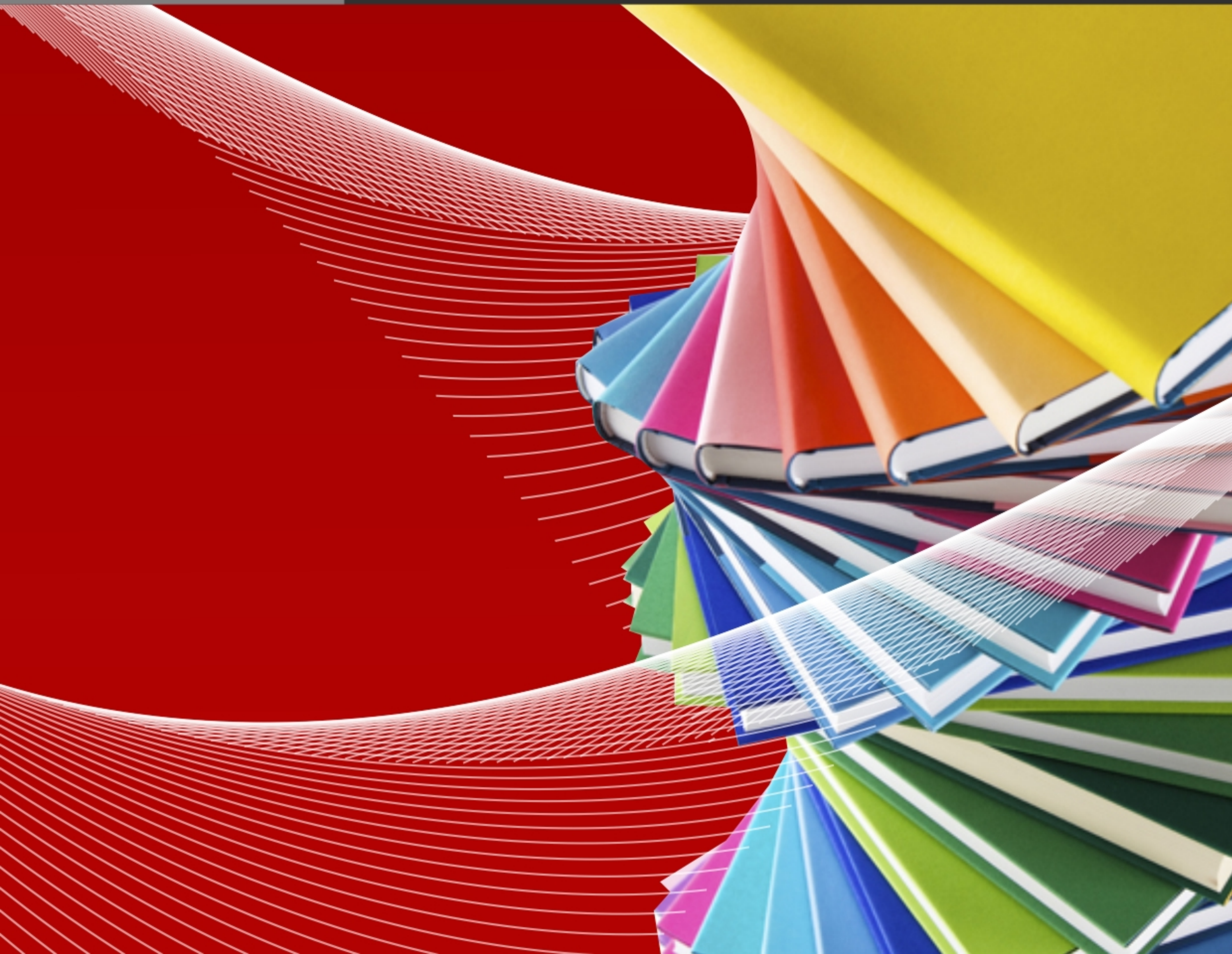
澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

Outstanding Academic Papers by Students
學生優秀作品



UNIVERSITY OF MACAU

FACULTY OF BUSINESS ADMINISTRATION

`

MANAGING CHINESE COMMODITY FUTURES PORTFOLIO: A STOCHASTIC PROGRAMMING APPROACH

Ng Fok Cheong

MB0430191

Thesis presented to the

Faculty of Business Administration

University of Macau

In partial fulfillment for granting the MSc in Finance Degree

2014

Managing Chinese commodity futures portfolio: a stochastic programming approach

Acknowledgement

The four-year study of the master degree in University of Macau is definitely one of the most memorable experiences in my life so far.

Above all, I would like to thanks my supervisor Dr. Steve Lo. Dr. Lo guided me a clear direction in my thesis. Without his advice and support, I cannot have this thesis finished. As a tutor and as a friend, I am extremely grateful for his kind help during the past years.

The classmates that I met in the college are wonderful. I would like to thank them for the kindness, true friendship and support during the study. They gave me the best memories in my master studies.

Four-year was not a short time. I would like to thank all those people that encouraged me to finish this thesis in the past three years. Without their support, I may have given up. The completion of this thesis is the best achievement since I graduated from my bachelor in 2008.

Abstract

We present a dynamic stochastic programming model to manage future contracts in Chinese commodity future markets. We simulate the uncertainty in asset prices by means of the scenario trees that approximate the empirical joint distributions implied by historical market data. We firstly propose an improved algorithm that generates a discrete joint distribution consistent with the first four marginal moments and correlation matrix of random variables. We point out that algebra modeling language and our moment matching algorithm make it possible for any non-specialist users to build models in order to solve complex sequential decision making problems. Secondly, we propose a stochastic programming model for managing the futures contract. Empirical results based on the Chinese commodity futures contracts compare the performance of our proposed model with those of other popular trading strategies presented in finance literature. We also show that the multi-stage model outperforms single-stage models in terms of the stability of return series generated.

Keywords: Stochastic programming; Risk managements; Chinese future contracts; Scenario generation.

I. Introduction

Mathematical Programming, or Mathematical Optimization is a branch of quantitative methods exploring optimal ways to achieve a certain objective when facing with constraints. Normally, it solves problems that maximizing or minimizing a utility function that related to decision variables while subjected to several constraints. In financial industry, a typical problem that studying on how to allocate the portfolios to reduce the risk while maintain a certain level of expected return can be well solved by an optimization model. In the traditional mathematical model, all input parameters are constants. However, we are always facing uncertainties in the real market. When the uncertainties of the parameters are incorporated into a mathematical programming model, the model could be considered as a stochastic programming one. The uncertainties in the model are presented by their joint distributions and stochastic process for the single- and multiple-period models respectively.

