Modeling Price Volatility Based on a Genetic Programming Approach

Shusheng Ding, Yongmin Zhang and Meryem Duygun

Business School, Ningbo University, China, and 1 Nottingham University Business School, University of Nottingham, UK

Corresponding author email: zhangyongmin@nbu.edu.cn

Business profitability is highly dependent on risk management strategies to hedge future cash flow uncertainty. Commodity price shocks and fluctuations are key risks for companies with global supply chains. The purpose of this paper is to show how artificial intelligence (AI) techniques can be used to model the volatility of commodity prices. More specifically, the authors introduce a new model – LIQ-GARCH – that uses genetic programming to forecast volatility. The newly generated model is then used to forecast the volatility of the following three indexes: the Commodity Research Bureau (CRB) index, the West Texas Intermediate (WTI) oil futures prices and the Baltic Dry Index (BDI). The empirical model performance tests show that the newly generated model in this paper is considerably more accurate than the traditional GARCH model. As a result, this model can help businesses to design optimal risk management strategies and to hedge themselves against price uncertainty.

Introduction

Among the several functions of a company, risk management is quite an important one, as it contributes directly to value creation (Allayannis and Weston, 2001; Lewellen, 2006), where volatility forecasting becomes exceedingly relevant (Christoffersen and Diebold, 2000). Financial hedging (or hedging henceforth) is the main strategy used by businesses to reduce the adverse impact of price fluctuations on profit margins (Gordon, Loeb and Tseng, 2009; Liebenberg and Hoyt, 2003). Meanwhile, successful hedging strategies rely on the capability of the company to forecast (with some degree of accuracy) the future volatility of commodity prices.

In spite of its importance, this is an area of risk management where artificial intelligence (AI) techniques have not been widely used. In the current finance literature, generalized autoregressive conditional heteroskedasticity (GARCH) models are widely used to model the volatility of financial time series (e.g. Kambouroudis, McMillan and Tsakou, 2016; Prokopczuk and Simen, 2014). However, existing GARCH models suffer from three limitations: (a) they assume that the volatility of the market and correlations among assets change slowly or not at all; (b) they cannot control for extreme events; and (c) they cannot include all the information from the market. This point is particularly relevant when trying to control for the liquidity position in the market, as incorporating liquidity into GARCH models is necessary so that forecasts can be updated regularly using all the information extracted from traders’ expectations on future supply and demand conditions.

The purpose of this paper is to show how AI techniques can be used to model the volatility of commodity prices; more specifically we propose a new estimator of the GARCH model that allows the incorporation of existing information on the market liquidity in the estimation procedure.
The proposed estimator – named LIQ-GARCH – uses a genetic programming (GP) approach to the model estimation that has been widely used to predict stock returns (Manahov, Hudson and Hoque, 2015) and energy consumption (Castelli, Vanneschi and De Felice, 2015). Our estimator is compared with the standard GARCH estimator for three time series: namely, the Commodity Research Bureau (CRB) index, the West Texas Intermediate (WTI) oil futures prices and the Baltic Dry Index (BDI). These three time series have been chosen because they are extensively used by businesses around the world to hedge price risk along the supply chain. Our model exhibits overwhelmingly superior performance in forecasting the volatility of the three time series against the standard GARCH model.

The contribution of the paper is the following. First, to the best of our knowledge, our LIQ-GARCH estimator is the first to extend the prevalent GARCH estimator by including market liquidity information. Second, our paper shows how AI techniques can be used to improve forecasting models and support the development of risk management strategies when dealing with fluctuating commodity prices.

The remainder of the paper is organized as follows. The next section gives a short summary of the managerial implications of the model. The third section reviews the existing evidence on the relationship between volatility, liquidity and business performance. The fourth section introduces the LIQ-GARCH estimator, while the fifth section illustrates the data and the variables used in the empirical analysis. The sixth section presents the results, while discussing the possible uses of the new estimators to manage risk along the supply chain. The final section offers some concluding remarks.

Managerial implications of LIQ-GARCH model

Given that risk management constitutes an important aspect of company management, corporate hedging clearly plays a vital role in company management, since it can help to mitigate firms’ risk bearing (Allayannis and Weston, 2001; Lewellen, 2006). In fact, it is well documented that corporate hedging can increase firm value and improve business performance (Bessembinder, 1991; Smith and Stulz, 1985). Existing studies have unveiled a positive relation between derivatives hedging and firm performance (see Bartram et al., 2011; Haushalter et al., 2007; Pérez-González and Yun, 2013). More recently, Chen, Han and Zeng (2017) also provide robust evidence in support of the view that companies who use derivative hedging achieve higher returns than non-users do.

Enhanced risk management can act as one possible assistance toward the management challenge of performance improvement in a highly volatile environment mentioned in Parnell et al. (2012). More importantly, accurate volatility forecasting can aid managers’ proactive management against predicted risk (Jung, Lim and Oh, 2011). Recently, big data analytics has also proven to be a useful tool in enhancing risk and operations management (see Cerchiello and Giudici, 2016; Choi, Chan and Yue, 2017; Choi, Wallace and Wang, 2018).

Our LIQ-GARCH model is based on two key ingredients: big data and AI technologies are capable of analyzing high-frequency data and liquidity information that can capture the supply and demand conditions of a good. The managerial implications of our model can be summarized in the following two aspects. On the one hand, our model can deliver a more accurate view of price volatility of a good at a higher predictive frequency than traditional models. Managers can make more robust, cost-saving hedging decisions using our model. Since hedging can be costly when volatility is negligible, it is best not to take hedging positions when volatility is low and vice versa. Therefore, our model could support managers in making decisions, i.e. whether and when taking hedging positions is necessary, so as to save significant costs in risk and operations management. On the other hand, because our model delivers a more accurate view on supply and demand conditions of a good via liquidity information, managers can more effectively manage inventory of goods for both input and output channels to suit prevailing market conditions better. Therefore, our model can also help managers to improve the inventory management and avoid the disruption of critical supply chain networks.

