

(a) Conditional Christoffersen test statistics

Model	VaR level (%)	CL1	HO1	NG1	CT1	C1	W1	S1	RR1	SM1	BO1
RiskMetrics	1	0.00	0.00	0.00	27.14	18.31	0.53	11.77	7.44	5.43	10.41
	5	2.38	2.15	8.56	17.48	7.56	3.49	4.79	11.25	12.61	5.38
	95	1.13	3.81	1.70	12.83	2.16	8.90	4.74	11.43	7.98	6.91
	99	7.47	13.32	59.74	18.16	23.01	25.72	24.23	35.45	33.48	33.48
Historical simulation	1	10.40	0.00	2.87	20.88	14.48	3.59	18.89	3.13	10.93	13.24
	5	4.67	17.55	4.29	53.40	39.10	4.57	23.17	28.66	41.06	20.13
	95	10.57	0.40	2.55	29.99	16.70	25.08	6.36	30.01	6.28	6.23
	99	6.08	0.00	2.41	12.14	10.40	15.17	1.70	11.68	2.85	0.51
Volume-adjusted QR	1	0.00	0.00	0.00	10.04	6.82	0.00	3.13	0.53	0.51	6.57
	5	3.16	0.11	0.11	15.02	5.46	0.16	2.99	18.72	14.45	1.39
	95	0.38	0.40	0.26	14.45	0.48	6.77	5.46	11.40	0.77	0.18
	99	0.00	0.00	0.51	2.87	0.00	2.87	0.00	10.64	0.58	0.58

TABLE 4 Continued.

(b) Unconditional Kupiec test statistics

Model	VaR level (%)	CL1	HO1	NG1	CT1	C1	W1	S1	RR1	SM1	BO1
RiskMetrics	1	8.95	7.43	1.47	24.02	9.76	0.12	8.95	7.43	5.40	2.68
	5	0.62	1.83	7.78	5.07	1.15	3.43	1.65	0.74	7.78	2.68
	95	0.85	3.69	1.26	0.27	1.26	0.85	0.06	1.75	6.32	5.68
	99	7.43	13.27	59.59	16.21	22.83	19.40	24.02	27.75	33.06	33.06
Historical simulation	1	0.12	0.41	0.04	0.04	0.62	0.16	2.25	0.01	0.00	0.51
	5	0.42	0.63	1.29	0.51	0.06	1.45	3.23	0.03	2.66	0.09
	95	0.63	0.00	0.86	2.55	1.29	3.47	0.74	0.51	0.99	0.01
	99	0.62	1.15	2.25	0.16	0.12	0.12	1.48	0.06	0.41	0.04
Volume-adjusted QR	1	0.01	0.00	0.04	0.07	0.01	0.07	0.01	0.00	0.04	0.00
	5	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.00	0.00	0.00
	95	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.00
	99	0.04	0.04	0.04	0.04	0.04	0.04	0.07	0.04	0.01	0.01

TABLE 4 Continued.

(a) Conditional Christoffersen test statistics

Model	VaR level (%)	LH1	FC1	HG1	GC1	SI1	CC1	KC1	OJ1	SB1	GSCI
RiskMetrics	1	2.79	27.86	16.26	25.30	36.62	11.32	19.49	47.35	16.34	0.00
	5	35.74	9.83	1.08	1.00	0.32	1.89	0.81	2.11	2.08	0.20
	95	24.25	6.62	0.23	2.03	1.94	14.56	3.46	0.29	0.30	1.38
	99	10.58	0.00	0.00	13.19	13.87	47.83	35.38	25.43	27.90	8.98
Historical simulation	1	10.31	8.11	3.59	1.32	11.68	0.00	0.00	16.67	12.14	10.64
	5	43.38	10.67	22.47	5.65	16.67	0.36	1.80	11.36	9.30	10.64
	95	11.03	11.90	2.79	11.71	1.64	18.31	13.44	10.78	8.68	3.28
	99	3.33	3.33	6.81	12.66	7.12	12.61	19.26	1.32	1.64	2.82
Volume-adjusted QR	1	16.01	0.53	0.58	0.58	12.17	9.88	0.00	12.17	12.17	0.53
	5	16.37	10.90	8.20	0.38	2.11	0.13	1.12	4.28	0.81	0.16
	95	3.10	0.48	1.24	0.77	0.13	15.07	1.32	1.39	1.39	0.55
	99	0.00	0.00	0.00	6.37	2.87	15.64	0.58	0.58	2.87	0.00

TABLE 4 Continued.