Optimization techniques are now one of the most widely used quantitative approaches in decision analysis in many areas. Research topics ranging from agriculture to economics, from engineering to operational scheduling can be solved efficiently by mathematical optimization models. With the rapid development of computer technology, financial practitioners both from academic institutes and industry started to explore the effect of stochastic programming models in financial markets. Ziemba (2003) pointed out that risks are well diversified and the extreme scenarios are taken into consideration in the stochastic programming models. Thus the value of portfolios are well protected from unpredicted outcomes in the extreme cases while the models were also demonstrated to have good performances in normal times. To explore the effect of stochastic programming model in international portfolio managements with multi-currency and bonds, Topaloglou, Vladimiro and Zenios(2006)'s model showed that the approach is a flexible and effective tools in the international markets. Kouwenberg (2001) proposed a asset liability management optimization model for a Dutch pension fund. Their study showed that an appropriate scenario generation method is critical in the stochastic programming model. The performance could be improved significantly if an appropriate method is applied.

Meanwhile, financial practitioners found it difficult to handle and implement stochastic models in financial markets. The challenge of a stochastic programming model is twofold: 1). generate scenarios as the input of the stochastic programming model, as most of the input variables in the financial markets are uncertain. 2). build and solve the model. The mathematical algorithms involved to solve a stochastic model is very complicated, in this paper, we rely on commercial statistical software and solvers to tackle these difficulties. We show that computer software enable non-specialist to build their own models and get the optimal solutions efficiently.

Scenario generation plays an essential role in the stochastic programming model. Kaut and Wallace (2003) provided several scenario generation approach and discussed on the evaluation of the suitable methods for any given cases. A moment-matching method proposed by Høyland and Kaut (2002) is suitable if we do not know the distribution function of the marginal. The solution was also applied by Yin and Han (2013) who developed a multiple stages stochastic programming model to apply options to hedge the risk in international portfolio risk managements. Moment-matching solution is employed in our scenario trees' construction. The procedure allows for distributions of different types and various realistic constraints. However, we found that the moment-matching algorithm may show significant deviations if the scenario sample is not sufficient. We made a slight modification on the algorithm so that the method would produce well-matched statistical properties regardless of the sample size. We developed a fully-automated SAS program to implement this scenario generation method.

We also noticed that although stochastic programming models have been implemented and tested in many financial areas, commodity future markets has not drawn much of the attention from stochastic programming modelers. Considering the transaction cost are significant lower and the justifiability of short position takings in future market, it is realistic that we can implement a stochastic programming model which re-balance the portfolio in every decision-making points. We developed a stochastic programming model in Chinese Commodity futures markets and showed that stochastic programming is a powerful analytical method to make sequential portfolio management decisions under uncertainty in these markets.

Financial activities always involve risk. Futures are extreme risky as they are always traded with margin. Trading rules in future market allows investors to purchase 100% of the contract value by depositing only 5-20% down in the margin account. Thus, each margin account covers only 5-20% of the values for underlying assets. Depending on the margin ratio, futures trading can give investors 5 to 20 times leverage. Leverage is a great thing when price movement of underlying asset is in the investors' favor. However, when the price movement is in an unexpected way, investors would experience leveraged losses. Apparently risk management is one of the most concerned topics in future markets. In our model, conditional value-at-risk (CVaR) was applied as the risk measure, which often used to reduce the probability a portfolio will incur large losses. Several advantages of conditional value-at-risk (CVaR) are mentioned by Rockafellar and Uryasev (2002). As a risk indicator, CVaR shows significant advantage over value-at-risk(VaR) as CVaR is able to capture excess losses beyond VaR. CVaR is also designed to measure the discrete loss distribution and make the large-scale problems practically and effectively calculated. An example was provided to illustrate the numerical efficiency and stability of CVaR in this paper. Mathematically speaking, CVaR is derived by taking a weighted average between the value at risk and losses exceeding the value at risk. With these advantages, CVaR is employed as the risk indicator and minimized in the end of planning horizon in our model.

To be summarized, this paper provides a stochastic programming model implemented in Chinese Commodity future markets and shows that the approach is an effective way to manage commodity futures. CVaR is employed in the objective function to minimize the expected losses in the extreme cases. We also develop automated computer programs to conduct scenario generation process and solve the models.

This paper is presented with following structure. In Section 2 we introduce the problems we are exploring and the modeling approaches. Section 3 explains the future contract return series construction and scenario tree that was used as the input to represent the uncertainty in the single- and multi-stage future contracts management optimization model. We further provide our scenario generation algorithm which is fully automated in SAS. In section 4 we create the single- and multi period stochastic programming models in future markets and examine its key features. The model is then solved by CPLEX in AMPL. The results and findings of the model are discussed in Section 5. Section 6 concludes the paper and points out the questions remaining for further studies.

II. Problem description and modeling approach

We look into a problem of an investor who is concerned with the dynamic management of future contracts in China commodity future markets in 18-month planning horizon. The investor would like to generate a certain amount of profit while control the downside risk exposure. Initially the investor starts with a certain amount of cash and has full understanding of price information in the history. He assesses movements of the market price based on historical information and decides the portfolio composition to maximize his utilities. At next decision period, the investor rebalances the portfolio to response the new information that revealed in the last periodic intervals. This problem has a dynamic structure, as the investor would close out all positions and re-allocate the funds in the future markets. The rebalancing decisions are realized by a sequence of long and short of future contracts in the markets.

The investor's perception of market price movements are described as the joint distribution of the historical prices movements, which are used to project the uncertain price movements in the future. The projected price movements are represented as a scenario three, which is served as the key decision-making basis of the future investor. In each decision period, the investor re-construct a portfolio to maximize his utilities based on the projected outcome in the scenario trees. Here, the utility indicates the CVaR of portfolio, which is generally used to describe the expected large losses beyond VaR at the end of planning horizon. The reason we choose this risk measure is that use CVaR in the objective function enables us to effectively control the downside risk, which is often encountered by future market investors.