Financial hedging and business performance: a review

Traditionally, risk management has relied on a mixture of quantitative techniques and expert
judgments where accounting and planning for liquidity shocks have been handled indirectly through scenario planning, risk budgeting and portfolio theory. However, risk management has recently started to benefit from AI with the result that the traditional quantitative techniques used for risk management have started to be replaced by a variety of analytical techniques. These new techniques are particularly relevant to businesses with complex supply chains spanning several countries. While playing a vital role in fostering international trade and economic growth, global supply chains create new risks as well: indeed, in a world where markets are highly integrated, minor supply chain disruptions can have major impacts on the performance of the supply chain as markets react to negative shocks with increased speed and volatility (Tummala and Schoenherr, 2011).

Among the many risks that may affect the performance of supply chain, volatility of the commodity prices is a key one. Indeed volatile commodity prices cause fluctuations in the cost of raw materials that, if not properly managed, can adversely affect profit margins. It is well known that, in some industries, the exposure to the commodity price risk exposure is quite substantial, as in gold price for mining companies (Tufano, 1998) and non-energy commodity price for automobile companies (Oxelheim and Wihlborg, 1995). The volatility of commodity prices can be detrimental to some companies. For instance, Ford Motor Co. wrote off US$1bn value of its metal reserves in 2002 because of the unexpectedly sharp decrease in the metal price (White, 2002). In addition, extreme volatility might result in bankruptcy even for well-capitalized companies (Bessembinder and Lemmon, 2002).

Businesses usually manage their risk by taking hedge positions. In the automotive industry, the biggest risk is the volatility of metal price, which is often hedged against commodity futures and options. Oil and airline companies have substantial risk exposures to oil price fluctuations (Jin and Jorion, 2006; Phan, Nguyen and Faff, 2014). Oil futures are often used by transport and power utility companies to hedge against the risk of oil price fluctuations. Carter, Rogers and Simkins (2006) quantify the ‘hedging premium’ for airline companies and find that jet fuel hedging can reduce the underinvestment costs for airline companies. Furthermore, Mohanty et al. (2014) show that oil price volatility significantly influenced a number of other industries, including airlines, recreational services, restaurants and bars. As a result, accurate forecasting of oil price volatility is of great importance.

Shipping companies need to hedge against the freight rate volatility as well (Zhang and Shen, 2016, 2017). Therefore, forecasts of shipping index plays a vital role in the management of shipping companies (Duru, 2010). Dry bulk freight futures contracts, which are traded on the International Maritime Exchange (Prokopczuk, 2011), are useful risk mitigation strategies for shipping companies. Samitas and Tsakalos (2010) provide supporting evidence that the use of derivatives hedging can minimize shipping firms’ risk exposure and ensure their growth. However, the forecasting is still challenging, owing to the complexity of the bulk shipping market, especially for precise predictions (Goulielmos and Psifia, 2013).

When the expected volatility of prices is large, firms tend to increase their hedging positions to counterbalance the adverse effect of large future price swings. The increased hedging positions could help firms to limit their future losses and, as a result, the risk attached to cash flow volatility can be reduced. In this way, the firm can reduce the probability of financial distress and increase the financial flexibility (Gao et al., 2015). When expected volatility is negligible, the price will remain stable or follow historical trends. In this case, price swings are predictable. As a result, it would be costly for companies to hold option positions (Howard and D’Antonio, 1994). Companies then need not buy options or take futures positions to hedge future price uncertainty and can thereby reduce costs from decreasing their hedging positions when the future volatility tends to be trivial.

The contribution of hedging to value creation and improved business performance is well established (Bessembinder, 1991; Smith and Stulz, 1985). Several studies have shown that there exists a positive relationship between derivatives hedging and firm performance (see Allayannis et al., 2012; Bartram, Brown and Conrad, 2011; Haushalter et al., 2007; Pérez-González and Yun, 2013). More recently, Lau (2016) has shown that hedging can strengthen company’s ROA and ROE, while Chen, Han and Zeng (2017) find that hedging companies announce higher returns than non-users. Typically, hedging strategies are supported by a variety of econometric models aiming at forecasting the volatility of commodity prices. In spite of the
fact that they are widely used for this purpose, they suffer from a variety of limitations. GARCH models are not designed to handle systemic changes caused by jumps in the availability of liquidity or changes in the market micro-structure.\(^1\) For instance, information on the liquidity of the market is quite important for estimating volatility of prices over time accurately, as the degree of liquidity in a market is very informative of the traders’ expectations on future demand and supply on the market (Easley et al., 1996; Welker, 1995). The relationship between liquidity and volatility has been widely analyzed by several authors. Fleming and Remolona (1999) have investigated the relationship among liquidity, volatility and public information in the US Treasury market and have shown that that volatility and liquidity respond simultaneously to the release of new information. More recently, Collin-Dufresne and Fos (2016) have explored the relationship between liquidity and noise trading volatility and found that liquidity is an important driver of trading volatility. In addition, it has been found that the variation in the market liquidity (so called liquidity risk) is correlated with the informational content of the prices (Ng, 2011). Recent studies such as Zhang, Ding and Scheffel (2018) and Zhang and Ding (2018) have shown the significant liquidity effect on price volatility in commodity markets.