(b) Unconditional Kupiec test statistics

Model	VaR level (%)	LH1	FC1	HG1	GC1	SI1	CC1	KC1	OJ1	SB1	GSCI
RiskMetrics	1	2.68	27.75	16.21	24.02	34.45	10.59	19.40	41.72	14.22	4.21
	5	31.15	4.88	0.27	0.06	0.20	1.83	0.10	0.00	1.00	0.14
	95	23.11	4.88	0.00	0.01	1.83	1.26	1.26	0.06	0.14	0.34
	99	9.76	15.21	14.22	10.59	13.27	33.06	33.06	25.24	27.75	8.95
Historical simulation	1	4.73	6.34	0.16	0.51	0.06	0.41	0.01	5.50	0.16	0.04
	5	0.42	0.01	2.66	0.33	0.19	0.03	0.53	5.85	0.43	0.51
	95	2.01	1.63	0.03	1.13	0.03	3.40	0.99	0.14	1.63	0.73
	99	0.06	0.06	0.01	0.31	0.06	1.82	0.87	0.51	0.76	0.12
Volume-adjusted QR	1	0.00	0.00	0.01	0.01	0.04	0.01	0.04	0.04	0.04	0.00
	5	0.00	0.00	0.01	0.01	0.00	0.01	0.01	0.01	0.00	0.00
	95	0.00	0.01	0.00	0.01	0.01	0.01	0.01	0.00	0.00	0.01
	99	0.01	0.04	0.01	0.04	0.04	0.04	0.01	0.01	0.04	0.07

Critical values are 6.63 (1% level), 3.84 (5% level) and 2.71 (10% level) for the Kupiec test and 9.21 (1% level), 5.99 (5% level) and 4.61 (10% level) for the Christoffersen test. For ticker codes, see Table 1 on page 58.

TABLE 5 Number of violations (all commodities, all quantiles) at different significance levels for the (a) Kupiec and (b) Christoffersen test.

(a) Kupiec test

Model	Significance level	Number of violations	Significance level	Number of violations	Significance level	Number of violations
RiskMetrics	10%	44	5%	39	1%	30
Historical simulation	10%	4	5%	1	1%	0
Volume-adjusted QR	10%	0	5%	0	1%	0

(b) Christoffersen test

Model	Significance level	Number of violations	Significance level	Number of violations	Significance level	Number of violations
RiskMetrics	10%	50	5%	46	1%	37
Historical simulation	10%	53	5%	51	1%	43
Volume-adjusted QR	10%	23	5%	21	1%	16

Critical values are 6.63 (1% and 99% level) and 3.84 (5% and 95% level) for the Kupiec test and 9.21 (1% and 99% level) and 5.99 (5 % and 95% level) for the Christoffersen test.

For most commodities, the 5% and 95% VaR predictions from RiskMetrics perform rather well. Regarding the 1% and 99% VaR predictions, RiskMetrics appears not to be an adequate model for describing the dynamics of the left and right tails of the distributions. Only for some commodities can the left or right tails be predicted. This should come as no surprise because RiskMetrics is based on a normal distribution assumption. Even if volatility clustering is taken into account in the model, the conditional distribution, given a value for the volatility, is always normal. The descriptive statistics presented above revealed fat tails and both negatively and positively skewed return distributions. Because RiskMetrics does not allow for these empirical facts, its VaR predictions are not estimated correctly. The errors are higher the higher/lower the quantiles are that we investigate. At the 1% significance level, we have model violation in thirty out of sixty cases using the Kupiec test and fifty out of sixty cases using the Christoffersen test. At the 5% significance level, we have model violation in thirty-nine out of sixty cases under the Kupiec test and fifty out of sixty cases based on the Christoffersen test. At the 10% significance level, we have model violation in forty-four out of sixty cases for the Kupiec test and fifty out of sixty cases for the Christoffersen test. This backtesting of RiskMetrics for commodities reveals a weak performance. This is also consistent with the findings of Giot and Laurent (2003).