The financial decision-making problems we discussed above could be well modeled as a stochastic programming model. Stochastic programming assumes that the uncertain parameters are random variables with known probability distributions. This information is then used to transform the stochastic program into a so-called deterministic equivalent which might be a linear program or a nonlinear program. The randomness of input parameters in the model are represented by a discrete scenario tree which depict the joint distribution and stochastic process of price movements. As we indicate below, the return of commodities are asymmetric and fat-tailed. The marginal distribution (mean, variance, skewness and kurtosis) and correlation matrix are captured and simulated by a moment-matching technique proposed by Høyland and Kaut (2002), which are not restricted to any particular distributional assumptions. This flexibility in scenario generation and representing the uncertainties of input parameters is definitely one of the major advantage of stochastic programming approach.

In additional, stochastic programming enables the model to accommodate different utility objective to capture the investor's risk bearing preference. CVaR which quantifies the large losses in extreme cases, can be expressed by a simple formula and readily be incorporated into our stochastic programming model. The formula makes CVaR practical and effectively dealing with risks.

Commercial computational and statistical software play crucial roles in the solution of our stochastic model. An iterative algorithm is applied to conduct moment-matching techniques in the scenario tree generation process. The whole process is realized by a fully-automated SAS project, which takes approximately 45 minutes to generate a scenario tree with two-stages and 150x100 scenarios in a personal computer. The algorithm involved to solve a stochastic programming model is extremely complicated. However, our model can be solved effectively by commercial solvers in AMPL.

Both single- and multi-stage models are adopted in the Chinese future contract management problem. Multi-stage model captures long tern price movement and help investors to make decisions under long term market conditions and avoid myopic reaction to short-term market fluctuations that might show risky. These advantages of multi-stage model will be shown in the comparison of single and multi-stage models. All of models and trading strategies are implemented in the real future contract data in the Chinese future markets, we further compare the performance of our model in 18-month horizon with the returns generated by the trend following strategies which provided by Szakmary, Shen and Sharma (2010) and show that our model generates a much more stable and reliable positive return series.

III. Return Series and Scenario Tree Construction

1) Return Series

Comparing with the future markets in eastern countries, Chinese commodity future markets can only be viewed as a rising industry. We extract daily settlement price for 15 futures which are currently actively traded in Chinese commodity future markets and covering trading period longer than 6 years. The statistical properties that evaluated from historical data are expected to be an accurate and stable estimation of the true values if the sample size is sufficient. The futures in our study are traded on three different exchanges, where the underlying assets are all single deliverable commodities. To obtain infinite daily price series, or equivalently, daily return series, we follow the procedures proposed by Szakmary, Shen and Sharma (2010). For each commodity future, before the month of contract expiration, current contract is rolled over to the next contract on the last trading day. Meanwhile, during the contract roll over, the return are always calculated from the data in the same contract. Thus price both from the current and next contract are extracted on the roll-over days. And to build a daily price value index, initial price value of each commodity is set to be 1, price values in the subsequent periods are derived by multiplying the price value in the previous period with the return.

We convert the daily series to a monthly frequency by extracting the price index on the lasting trading day of each month. Table 1 lists the 15 individual commodity futures and their general information including exchanges, start dates, transaction cost, margin percentage and contract size. Meanwhile, the summary statistics and the result of normality tests of the monthly return series are provided in Table 1 as well.

The 15 individual commodity futures included in our study covers products from different industries, including agriculture, industrial, metal and energy. The summary statistics listed in Table 1. show there are significant differences in the central location and volatility in the returns across the 15 commodities. The average monthly returns are ranging from -1.20% (Palm Oil) to 1.44% (No.2 Soy Bean), while standard

Table 1.
Market information and summary statistics

		Start	Latest			Transaction	Monthly return on long positions					Avg. volume (no. of contracts)	
Commodity	Exchange	Date	Price	Margin	Size	Cost	Mean(%)	Std.Dev(%)	Skew(%)	Kurt(%)	Kolmogorov-Smirnov	Open Interest	Volume Traded
CORN	DCE	22Sep2004	2,363	10%	10	1.2	-0.60	2.79	-6	-58	0.052	54	12
LLDPE	DCE	31Jul2007	11,005	10%	5	2.5	0.09	8.31	-190	1,009	0.106*	34	9
NO.1 SOYBEAN	DCE	15Mar2002	4,548	10%	10	2.0	0.02	4.59	-139	779	0.080*	61	10
NO.2 SOYBEAN	DCE	22Dec2004	4,122	10%	10	2.0	1.44	5.78	-102	564	0.108**	2	1
PALM OIL	DCE	29Oct2007	5,056	10%	10	2.5	-1.20	6.89	-3	83	0.060	1,004	14
SOYBEAN MEAL	DCE	17Jul2000	3,933	10%	10	2.0	2.09	7.75	68	123	0.051	1,075	447
SOYBEAN OIL	DCE	09Jan2006	6,882	10%	10	2.5	0.31	7.93	-9	151	0.106**	53	9
ALUMINIUM	SHFE	28Feb1994	14,430	15%	5	3.0	-0.51	4.45	30	243	0.096**	31,036	2,744
COPPER	SHFE	26May1993	52,730	15%	5	10.0	0.34	8.25	-151	732	0.073**	60,836	17,127
FUEL OIL	SHFE	24Aug2004	4,017	30%	50	1.0	-0.58	6.56	-101	326	0.085*	17	4
NATURAL RUBBEI	SHFE	27Jun1995	18,055	20%	10	4.0	0.32	8.32	-28	133	0.110**	464	72
ZINC	SHFE	26Mar2007	14,810	15%	5	3.0	-0.80	8.20	-124	417	0.092	13,396	1,918
COTTON #1	ZCE	01Jun2004	20,330	25%	5	4.3	0.29	6.92	257	1,252	0.155**	3,421	226
PTA	ZCE	18Dec2006	7,670	15%	5	4.0	0.09	7.77	-51	315	0.076	200	63
SUGAR	ZCE	06Jan2006	5,341	15%	10	4.0	0.61	6.51	134	473	0.109**	3,204	287

Note:

1. Exchange: DCE is Dalian Commodity Exchange, ZCE is Zhengzhou Commodity Exchange and SHFE indicates Shanghai Futures Exchange.;
2. Latest prices indicate the price on Sep 30, 2013 for the contracts that will expire in Oct/Nov 2013.
3. Margin is the lowest amount, in terms of the percentage of contract value, on the last month before the contract expiration as required by the exchanges.
4. Size is the deliverable quantity of underlying commodity in a future contract.
5. Transaction cost is fees per contract that paid by the investors.
6. Kolmogorov-Smirnov is a normality test in non-parametric way.
   * indicates statistically significant at the 5% level.
   ** indicates statistically significant at the 1% level.
7. Average open interest/volume traded refer to daily average no. of contracts from Aug 30, 2013 to Sep 30, 2013 in the latest expired contracts.