### Using GP to forecast volatility

In econometric and financial theories, volatility measures the variation degree of a price series \(\{P_t, t = 1, 2, 3 \ldots\}\) over time, where the standard deviation is usually used as a proxy of volatility. We define return \(r_t\) as:

\[
r_t = \ln \left( \frac{P_t}{P_{t-1}} \right), \quad t = 1, 2, \ldots
\]

Consider a conditional normal distributed residual model with time-varying volatility \(\sigma_t\):

\[
r_t = \varphi + \varepsilon_t \quad \varepsilon_t | I_{t-1} \sim N(0, \sigma_t^2) \quad (1)
\]

where \(I_{t-1}\) represents the information available at time \(t-1\), and \(\varphi\) is the long run mean of the return series. The ARCH model specifies the conditional volatility \(\sigma_t\) that satisfies:

\[
\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2, \quad \alpha_0 > 0, \quad 0 < \alpha_1 < 1
\]

A generalized ARCH model is denoted as GARCH (1,1), which can be described as:

\[
\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \sigma_{t-1}^2 \quad (2)
\]

where \(\alpha_0 > 0, \alpha_1 + \alpha_2 < 1\) for stationarity of the above process.

If the functional form of equation (2) is unknown, the estimator based on GP can be used, as it does not require assumptions on the functional form of the equation to be estimated. In AI, genetic programming is a technique whereby programs are encoded as a set of genes that are then modified using an evolutionary algorithm – it is an application of genetic algorithms where the possible solutions consist of computer programs (Hirsh et al., 2000; Rasheed, Hirsh and Gelsey, 1997). The methods used to encode a computer program in an artificial chromosome and to evaluate its fitness with respect to the predefined task are central to the GP technique. In addition, it is well suited to working with high-dimensional data (Viegas et al., 2018). Genetic programming can be viewed as an extension of the ‘genetic algorithm’, a model for testing and selecting the best choice among a set of results. Genetic programming makes the program or ‘function’ the unit that is tested. Our GP estimator works as follows: it first generates a random population of functions, and then it evaluates the quality (fitness) of each individual function by evaluating the difference between the generated function and the targeted function. Next, one or two function(s) will be probabilistically selected based on their fitness in order to participate in the genetic operations. Normally, there are two genetic operations, one is called crossover and another is called mutation. The crossover operation is used to create a new function (called offspring) by randomly choosing some subitems from two selected functions (called parents, which are usually polynomials) and recombining the subitems from the two functions together. The mutation operation is used to create a new offspring by choosing some random subitems from one selected function and altering them. After new individuals are

---

\(^1\)Market microstructure is the process by which investors’ demands and expectations can be ultimately translated into asset prices and trading volumes (Garman, 2015; Madhavan, 2015).
created, their fitness will be calculated again, and genetic operations will also be performed again to evaluate the newly generated functions. This whole process is repeated until an acceptable solution is found or another termination criterion is satisfied (usually up to some certain number of generations). The best individual function will be returned as the solution.

Our starting function is as follows:

\[
 f(L_{i-1}^2, \varepsilon_{i-1}^2, \sigma_{i-1}^2) = \sigma_i^2 
\]

with the following objective function:

\[
 \min \sum_{t=0}^T [f(L_{i-1}^2, \varepsilon_{i-1}^2, \sigma_{i-1}^2) - \sigma_i^2] 
\]

and the objective function is subject to:

\[
 f(L_{i-1}^2, \varepsilon_{i-1}^2, \sigma_{i-1}^2) \geq 0 
\]

where \( L_{i}^2, \varepsilon_{i}^2, \sigma_{i}^2 \) are the squared liquidity, the squared residuals and the realized variance at time \( t \), respectively. Using the settings and the procedural of GP detailed in Appendix 1, we ran our GP system 50 times. Eventually, the GP procedure generates the following model:

\[
 \sigma_i^2 = \alpha_0 + \alpha_1 \sigma_{i-1}^2 + \alpha_2 X_{i-1} 
\]

where \( X_i = (1 - L_i^2) \ast (L_i^2 - \sigma_i^2 - L_i^2) \), and \( L \) is the market liquidity. We name the new model LIQ-GARCH \((1, 1)\), i.e. liquidity-adjusted GARCH model.

### Data, variables and the empirical methodologies

#### Data and variables

For our empirical analysis, we use three time series: the CRB index, the WTI oil futures prices and the BDI. All the indexes are observed daily, although over different periods of time. The sample period of CRB index runs from 1 January 1995 to 30 November 2017, the sample period of the WTI index is from 1 January 2000 to 30 November 2017, and the sample period for BDI is from 1 January 1990 to 30 November 2017. The total number of observations is 17,606, and it is a large enough data set to illustrate our GP procedure.

Empirically, volatility measures are based on market returns, and returns can be defined through the market price series, which is \( r_t = \ln(\frac{P_t}{P_{t-1}}) \). Therefore, we estimated the volatility of the three time series via the sample return standard deviation (Christensen and Prabhala, 1998), which can be defined as:

\[
 \sigma_t = \sqrt{\frac{1}{T-1} \sum_{i=1}^T (r_{t-i} - \bar{r}_t)^2} 
\]

where \( r_{t-i} = \ln(P_{t-i}/P_{t-i-1}) \), with \( P_{t-i} \) representing the CRB index, WTI oil price and BDI, respectively, at day \( t-i \), and \( \bar{r}_t = \frac{1}{T} \sum_{i=1}^T r_{t-i} \). We take \( T = 21 \) as the monthly rolling average. The returns and volatilities of all three indexes are summarized in Table 1, and Figure 1 plots the three indices. The realized volatility of the three series is presented in Figure 2.