The historical simulation risk model performs somewhat better. The model performs very well in the unconditional coverage test. For the conditional coverage test, the performance is not at all good; in fact, the historical simulation performance is worse than that of RiskMetrics. This means that, although historical simulation does capture the right number of exceedances, there are extended periods in which the model underpredicts or overpredicts VaR. Again, this result is intuitive because historical simulation does not capture the feature of time-varying volatility. The conditional distribution is always equal to the empirical one over the whole data sample. Whether conditional volatility is high or low does not matter.[4] At the 1% significance level, we have no model violations in sixty cases using the Kupiec test and fifty-three out of sixty cases under the Christoffersen test. At the 5% significance level, we have model violation in one out of sixty cases for the Kupiec test and fifty-three out of sixty cases for the Christoffersen test. At the 10% significance level, we have model violation in four out of sixty cases for the Kupiec test and, again, fifty-three out of sixty cases for the Christoffersen test. This clearly shows that we are able to obtain the correct unconditional coverage with historical simulation, but we are by no means able to obtain adequate conditional coverage.

[4] One way of handling this would be to implement a "volatility filtered" historical simulation (see Alexander (2009) for more details and examples). This is not performed in this analysis and is left for future research.

The QR risk model performs better than both RiskMetrics and historical simulation. There are no violations in the unconditional coverage tests. This means a perfect fit with regard to backtesting the number of violations. For the unconditional tests, we have twenty-three out of sixty cases for the different significance levels. Although this is not optimal, it is less than half of the violations of the other two model alternatives. For some commodities, we have a close-to-perfect fit. The remaining problem with conditional coverage for the other commodities is a result of the changing relationship between returns and volatility over time. In the QR, we model this relationship to be stable over time, which may be too restrictive an assumption.

5 CONCLUSIONS

Correct modeling and forecasting of risk are obviously of great importance to commodity market investors and hedgers. Commodities are not a homogeneous asset class, as each commodity is driven by specific supply-and-demand conditions leading to very different dynamics and return distributions. Standard risk models such as RiskMetrics and historical simulation have important weaknesses. The former does not capture the true empirical return distribution conditional on the changing volatility. The latter captures the empirical return distribution but does not take into account time-varying volatility. Backtesting these models for nineteen different commodities and a major commodity index using more than twenty years of daily data clearly indicates that these models are problematic for modeling VaR, especially for very low and very high quantiles. A robust and easy-to-implement alternative proposed in this paper is the QR model, which involves running an exponential weighted moving average volatility model (similar to RiskMetrics) and then estimating a linear QR model using this volatility as an input. The conclusion from our empirical analysis is that the QR approach provides far better results than the two standard models for analyzing tail risk. However, there are still some problems regarding the estimation of conditional coverage for some commodities (this is particularly the case for orange juice, cotton, rice and corn). This is because of the time-varying relationship between returns and volatility for these commodities.

Regarding univariate risk modeling of commodity portfolios, the simple QR model could be extended to cover volatility-filtered historical simulation (see Alexander 2009), and a dynamic QR (DQR) model updating the return–volatility relationship over time. Taylor (2008) provided suggestions on how to progress along these lines. Another path would be to look into how VaR for commodity portfolios should be modeled in a multivariate setting. Börger *et al* (2009) and Kat and Oomen (2006, 2007b) are relevant starting references for this approach. Kuralbayeva and Malone (2012) have explored the importance of various global and commodity-specific determinants in explaining extreme movements (tails) in commodity spot prices using a QR approach. The model used in our paper could be expanded to include such explanatory factors. It is also quite possible that the quantile sensitivities may change over time. A starting point for an improved VaR model could be along the lines suggested by Taylor (2008), who used QR with exponential weighted parameters.

DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

ACKNOWLEDGEMENTS

The authors are grateful for useful and constructive suggestions and comments on an earlier version of this paper provided by this journal's referee and the participants at the FIBE conference in Bergen in January 2015.

REFERENCES

Alexander, C. (2009). *Market Risk Analysis: Volumes I, II, III and IV*, Vol. 4. Wiley.

Bodie, Z., and Rosansky, V. I. (1980). Risk and return in commodity futures. *Financial Analysts Journal* **36**(3), 27–39.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**(3), 307–327.

Börger, R., Cartea, Á., Kiesel, R., and Schindlmayr, G. (2009). Cross-commodity analysis and applications to risk management. *Journal of Futures Markets* **29**(3), 197–217.