Table 2. Pearson Correlation Matrix of monthly return of commodity futures over the period Apr 2007 - Sep 2013.

	CORN	LLDPE	NO.1 SOYBEAN	NO.2 SOYBEAN	PALM OIL	SOYBEAN MEAL	SOYBEAN OIL	ALUMINIUM	COPPER	FUEL OIL	NATURAL RUBBER	ZINC	COTTON #1	PTA	SUGAR
CORN	1														
LLDPE	0.00	1													
NO.1 SOYBEAN	0.28	0.23	1												
NO.2 SOYBEAN	0.34	0.38	0.66	1											
PALM OIL	0.18	0.43	0.44	0.47	1										
SOYBEAN MEAL	0.28	0.47	0.54	0.45	0.38	1									
SOYBEAN OIL	0.06	0.45	0.53	0.53	0.65	0.37	1								
ALUMINIUM	0.26	0.43	0.24	0.32	0.27	0.27	0.23	1							
COPPER	0.22	0.67	0.23	0.44	0.44	0.36	0.42	0.72	1						
FUEL OIL	0.25	0.49	0.35	0.50	0.44	0.30	0.29	0.29	0.46	1					
NATURAL RUBBER	0.16	0.59	0.25	0.46	0.48	0.37	0.38	0.48	0.68	0.41	1				
ZINC	0.13	0.71	0.27	0.41	0.46	0.34	0.47	0.60	0.80	0.34	0.56	1			
COTTON #1	0.12	0.25	0.12	0.13	0.25	0.12	0.24	0.22	0.26	0.11	0.40	0.35	1		
PTA	0.11	0.67	0.14	0.25	0.53	0.27	0.41	0.48	0.61	0.44	0.63	0.55	0.35	1	
SUGAR	0.19	0.11	0.22	0.31	0.27	0.05	0.30	0.38	0.37	0.15	0.41	0.40	0.46	0.23	1

deviations are ranging from 2.8% to 8.3%. Although the commodities display different means and variances, the skewness and kurtosis suggest the commodities are asymmetry and fat tailed. The Kolmogorov-Smirnov test for null hypothesis that the return is normal distribution, suggests we can believe that among the 15 commodities, at least 10 return series are not normal distributed. The correlation matrix of the random variables over the period Apr 2007 - Sep 2013 are shown in Table 2. The observed correlations across different commodities are relatively low. Hence, total risk in the diversified commodity future portfolio can be well reduced. All these results suggest that moment-matching is a suitable solution for the scenario generation. Finally, we point out that the average trading volumes among the 15 selected show large variations, ranging from over 17,000 contracts traded per day in Copper down to 1 contract only in No. 2 Soybean. The low open interest and trading volume in some of these commodities suggest that our trading strategies may difficult to implement in real markets. However, since the markets are growing, it is still worth to test the efficacy of these strategies.

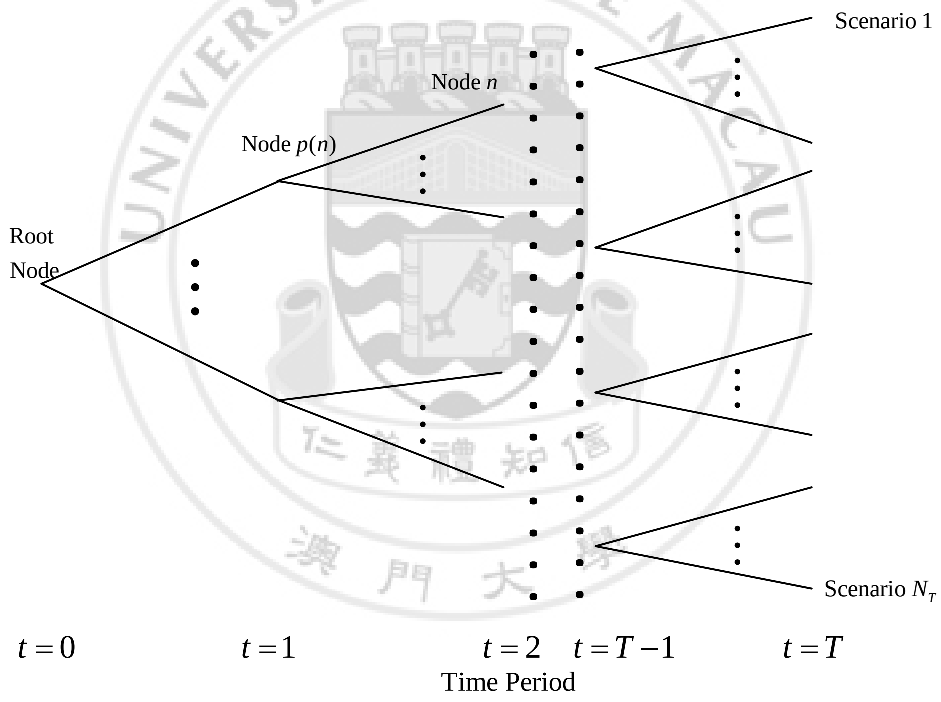
Once we obtain the monthly basis return series, we are able to estimate the statistical properties which are served as inputs of scenario generation. In each period for portfolio re-balancing, based on the historical data in the past 5 years, we re-calculate the mean, variance, skewness and kurtosis for each commodity monthly return series and correlation matrix for these series. Then the 4 moments and correlation matrix are served as the key parameters to generate the price scenarios, which are used as the inputs to generate trading signals in the stochastic programming model.

2) Scenario tree

Quality of the solution obtained by solving a stochastic programming model highly depends on how well the uncertainty influencing the portfolio construction is represented. One of the major advantages of a stochastic programming model is that the input random variables are not restricted to a particular distribution assumption. The key input random variables that influenced decision problems are the price movements, or equivalently, the returns for each commodity. Discrete evolution of the price movements during the planning horizon could be modeled as a scenario tree which shown in Figure 1. The possible outcomes in a scenario tree are projected based on the statistical model that employed with historical market data as the input parameters, as well as the opinions of experts.