Finally, we estimate the liquidity of the three markets. We adopt a widely used proxy for liquidity: the bid–ask spread (BAS), which is positively correlated with price volatility in financial markets (Bollerslev and Melvin, 1994; Wang and Yau, 2000). In our paper, we use the effective spread estimator developed by Roll (1984) and used in a number of financial papers such as Goyenko, Holden and Trzcinka (2009) and Corwin and Schultz (2012). The proxy uses the autocovariance of the daily price changes as an effective measure of the BAS. Roll’s (1984) starting point is that the
traded assets have fundamental value $V_t$ (Lux and Marchesi, 1999):

$$V_t = V_{t-1} + \eta_t$$  \hspace{1cm} (6)

where $\eta_t$ reflects new information arrival, which is assumed to be independent of the previous period information under the efficient market hypothesis. Next, Roll (1984) denotes $S_t$ as the last observed trade price on day $t$ and assumes that $S_t$ follows the following process:

$$S_t = V_t + \frac{1}{2} E Q_t$$  \hspace{1cm} (7)

where $E$ is the effective spread, and $Q_t$ is a buy/sell indicator for the last trade that equals +1 for a buy and -1 for a sell. He further assumes that $Q_t$ is equally likely to be +1 or -1, and $Q_t$ is also serially uncorrelated, and independent of $\eta_t$. Then he takes the first difference of equation (7) and plugs in the result from equation (6), which yields:

$$\Delta S_t = \frac{1}{2} E \Delta Q_t + \eta_t$$  \hspace{1cm} (8)

where $\Delta Q_t$ measures the change of the order type from two consecutive days, and $\Delta$ is the change operator: namely, $\Delta Q_t = Q_t - Q_{t-1}$ (Goyenko, Holden and Trzcinka, 2009).
As a result,\

\[
\text{Cov}(\Delta S_t, \Delta S_{t-1}) = \frac{-1}{4} E^2, \quad \text{or equivalently,} \\
\text{spread} = 2\sqrt{-\text{Cov}(\Delta S_t, \Delta S_{t-1})}
\]

However, as the autocovariance is positive, the formula is undefined. We therefore use a modified version of the Roll estimator (Goyenko, Holden and Trzcinka, 2009):

\[
l = \text{spread} = \begin{cases} 
2\sqrt{-\text{Cov}(\Delta S_t, \Delta S_{t-1})}, & \text{Cov}(\Delta S_t, \Delta S_{t-1}) \leq 0 \\
0, & \text{Cov}(\Delta S_t, \Delta S_{t-1}) > 0 
\end{cases}
\]

For one particular day’s liquidity, it is effectively the average of the previous month’s (the past 21 days) liquidity measures. If \( L_t \) is the rolling average of liquidities from the past 21 trading days, this is then equal to:

© 2019 British Academy of Management.
Table 2. Regression results for (a) CRB series for the whole sample period (1 January 1995 to 30 November 2017), (b) WTI oil series for the whole sample period (1 January 2000 to 30 November 2017) and (c) BDI series for the whole sample period (1 January 1990 to 30 November 2017).^a

|        | (a) σ_t^2 | (b) σ_t^2 | (c) σ_t^2 |
|--------|-----------|-----------|-----------|
| σ_{t-1}^2 | 0.97***   | 0.98***   | 0.99***   |
| X_{t-1}   | 6.77E-10***| 3.31E-10**| -8.05E-13*|
| F-value   | 99999.99***| 90988.86***| 17311.68***|

^a The result confirms the marvelously fitness of LIQ-GARCH model via statistically significant coefficients of both independent variables as well as tremendous F-value. *, **, *** indicate statistical significance at 10%, 5% and 1% levels, respectively.

\[ L_t = \frac{1}{21} \sum_{i=1}^{21} l_{t-i}, \]

where \( l_{t-i} \) is the liquidity measure at day \( t-i \).

The full sample regression results of the LIQ-GARCH model in equation (4) are reported in Table 2(a), (b) and (c) for CRB, WTI oil and BDI series, respectively. From this table, it is observable that the LIQ-GARCH model is occupied with outstanding data fitness ability, since all independent variable coefficients are statistically significant at the 10% level with a huge F-value for all three series.

### Empirical methodologies

To test the performance of our proposed LIQ-GARCH (1,1) model against the standard GARCH (1,1) model, we use three different empirical methodologies. First, we use the full sample data to estimate the model parameters and compute the model’s fitting errors for each index. Under the second methodology, we estimate parameters for each year included in our sample. Then, we forecast the one-day-ahead volatility for each year. Finally, we compare the in-sample and out-of-sample volatility forecasts with the LIQ-GARCH model.

| Year   | GARCH(MSE) | LIQ-GARCH(MSE) | Improvement rate (%) | P-value |
|--------|------------|-----------------|----------------------|---------|
| Average| 4.08E-08   | 2.17E-08        | 46.72                | 0       |
| 2000   | 3.34E-08   | 1.91E-08        | 42.67                | 0.06    |
| 2001   | 2.90E-08   | 1.93E-08        | 33.38                | 0.04    |
| 2002   | 1.93E-08   | 1.92E-08        | 0.23                 | 0.45    |
| 2003   | 3.16E-08   | 1.92E-08        | 39.21                | 0.05    |
| 2004   | 2.76E-08   | 1.91E-08        | 30.66                | 0.09    |
| 2005   | 2.29E-08   | 1.63E-08        | 28.59                | 0.06    |
| 2006   | 2.37E-08   | 1.92E-08        | 18.79                | 0.07    |
| 2007   | 2.69E-08   | 1.96E-08        | 27.06                | 0.01    |
| 2008   | 2.38E-07   | 7.19E-08        | 69.83                | 0       |
| Average| 5.03E-08   | 2.48E-08        | 32.27                | 0       |
| 2009   | 8.39E-08   | 1.07E-09        | 98.73                | 0       |
| 2010   | 1.24E-08   | 1.59E-10        | 98.72                | 0       |
| 2011   | 1.77E-08   | 1.01E-09        | 94.28                | 0       |
| 2012   | 3.94E-09   | 1.44E-10        | 96.35                | 0       |
| 2013   | 1.00E-09   | 3.96E-11        | 96.04                | 0       |
| 2014   | 2.31E-09   | 7.97E-11        | 96.55                | 0       |
| 2015   | 1.80E-08   | 2.36E-10        | 98.69                | 0       |
| 2016   | 1.00E-08   | 1.60E-10        | 98.40                | 0       |
| 2017   | 1.54E-09   | 2.85E-11        | 98.14                | 0       |
| Average| 1.68E-08   | 3.25E-10        | 97.32                | 0       |