Brooks, C., and Prokopczuk, M. (2013). The dynamics of commodity prices. *Quantitative Finance* **13**(4), 527–542.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review* **39**(4), 841–862.

Deaton, A., and Laroque, G. (1992). On the behaviour of commodity prices. *Review of Economic Studies* **59**(1), 1–23.

Dunsby, A., Eckstein, J., Gaspar, J., and Mulholland, S. (2008). *Commodity Investing: Maximizing Returns Through Fundamental Analysis*, Vol. 425. Wiley.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**(4), 987–1007.

Engle, R. F., and Granger, C. W. J. (1987). Cointegration and error correction: representation, estimation and testing. *Econometrica* **55**, 251–276.

Engle, R. F., and Manganelli, S. (2004). CAViaR: conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics* **22**(4), 367–381.

Erb, C. B., and Harvey, C. R. (2006). The strategic and tactical value of commodity futures. *Financial Analysts Journal* **62**(2), 69–97.

Fabozzi, F. J., Füss, R., and Kaiser, D. G. (2008). *The Handbook of Commodity Investing*, Vol. 156. Wiley.

Füss, R., Adams, Z., and Kaiser, D. G. (2010). The predictive power of value-at-risk models in commodity futures markets. *Journal of Asset Management* **11**(4), 261–285.

Geman, H. (2005). *Commodities and Commodity Derivatives: Modeling and Pricing for Agriculturals, Metals and Energy*. Wiley.

Giot, P., and Laurent, S. (2003). Market risk in commodity markets: a VaR approach. *Energy Economics* **25**(5), 435–457.

Gorton, G. B., and Rouwenhorst, K. G. (2004). Facts and fantasies about commodity futures. Working Paper, National Bureau of Economic Research.

Gorton, G. B., Hayashi, F., and Rouwenhorst, K. G. (2013). The fundamentals of commodity futures returns. *Review of Finance* **17**(1), 35–105.

Greer, R. J. (2000). The nature of commodity index returns. *Journal of Alternative Investments* **3**(1), 45–52.

Gregoriou, G. N., Karavas, V., Lhabitant, F.-S., and Rouah, F. D. (2011). *Commodity Trading Advisors: Risk, Performance Analysis, and Selection*, Vol. 281. Wiley.

Kat, H. M. (2006). Is the case for investing in commodities really that obvious? Report, Alternative Investment Research Center, Cass Business School, City University, London.

Kat, H. M., and Oomen, R. C. (2007a). What every investor should know about commodities: part I. *Journal of Investment Management* **5**(1), 4–28.

Kat, H. M., and Oomen, R. C. (2007b). What every investor should know about commodities: part II, multivariate return analysis. *Journal of Investment Management* **5**(1), 40–64.

Koenker, R., and Bassett, G. Jr. (1978). Regression quantiles. *Econometrica* **46**, 33–50.

Kroner, K. F., Kneafsey, K. P., and Claessens, S. (1995). Forecasting volatility in commodity markets. *Journal of Forecasting* **14**(2), 77–95.

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives* **3**(2), 73–84.

Kuralbayeva, K., and Malone, S. W. (2012). The determinants of extreme commodity prices. Annual Meetings Paper, Midwest Finance Association.

Lambert, P., and Laurent, S. (2001). Modelling financial time series using GARCH-type models with a skewed student distribution for the innovations. Discussion Paper 0125, Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

Miffre, J. (2011). Long-short commodity investing: implications for portfolio risk and market regulation. Working Paper, EDHEC-Risk Institute.

Pindyck, R. S. (1994). Inventories and the short-run dynamics of commodity prices. *RAND Journal of Economics* **25**(1), 141–159.

Steen, M., and Gjøberg, O. (2013). Are commodity markets characterized by herd behaviour? *Applied Financial Economics* **23**(1), 79–90.

Steen, M., and Gjølberg, O. (2014a). Working Paper, Norwegian Center for Commodity Market Analysis, Norwegian University of Life Sciences.

Steen, M., and Gjølberg, O. (2014b). Warmer, wetter, wilder? Weather evidence from the grain markets 1960-2014. Working Paper, Norwegian Center for Commodity Market Analysis, Norwegian University of Life Sciences.

Taylor, J. W. (2008). Using exponentially weighted quantile regression to estimate value at risk and expected shortfall. *Journal of Financial Econometrics* **6**(3), 382–406.