Generally, a planning horizon is divided into T periods, each period is corresponding to a time which portfolio rebalance decisions are made. We use monthly trading period in our model. That means, the future portfolios are rebalanced every month. Starting from a root node in Period 0 which the investor has full knowledge of all realized market information, as well as a certain amount of funds in hand. That is, the input parameters in Period 0 are known with certainty. The investor's perception of uncertainties in the subsequent period are depicted as a tree which branches out from root node. The subsequent nodes in the tree represent possible outcomes in periods $t = 1,...,T$. Each node is associated with a unique predecessor. Random variables in the nodes associated with a same predecessor follow a same joint conditional distribution. Parameters in the nodes emanating from same node are projected by same joint distribution with a statistical model. Each node in the end of planning horizon (terminal node) represents a scenario. Each scenario associated with a unique terminal node distinguished a unique evolution process of random variables. The terminal nodes are not necessary to be generated with same probability, and the tree does not need to be symmetric and binomial as well. The size of a multi-stage scenario tree grow substantially as trading periods and nodes increase.

Figure 1. A typical scenario tree



We define below notations for the scenario tree:

N is the set of nodes in the tree;

$n \in N$ is a typical node in in the tree, $n = 0$ reprenets the root node.

T is the set of trading periods in the tree; $t_0 \in T$ is the initial time.

$N_t \in N$ is the set of nodes in period $t = 0,1,...T$ in the tree, N_T is the set of terminal nodes;

$p(n) \in N$ is the unique predecessor node of node $n \in N \setminus \{0\}$;

p_n is the probability of state associated with node *n*.

As we mentioned above, the stochastic process of random variables are projected in scenario trees. In each node $n \in N \backslash n_0 \backslash N_T$ (intermediate nodes), portfolios are rebalanced based on the realized market information and the postulated outcomes that emanating from this node. The realization of random variables in the intermediate nodes are modeled by statistical solutions and experts' opinions. In the terminal nodes, we only calculate the value of portfolio based on the realized marketing information in the end of planning horizon.

The statistical model that employed to project random variables in the intermediate nodes is critical in the construction of a scenario tree. We consider the movement of market prices in commodity future contracts, or equivalently the returns of these commodities in the scenario tree. The correlation matrix presented in Table 2 shows that the returns of commodities are correlated. Moreover, the Kolmogorov-Smirnov test results shown in Table 1. suggest that the random variables are generally not normal distributed. Rather than normal distributed, skewness and excess kurtosis are observed in the sample data. These statistical features of data should be fully captured in the representation of uncertainties in a scenario tree. Given the statistical characteristics we discussed above, the moment-matching technique which is suitable in this situation, is applied in our scenario generation process.

Specifically, the first four moments and correlation matrix of the random variables are matched with the historical market data in our model.

The first four moments are:

- Mean: depicts the central location of a random variable.
- Variance: measures the dispersion of a random variable.
- skewness and kurtosis: measure the shape of a random variable's probability distribution. Specially, skewness describes the asymmetry about the mean of a random variable, while kurtosis indicates the "peakedness".

We should point out that the four moments of random variables and their correlation matrix could not uniquely determine the joint distribution. That means, two different joint probability distribution functions could have same mean, variance, skewness, kurtosis and correlation matrix. However, statisticians suggest that first four moments and correlation matrix are sufficient to represent the uncertainties of random variables.

All random variables in nodes in each time period $t = 1...T$ should follow a same joint distribution, which is modeled and generated by the moment-matching procedure. Size of the tree depends on the number of time period as well as number of realized outcomes(nodes) in each time period. Higher accuracy of the randomness is captured with more outcomes postulated in the scenario tree. However, the size of a scenario tree is increased exponentially with the increase of nodes in the intermediate periods. Our model presents in the later section is flexible to involve arbitrary number of decision periods. However, it is extremely time exhausting and high computational effort requirements to solve a large scale stochastic model. Generally, a large model is solved by professional computational machines. Due to system limitation, our model is applied in a single-stage and a two-stage case respectively. The models are then solved in a regular personal computer.

The moment-matching procedures are initially provided by Kaut and Wallace (2003). To match statistical properties which calculated from the historical data, the procedure generally gives stable results in the case of large number of scenarios. However, if the postulated outcomes are small, say, 150 in the second stage, significant deviations are shown. To ensure a stable and well-matched outcome, we slightly modify the iterative algorithm that provided by Kaut and Wallace (2003). The improved solution provided a way to generated postulated outcomes with statistical properties well-matched with the historical data within a certain accuracy. Please refer to the Appendix for our modified algorithm. We should point out that the convergence issue of this iterative algorithm is complicated. Mathematically, we are not able to provide an analytical proof for the convergence. However, in the extensive experiments in the future return scenario generation, we never met an unsolvable case. The iterative process is fully-automated by a SAS project designed by us. The process will be stopped once the convergence criteria are achieved or a pre-specified number of iterations are performed. The project restarts the process in the cases of divergences. Empirical test suggests that the divergences will ultimately achieved with multiple experiments. We ensure that all scenario trees has well-matched statistical properties in our model. The major concern for the convergence is the performance of computer in the scenario generation but not the risk of using a tree with miss-matched properties.

As the inputs for our stochastic programming model, we generate both single and two- stage scenario trees. At time period 0, we calculate the first four moments as well as the correlation matrix of commodity returns, which are served as the key parameters of our moment-matching procedure to generate outcomes in period 1 of the scenario tree. 10,000 scenarios are generated for the single-stage tree, while 150 scenarios are projected for the two-stage tree. We generate scenarios incrementally in the two-stage case. For each node in period 1 of the two-stage tree, we re-calculate the statistical properties based on the projected parameters in the predecessor node as well as the realized market information in root node. A number of outcomes for the successors that matched the re-calculated statistical properties are further generated. Same procedures are applied in all of nodes in period 2 of the two-stage tree. We generate a two-stage tree with 150×100 scenarios. Without loss of generality, each scenario is generated with equal probability.