^a This table presents the volatility prediction results comparison for GARCH model and LIQ-GARCH model regarding the CRB index. MSE stands for the mean squared error and the improvement rate is defined as \( \frac{\text{MSE}_{\text{GARCH}} - \text{MSE}_{\text{LIQ-GARCH}}}{\text{MSE}_{\text{GARCH}}} \). P-values are the paired test results between MSE of GARCH model and MSE of LIQ-GARCH model. The full sample period is from 1 January 1995 to 30 November 2017 with in-sample test method employed. The in-sample test period is 2000–2008, and the tested year is the same as the sample year within the period. The out-of-sample test period is 2009–2017, and the sample year is one year ahead of the tested year within the period. For all sample volatility forecasting, we use one-day-ahead prediction approach. The yearly t-statistic is achieved by comparing the daily data of two series: namely, the GARCH and LIQ-GARCH estimated volatility.

© 2019 British Academy of Management.
Table 4. Model comparison in forecasting volatility for the time series of WTI oil price

| Year   | GARCH(MSE)    | LIQ-GARCH(MSE)    | Improvement rate (%) | P-value |
|--------|---------------|-------------------|----------------------|---------|
| Yearly sub-sample in-sample fitting | 1.75E-06        | 3.82E-08          | 97.81                | 0       |
| 2000   | 2.16E-06       | 1.20E-08          | 99.44                | 0       |
| 2001   | 3.44E-07       | 2.20E-08          | 93.60                | 0       |
| 2002   | 2.65E-07       | 5.41E-09          | 97.96                | 0       |
| 2003   | 2.95E-07       | 1.86E-08          | 93.69                | 0       |
| 2004   | 2.70E-07       | 1.50E-08          | 94.44                | 0       |
| 2005   | 3.20E-07       | 2.95E-08          | 90.78                | 0       |
| 2006   | 2.90E-07       | 2.76E-08          | 90.48                | 0       |
| 2007   | 3.21E-07       | 7.82E-09          | 97.56                | 0       |
| 2008   | 2.12E-06       | 4.32E-07          | 79.62                | 0.02    |
| Average| 7.09E-07       | 6.33E-08          | 93.07                |         |

Out-of-sample: prediction based on previous year fitting

| Year   | GARCH(MSE)    | LIQ-GARCH(MSE)    | Improvement rate (%) | P-value |
|--------|---------------|-------------------|----------------------|---------|
| 2009   | 2.50E-06       | 4.20E-07          | 83.20                | 0.01    |
| 2010   | 3.70E-07       | 9.70E-08          | 73.78                | 0       |
| 2011   | 1.05E-06       | 6.42E-08          | 93.89                | 0       |
| 2012   | 4.20E-07       | 2.07E-08          | 95.07                | 0       |
| 2013   | 4.22E-08       | 4.41E-09          | 89.55                | 0       |
| 2014   | 6.60E-07       | 2.38E-08          | 96.39                | 0       |
| 2015   | 2.22E-06       | 2.44E-08          | 98.90                | 0       |
| 2016   | 2.53E-06       | 1.12E-08          | 99.56                | 0       |
| 2017   | 9.41E-08       | 9.88E–10          | 98.95                | 0       |
| Average| 1.10E-06       | 7.41E-08          | 92.14                |         |

This table presents the volatility prediction results comparison for GARCH model and LIQ-GARCH model regarding the WTI oil futures. MSE stands for the mean squared error and the improvement rate is defined as: $\frac{\text{MSE}_{\text{LIQ-GARCH}} - \text{MSE}_{\text{GARCH}}}{\text{MSE}_{\text{GARCH}}}$. P-values are the paired test results between MSE of GARCH model and MSE of LIQ-GARCH model. The full sample period is from 1 January 2000 to 30 November 2017 with in-sample test method employed. The in-sample test period is 2000–2008, and the tested year is the same as the sample year within the period. The out-of-sample test period is 2009–2017, and the sample year is one year ahead of the tested year within the period. For all sample volatility forecasting, we use one-day-ahead prediction approach. The yearly t-statistic is achieved by comparing the daily data of two series: namely, the GARCH and LIQ-GARCH estimated volatility data can be quite noisy (Bollerslev, Patton and Quaedvlieg, 2016; Pong et al., 2004). The MSE can be defined as:

$$\text{MSE} = \frac{1}{T} \sum_{t=1}^{T} (\text{Observed}_t - \text{Predicted}_t)^2$$

where $T$ represents the number of observations embedded in the forecasting period, $\text{Observed}_t$ presents the observed variance from the market, and $\text{Predicted}_t$ presents the variance predicted from the models. Under the first methodology, MSE is calculated as the average during the full sample period while, under the second and third methodologies, MSE is calculated as the average during the specific year. We denote GARCH (MSE), LIQ-GARCH (MSE) as the MSEs for GARCH (1, 1) and LIQ-GARCH (1, 1), respectively. We also define the improvement rate for the LIQ-GARCH model compared with the GARCH model as:

© 2019 British Academy of Management.
Table 5. Model comparison in forecasting volatility for the time series of BDI

| Year                | GARCH(MSE)  | LIQ-GARCH(MSE) | Improvement rate (%) | P-value |
|---------------------|-------------|----------------|----------------------|---------|
| Full-sample (1995–2017) | 1.39E-07    | 5.61E-08       | 59.66                | 0       |
| Yearly sub-sample in-sample fitting |             |               |                      |         |
| 2000                | 8.72E-09    | 1.13E-10       | 98.70                | 0       |
| 2001                | 7.53E-09    | 1.14E-10       | 98.49                | 0       |
| 2002                | 7.42E-09    | 1.10E-10       | 98.52                | 0       |
| 2003                | 8.93E-09    | 2.78E-10       | 97.98                | 0.04    |
| 2004                | 2.15E-08    | 4.59E-10       | 97.86                | 0.04    |
| 2005                | 2.99E-08    | 1.10E-09       | 96.33                | 0       |
| 2006                | 6.74E-08    | 6.44E-08       | 4.51                 | 0.08    |
| 2007                | 5.48E-08    | 7.02E-10       | 98.72                | 0       |
| 2008                | 3.03E-07    | 7.48E-09       | 97.53                | 0.06    |
| Average             | 5.66E-08    | 8.31E-09       | 87.50                |         |
| Out-of-sample: prediction based on previous year fitting |             |               |                      |         |
| 2009                | 3.93E-06    | 1.44E-07       | 96.34                | 0       |
| 2010                | 4.49E-07    | 2.71E-08       | 93.95                | 0       |
| 2011                | 1.46E-07    | 1.12E-08       | 92.32                | 0       |
| 2012                | 4.02E-07    | 1.91E-08       | 95.26                | 0       |
| 2013                | 5.48E-07    | 2.52E-08       | 95.40                | 0       |
| 2014                | 8.20E-07    | 1.09E-07       | 86.73                | 0       |
| 2015                | 7.99E-07    | 5.06E-08       | 93.66                | 0       |
| 2016                | 8.56E-07    | 2.75E-08       | 96.79                | 0       |
| 2017                | 5.81E-08    | 1.84E-09       | 96.83                | 0       |
| Average             | 8.89E-07    | 4.61E-08       | 94.14                |         |

This table presents the volatility prediction results comparison for GARCH model and LIQ-GARCH model regarding the BDI. MSE stands for the mean squared error and the improvement rate is defined as \( \frac{MSE_{GARCH} - MSE_{LIQ-GARCH}}{MSE_{GARCH}} \). P-values are the paired test results between MSE of GARCH model and MSE of LIQ-GARCH model. The full sample period is from 1 January 1990 to 30 November 2017 with in-sample test method employed. The in-sample test period is 2000–2008, and the tested year is the same as the sample year within the period. The out-of-sample test period is 2009–2017, and the sample year is one year ahead of the tested year within the period. For all sample volatility forecasting, we use one-day-ahead perdition approach. The yearly \( t \)-statistic is achieved by comparing the daily data of two series: namely, the GARCH and LIQ-GARCH estimated volatility.

\[ \text{Improvement rate} = \frac{GARCH (MSE) - LIQ - GARCH (MSE)}{GARCH (MSE)} \]

Empirical results

Tables 3–5 show the values of the three MSEs indicators for the three time series. More specifically, Table 3 refers to the CBR index, while Tables 4 and 5 report the results for the WTI oil futures prices and for the BDI series, respectively. In general, our model outperforms the GARCH model in all cases. In the case of the CRB index, the improvement rate is around 46%, and it is statistically significant when the full sample is used for the estimation. Moreover, the improvement rate is around 32% on average when the sub-sample (2000–2008) is used. In the case of the out-of-sample forecasts (i.e. 2009–2017), our model dominates the GARCH model with a 97% improvement rate.

In the case of the WTI oil futures prices, our model outperforms the GARCH model, and the (statistically significant) improvement rate is around 97%. In addition, the improvement rate is around 93% on average in the case of the in-sample test during the period from 2000 to 2008. In the case of the out-of-sample forecast, our model outperforms the GARCH model with a 92% improvement rate.

Finally, in the case of the BDI series, our model outperforms the GARCH model: the improvement rate is around 60%, and the result is statistically significant. In addition, the improvement rate is around 88% on average for the in-sample test during the period from 2000 to 2008. In the case of the out-of-sample forecasts, our model has a 94% improvement rate. Each year, our model outperforms the GARCH model.

The accuracy of the new LIQ-GARCH model compared with the standard GARCH model can be explained as follows. First, GP is a flexible
analytical technique that can search for the best general functional form with best fitting to the data. Second, the LIQ-GARCH model uses the liquidity variable to predict the volatility. These results point out that liquidity plays a significant role in forecasting volatility as liquidity captures the supply and demand dynamics in the market, which are driving price volatility.

Conclusions

Among the many risks that may affect the performance of the supply chain, volatility of the commodity prices is a key one. Indeed the cost of raw materials can fluctuate as a result of the volatile commodity prices. While hedging is the main mitigation strategy used by businesses to reduce the adverse impact of volatile commodity prices, successful hedging strategies rely on the capability of the business to forecast the future volatility of commodity prices in such a way that all the information provided by the market is used.

This paper has proposed a new GARCH model that uses AI techniques to model the volatility of commodity prices and incorporate existing information on the market liquidity in the estimation procedure. The proposed estimator is compared with the standard GARCH estimator for three time series: namely, the CRB index, the WTI oil futures prices and the BDI. Our model exhibits overwhelmingly superior performance in forecasting the volatility of the three time series against the standard GARCH model.