IV. Future contract management model

Our model determines a sequence of investment decisions in Chinese commodity future markets. The decisions are made at discrete time point (monthly) in a given period. The problem is viewed as an investor who is interested in Chinese commodity future markets starting with an initial fund. In the initial period, the investor enters the future market by taking positions in commodity future contracts based on his perceptions of market movements. At the beginning of next period, the investor closes out all positions in the market and re-evaluates the market information. The investor then re-constructs a new portfolio with the newly realized information during previous periods. For every projected outcome in the intermediate nodes, the investor makes different decisions based on the realized return and considers the projected realization of future prices. The decision in each node affects the outcomes in its successive nodes. A sequence of decisions made in the intermediate period end up at a terminal node. The portfolio at a terminal node represents a possible realized outcome which is resulted by a series of decisions made by the investor during the planning horizon. A sequence of decisions during the period distinguishes a unique scenario, which result in a unique node in the terminal period. All of the postulated realization in terminal nodes formulate the sample of outcomes that resulted by the investor's decisions in the starting period. Since the investor is exposed to huge risk in the future market, the model minimizes the CVaR of portfolio losses while achieve an expected target return at the end of planning horizon.

In our stochastic programming model, the uncertainties are simulated by a scenario tree. The random variables in our model are the market price movements of commodities. We project the market prices based on the historical data employing the improved moment-matching algorithm. The market information that listed in Table 1 also served as deterministic parameters in our model.

We define the following notations in our stochastic programming model:

Sets:

F Set of commodity futures

User-specified parameters:

α Critical percentile for CVaR

μ Expected target portfolio return over the planning horizon.

$InitCash$ Initial available case in root node.

Deterministic input data:

$margin_f$ The lowest percentage of contract value that required by the exchange in the last month before contract expiration for commodity $f \in F$.

$Size_f$ Quantity of underlying deliverable commodity per contract $f \in F$.

$TranCost_f$ A fixed amount that paid by future investor for every commodity contract $f \in F$

$price_f^0$ Price of commodity $f \in F$ in initial period (root node).

Scenario Dependent Data:

$Price_f^n$ Price of commodity $f \in F$ at node n .

p_n Probability of node n, for simplicity, we generate symmetry scenarios with equiprobability, thus for $n \in N_T$, $p_n = 1/N_T$.

Decision Variables:

$long_f^n$ No. of commodity contracts $f \in F$ that long in node $n \in N \setminus N_T$

$short_f^n$ No. of commodity contracts $f \in F$ that short in node $n \in N \setminus N_T$

Computed Variables:

$Cash_0$ Cash amount that is not invested in the market at the end of node $n \in N / N_T$

$Position_f^n$ The position of commodity $f \in F$ at the end of node $n \in N / N_T$, in terms of number of contracts. A positive value indicates long position in the future, while negative means short position.

$Value_n$ Total portfolio value at the beginning of node $n \in N$

Δ_n Portfolio gain/loss during the period from node $p(n) \in N / N_T$ to $n \in N / 0$;

R_n Overall portfolio return over the full trading period at node $n \in N_T$.

L_n Overall portfolio loss over the trading period at node at node $n \in N_T$.

ζ Conditional Value at Risk.

Auxiliary variables:

y_n Excess loss of VaR at terminal node $n \in N_T$

z equals to VaR at the optimal solution.

Our stochastic programming model is developed as below:

$$\min\ \zeta = z + \frac{1}{1-\alpha} \sum_{n \in N_T} p_n y_n \tag{1}$$

s.t.

$$Position_f^n = long_f^n + short_f^n,\ \forall f \in F, n \in N / N_T \tag{2}$$

$$Value_0 = InitCash \tag{3}$$

$$\sum_{f \in F} \Big(\left(long_f^0 + short_f^0 \right) \times Size_f \times price_f^0 \times margin_f + \left(long_f^0 + short_f^0 \right) \times TranCost_f \Big) + Cash_0 = InitCash \tag{4}$$

$$\Delta_n = \sum_{f \in F} (Postion_f^{p(n)} \times Size_f \times \left(Price_f^n - Price_f^{p(n)} \right) - 2 \times \left(long_f^{p(n)} + short_f^{p(n)} \right) \times TranCost_f), \forall n \in N / N_0 \tag{5}$$

$$Value_n = Value_{p(n)} + \Delta_n, \forall n \in N / N_0 \tag{6}$$

$$\sum_{f \in F} \Big(\left(long_f^n + short_f^n \right) \times Size_f \times price_f^n \times margin_f + \left(long_f^n + short_f^n \right) \times TranCost_f \Big) + Cash_n = Value_n, \forall n \in N / 0 / N_T \tag{7}$$

And finally, $\forall n \in N_T$, (8)

$$R_n = \frac{Value_n}{InitCash} - 1 \tag{9}$$

$$L_n = -R_n \tag{10}$$

$$y_n \geq L_n - z \tag{11}$$

$$y_n \geq 0 \tag{12}$$

$$\sum_{n \in N_T} p_n R_n \geq \mu \tag{13}$$

The model is a linear stochastic programming with recourses. It is developed to cater arbitrary stages case. Our model determines a sequence of portfolio rebalance decisions during the planning horizon. Decision variables in our model are number of commodity futures that the investor long and short in each node in the initial and intermediate periods. The postulated outcomes of random variables in subsequent nodes determine the composition of portfolio construction and directly impact the decision in the later periods. Random variables are the prices of commodities in the future, which are projected by our moment-matching technique. The projection of scenario tree is considered as one of the most critical part in our modeling solutions. Each outcome at the node in the terminal period distinguishes a unique scenario of the tree. Our model minimizes portfolio risk which measured by CVaR at the end of planning period, while a minimum return μ is set in the constraints as the investor's expected target at the end of period. The expected return is calculated over the equiprobably outcomes at terminal period. In the objective of our model, the conditional value at risk is measured by ζ, while z captures the VaR at optimal level.

Equation 2. and 3. describe the situation of the commodity future investor at initial period. The investor is interested in participating in commodity future market with an amount of cash. The initial value of the portfolio that the investor has equals to the cash amount in hand. Equation 4. explains the formulation of the portfolio in the stages before terminal periods. Taking into account of initial margin requirements as well as the transaction cost, the investor enters the market by taking positions in selected commodity contracts based on his perceptions of future market movements. The margin ratio gives investor leverage on the future trading. The future margin ratio in the last month of contract expiration day, which set by different exchanges are applied in the equations. The investor may choose to keep some cash in the pocket as he does not need to invest all of his money in the market to achieve his objective. The portfolio gain and loss during the trading period are defined in Equation 5.. As we mentioned previously, we use monthly trading period in our model. In each period, all future positions in the previous period are closed out in the starting and re-taken again based on the current market information and projected outcomes in the subsequent periods. Transaction cost is occurred twice during the process. The extremely low cost in future contract transaction makes our process realistic. We do not need to pay a large cost during the frequent taking and closing positions in futures contracts. Equation 6. captures the evolution of portfolio values during the planning horizon.