Our paper adds to the existing literature in several ways. First, to the best of our knowledge, our LIQ-GARCH estimator is the first to extend the prevalent GARCH estimator by including market liquidity information. Second, our paper shows how AI techniques can be used to improve forecasting models and support the development of risk management strategies when dealing with fluctuating commodity prices.

Our model exhibits overwhelmingly superior forecasting performance, with the improvement rate round 90% for both in-sample and out-of-sample tests compared with the GARCH model. This model can be used to develop optimal hedging strategies. Indeed, firms can increase their hedging positions to stabilize future cash flows if expected volatility is predicted to be large. Conversely, firms may reduce their hedging positions to save hedging costs if future volatility is expected to be negligible.

Appendix

The parameters of our GP system are as follows:

**Terminal set:** $L_{t-1}^2$, $e_{t-1}^2$, $\sigma_{t-1}^2$

**Function set:** $+,-, \times$

**Fitness measure:** the difference between the value of the individual function and the corresponding desired output $\sigma_t^2$

**GP parameters:** population = 10000, the maximum length of the program = 1000 (i.e. up to 1000 subitems within one polynomial function), probability of crossover operation = 0.8 (i.e. 80% of population functions will be mixed with other functions to generate new functions) and probability of mutation operation = 0.1 (i.e. 10% of population functions will be mutated to generate new functions)

**Termination criterion:** when the fitness measure reaches 0 or the system runs up to 100 generations, the system will terminate.

The detailed procedural for GP is provided as follows:

1. **Initialisation:** Initialise the population of the first generation;
2. **while** not find the “good enough” forecasted volatility function or not reach the maximum number of generations;
3. **do**
4. **for** each individual volatility function in the generation **do**
5. **Evaluation:** Evaluate each volatility function’s fitness;
6. **Select Parents:** Select the individual volatility functions from the population of the current generation to breed;
7. **Crossover:** Pair the selected parents up to produce offspring volatility functions;
8. **Mutation:** Randomly alter the volatility function with a given probability;
9. **Elitism:** Select the best volatility function from the population of the current generation and insert it into the next generation;
10. **Update Population:** Update the population of the current generation;

**Algorithm 1:** GP system for volatility forecasting

References

Allayannis, G., U. Leland and D. P. Miller (2012). ‘The use of foreign currency derivatives, corporate governance, and firm value around the world’, *Journal of International Economics*, 87, pp. 65–79.

Allayannis, G. and J. Weston (2001). ‘The use of foreign currency derivatives and firm market value’, *Review of Financial Studies*, 14, pp. 243–276.

Bessembinder, H. (1991). ‘Forward contracts and firm value: investment incentive and contracting effects’, *Journal of Financial and Quantitative Analysis*, 26, pp. 519–532.

Bessembinder, H. and M. L. Lemmon (2002). ‘Equilibrium pricing and optimal hedging in electricity forward markets’, *Journal of Finance*, 57, pp. 1347–1382.

Bollerslev, T. and M. Melvin (1994). ‘Bid–ask spreads and volatility in the foreign exchange market: an empirical analysis’, *Journal of International Economics*, 36, pp. 355–372.

© 2019 British Academy of Management.
Bollerslev, T., A. J. Patton and R. Quaedvlieg (2016). ‘Exploiting the errors: a simple approach for improved volatility forecasting’, *Journal of Econometrics*, **192**, pp. 1–18.

Bartram, S. M., G. W. Brown and J. Conrad, (2011). ‘The effects of derivatives on firm risk and value’, *Journal of Financial and Quantitative Analysis*, **46**, pp. 967–999.

Carter, D. A., D. A. Rogers and B. J. Simkins (2006). ‘Does hedging affect firm value? Evidence from the US airline industry’, *Financial Management*, **35**, pp. 53–86.

Castelli, M., L. Vanneschi and M. De Felice (2015). ‘Forecasting short-term electricity consumption using a semantics-based genetic programming framework: the South Italy case’, *Energy Economics*, **47**, pp. 37–41.

Cerchiello, P. and P. Giudici, (2016). ‘Big data analysis for financial risk management’, *Journal of Big Data*, **3**, pp. 1–18.

Chen, Z., B. Han and Y. Zeng (2017). ‘Financial hedging and firm performance: evidence from cross-border mergers and acquisitions’, *European Financial Management*, **23**, pp. 415–458.

Choi, T. M., H. K. Chan and X. Yue, (2017). ‘Recent development in big data analytics for business operations and risk management’, *IEEE Transactions on Cybernetics*, **47**, pp. 81–92.

Choi, T. M., S. W. Wallace and Y. Wang, (2018). ‘Big data analytics in operations management’, *Production and Operations Management*, **27**, pp. 1868–1883.

Christensen, B. J and N. R. Prabha (1998). ‘The relation between implied and realized volatility’, *Journal of Financial Economics*, **50**, pp. 125–150.

Christoffersen, P. F. and F. X. Diebold, (2000). ‘How relevant is volatility forecasting for financial risk management?’ *Review of Economics and Statistics*, **82**, pp. 12–22.

Collin-Dufresne, P. and V. Fos (2016). ‘Insider trading, stochastic liquidity, and equilibrium prices’, *Econometrica*, **84**, pp. 1441–1475.

Corwin, S. A. and P. Schultz (2012). ‘A simple way to estimate bid–ask spreads from daily high and low prices’, *Journal of Finance*, **67**, pp. 719–760.

Duru, O. (2010). ‘A fuzzy integrated logical forecasting model for dry bulk shipping index forecasting: an improved fuzzy time series approach’, *Expert Systems with Applications*, **37**, pp. 5372–5380.

Easley, D., N. M. Kiefer, M. O’Hara and J. B. Paperman (1996). ‘Liquidity, information, and infrequently traded stocks’, *Journal of Finance*, **51**, pp. 1405–1436.

Fleming, M. J. and E. M. Remolona (1999). ‘Price formation and liquidity in the US Treasury market: the response to public information’, *Journal of Finance*, **54**, pp. 1901–1915.

Garman, M. B. (1976). ‘Market microstructure’, *Journal of Financial Economics*, **3**, pp. 257–275.

Gao, T., A. Gupta, N. Gulpinar and Y. Zhu (2015). ‘Optimal hedging strategy for risk management on a network’, *Journal of Financial Stability*, **16**, pp. 31–44.

Gordon, L. A., M. P. Loeb and C. Y. Tseng (2009). ‘Enterprise risk management and firm performance: a contingency perspective’, *Journal of Accounting and Public Policy*, **28**, pp. 301–327.

Goulielmos, A. M. and M. E. Psifia (2013). ‘Forecasting short-term freight rate cycles: do we have a more appropriate method than a normal distribution?’, *Maritime Policy & Management*, **38**, pp. 645–672.

Goyenko, R., C. Holden and C. Trzcinka (2009). ‘Do liquidity measures measure liquidity?’, *Journal of Financial Economics*, **92**, pp. 153–81.

Haushalter, D., S. Klasa and W. F. Maxwell (2007). ‘The influence of product market dynamics on a firm’s cash holdings and hedging behavior’, *Journal of Financial Economics*, **84**, pp. 797–825.

Hirsh, H., W. Banzhaf, J. R. Koza, C. Ryan, L. Spector and C. Jacob (2000). ‘Genetic programming’, *IEEE Intelligent Systems*, **15**, pp. 74–84.

Howard, C. T. and L. J. D’Antonio (1994). ‘The cost of hedging and the optimal hedge ratio’, *Journal of Futures Markets*, **14**, pp. 237–258.

Jin, Y. and P. Jorion (2006). ‘Firm value and hedging: evidence from US oil and gas producers’, *Journal of Finance*, **61**, pp. 893–919.

Jung, K., Y. Lim and J. Oh (2011). ‘A model for measuring supplier risk: do operational capability indicators enhance the prediction accuracy of supplier risk?’, *British Journal of Management*, **22**, pp. 609–627.

Kambouroudis, D. S., D. G. McMillan and K. Tsakou (2016). ‘Forecasting stock return volatility: a comparison of GARCH, implied volatility, and realized volatility models’, *Journal of Futures Markets*, **36**, pp. 1127–1163.

Lau, C. K. (2016). ‘How corporate derivatives use impact firm performance?’, *Pacific-Basin Finance Journal*, **40**, pp. 102–114.

Lewellen, K. (2006). ‘Financing decisions when managers are risk averse’, *Journal of Financial Economics*, **82**, pp. 551–589.

Liebenberg, A. P. and R. E. Hoyt, (2003). ‘The determinants of enterprise risk management: evidence from the appointment of chief risk officers’, *Risk Management and Insurance Review*, **6**, pp. 37–52.

Lux, T. and M. Marchesi (1999). ‘Scaling and criticality in a stochastic multi-agent model of a financial market’, *Nature*, **397**, pp. 498–500.

Madhavan, A. (2000). ‘Market microstructure: a survey’, *Journal of Financial Markets*, **3**, pp. 205–258.

Manahov, V., R. Hudson and H. Hoque (2015). ‘Return predictability and the “wisdom of crowds”: Genetic Programming trading algorithms, the Marginal Trader Hypothesis and the Hayek Hypothesis’, *Journal of International Financial Markets, Institutions and Money*, **37**, pp. 85–98.

Mohanty, S., M. Narinda, E. Habis and E. Juhabi (2014). ‘Oil price risk exposure: the case of the US travel and leisure industry’, *Energy Economics*, **41**, pp. 117–124.

Ng, J. (2011). ‘The effect of information quality on liquidity risk’, *Journal of Accounting and Economics*, **52**, pp. 126–143.

Oxelheim, L. and C. G. Wihlborg (1995). ‘Measuring macroeconomic exposure: the case of Volvo cars’, *European Financial Management*, **1**, pp. 241–263.

Parnell, J. A., E. B. Dent, N. O’Regan and T. Hughes (2012). ‘Managing performance in a volatile environment: contrasting perspectives on luck and causality’, *British Journal of Management*, **23**, pp. S104–S118.

Pérez-González, F. and H. Yun, (2013). ‘Risk management and firm value: evidence from weather derivatives’, *Journal of Finance*, **68**, pp. 2143–2176.

Phan, D., H. Nguyen and R. Faff (2014). ‘Uncovering the asymmetric linkage between financial derivatives and firm value – The case of oil and gas exploration and production companies’, *Energy Economics*, **45**, pp. 340–352.
Shusheng Ding has a PhD in finance and his research focuses mainly on the financial modeling and investment with machine learning and deep learning, blockchain financing and cryptocurrency, financial econometrics and derivative pricing. He has published a number of papers in high-quality financial journals such as Quantitative Finance and Journal of Futures Markets.

Yongmin Zhang is Chair Professor in Applied Economics and Professor in Finance, Qianjiang Distinguished Professor at Ningbo University. Prior to this, he was Chair Professor in Finance and Director of Center for Global Finance at Nottingham University from 2011 to 2018. He also taught at University of Chicago and State University of New York for 17 years. He has authored about 80 papers. He obtained his PhD from the University of Chicago in 1997.

Meryem Duygun holds the Aviva Chair in Risk and Insurance at the Nottingham University Business School. She is the President of IFABS-International Finance and Banking Society. She leads the University of Nottingham Fintech Research Network. She has published two books and several articles in major journals, including Journal of Corporate Finance, Journal of Banking and Finance, European Journal of Operational Research among others. She is Subject Editor of Journal of International Financial Markets Institutions & Money.