Portfolio return and portfolio loss in the terminal nodes are defined in Equation 9. and 10. respectively. Constraints in Equation 11. and 12. implied that $y_n = \max\left[L_n - z, 0\right]$, which is the excess short fall over the planning horizon. Equation 13. defined the investor's expectation on the portfolio overall return over the planning horizon.

A sequence of contract rebalance decisions are determined by our future contract management optimization model. Starting with an initial fund in the root node, the model provides optimal strategies in each node by specifying not only the allocation of funds across the future contacts but also the positions in each contract. In the same time, the model makes investment decisions to minimize the excess shortfall in the end of planning horizon. Our model is straightforward and easy to understand, but is served as a pioneer in exploring the effect of managing commodity future contracts by a stochastic programming approach

V. Empirical Results

The stochastic programming models are solved by CPLEX in AMPL. We implemented single stage and two-stage models in Chinese commodity future markets. Through a series of numerical experiments in the real market, we explore the efficacy of stochastic programming approach in future contract managements. We repeatedly run the model in both single- and two-stage cases with alternative minimum target return constraints to compare the performances of stochastic programming models. The trend-following strategies that proposed by Szakmary, Shen and Sharma (2010) are also tested in the same market condition. We further compare the effect of stochastic models with those results generated by trend following strategies.

To identify the most effective approach, we examine the performance of stochastic programming models and trend following strategies in dynamic backtesting simulations. All solutions are run in real market data on a rolling horizon basics over a 18-month planning period starting from Mar 2012 to Sep 2013.

1) Trend following strategies

Before presenting the results of stochastic programming models, we implement two trend follow trading strategies which were proposed by Szakmary, Shen and Sharma (2010). They examined the efficacy of those strategies in commodity future markets as well. We briefly review the two trading strategies as below:

a) **Momentum Strategy**: Based on the returns to a long position in a formulation period, all selected commodities are ranked independently to determined the top commodities as well as the bottom ones. We took long positions in those top 1/3 commodities, and short positions in those bottom 1/3s. No position is taken in those middle commodities. The contract value of the top and bottom commodities are equally weighted. We use 3 and 6 months as the formulation periods to determined the past returns.

b) **Dual Moving Average Crossover (DMAC) Strategy**: The short-term moving average (STMA) and long-term moving average (LTMA) are employed as the indicators of the price trending. STMA exceeding LTMA indicates there is an upward trend in the price movement. Thus, we take long position in the commodity if STMA exceeds the LTMA by B percents as there is an upward trend identified, and short position if STMA fall below B percent of LTMA, no position is taken for the commodities that fall within the band. STMA of 1 months and LTMA of 3 and 6 months is applied in our simulation. For B we use B=0.025 and 0.05.

In the end of Mar 2012, starting with an initial cash endowment, each strategy is applied to determine the initial cash allocation among the selected commodities and the portfolio is re-balanced in the subsequent periods based on the realized real market data. The 18-month overall performances of momentum and DMAC strategies are shown in Figures 2 and 3.

Even trend following strategies generally recorded impressive results in the end of planning horizon, large fluctuations were seen during the trading periods. Portfolio values in all cases experienced huge monthly losses during the 18-month planning horizon, and especially struggled in year 2012. The hugest recessions occurred in Sep and Oct 2012 when the commodity prices were moving in complete different ways with that past trends had been showing. The performance of trend following strategies were much improved when the clock advanced to Year 2013. Each portfolio value under the six strategies soared in Apr 2013 and experienced steady growth in the subsequent months. However, the end portfolio value under DMAC with parameters ST=1, LT=3 and B=0.05 was just 41% of the initial value remaining. Actually, the value was almost quadrupled during the first nine months of Year 2013. However, the failure in Year 2012 was disastrous as nearly 90% of the wealth had already disappeared in the beginning of Year 2013. The portfolio values were not recovered even a prosperity was shown in the coming new year.

As shown in above case, big losses in the future trading is disastrous. How to keep the portfolio value change under stability and avoid big losses is remaining an important topic in future contract managements. In the following paragraphs, we present the solution of stochastic programming approach in this topic.

2) Stochastic programming models

We built both single- and multi- stage models in commodity future managements. The

Figure 2. Realized performance of momentum strategies.

First graph shows the cummulative portfolio values during the 18-month planning horizon, while the second one displays the monthly portfolio return in each period.

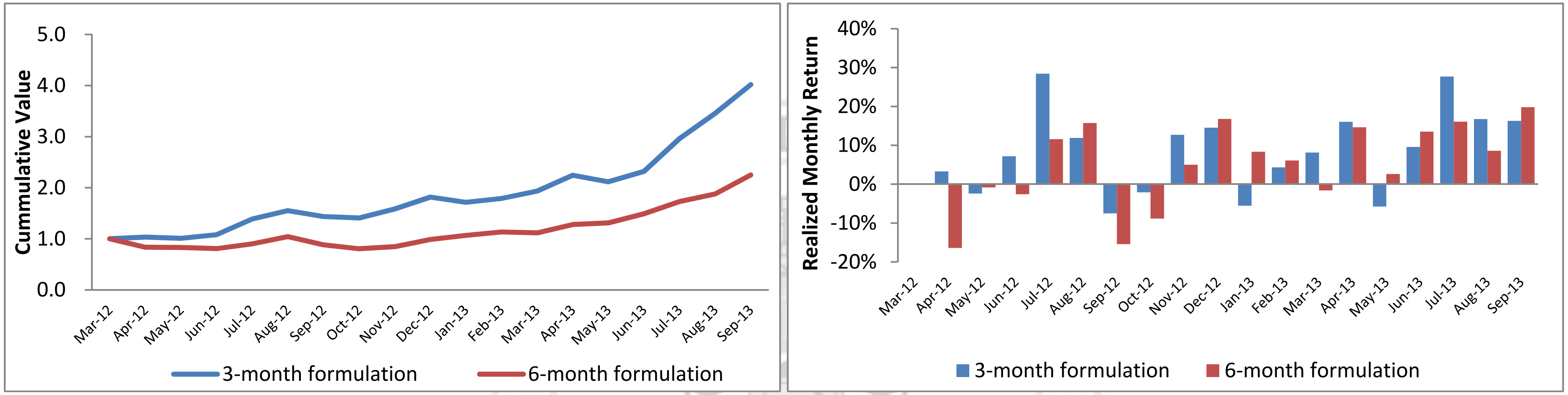
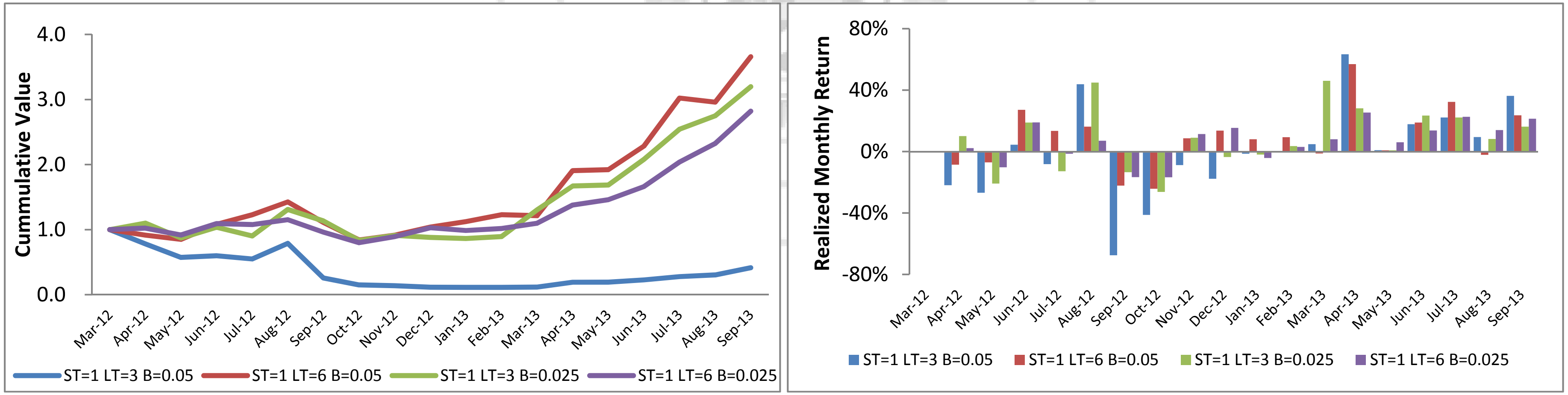


Figure 3. Realized performance of dual moving average crossover strategies:



solutions are consistent with those in the trend following strategies. In the starting of planning horizon - the end of Mar 2012, with the initial cash endowment, each model was implemented to determine the initial portfolio composition. The scenario tree was projected based on the statistics of the historical market information in the past 5 years. Each model was solved by AMPL with the scenario tree as the input and the decisions regarding the optimal portfolio construction were recorded. Then we come to the next month period, the actual realized return of the portfolio in hand was calculated based on the revealed real market information, taking margin rate and transaction cost into accounts. A new scenario tree was built by matching the statistical properties of the price movement in the past five years. The cash in hands which resulting from decisions in previous period was re-allocated to each future account to maximize the investor's utilities by solving the model. The process was repeated for each coming month during the 18-month period until Sep 2013. The ex post realized returns and the progression of portfolio values during the planning horizon were recorded. The backtesting simulation in real market data provided key performance indicator to assess the effect of stochastic programming model in Chinese commodity future markets.

We experimented below models with critical percentile $\alpha = 95\%$ for CVaR to assess the performance of stochastic programming approach:

- a two-stage model that used a scenario tree with 15,000 scenarios (150×100) generated by the moment-matching technique described above.
- a single period model that run in a tree with 10,000 scenarios.

The single period model considers the possible outcomes in the ahead month only, while two-stage model captures more information as the subsequent two-month period data are simulated. We firstly tested the model with expected monthly return target of 5%. We point out that, since margin rate is considered in the model, a target minimum return of 5% or more is achievable and realistic while the porfolio is exposed to a lervaraged higher risk. The realized returns of the two models applying real market information is presented in Figure 4. In both cases, targeted monthly

Figure 4. Realized performance of single- and two-stage mdoels in $\mu = 0.05$.

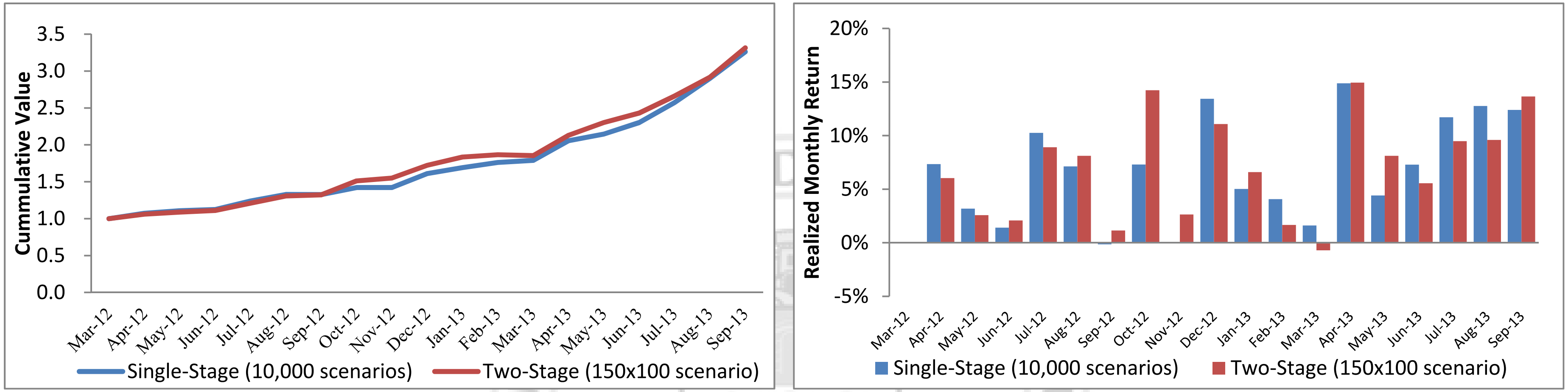


Figure 5. Realized performance of single- and two-stage models in $\mu = 0.06$.

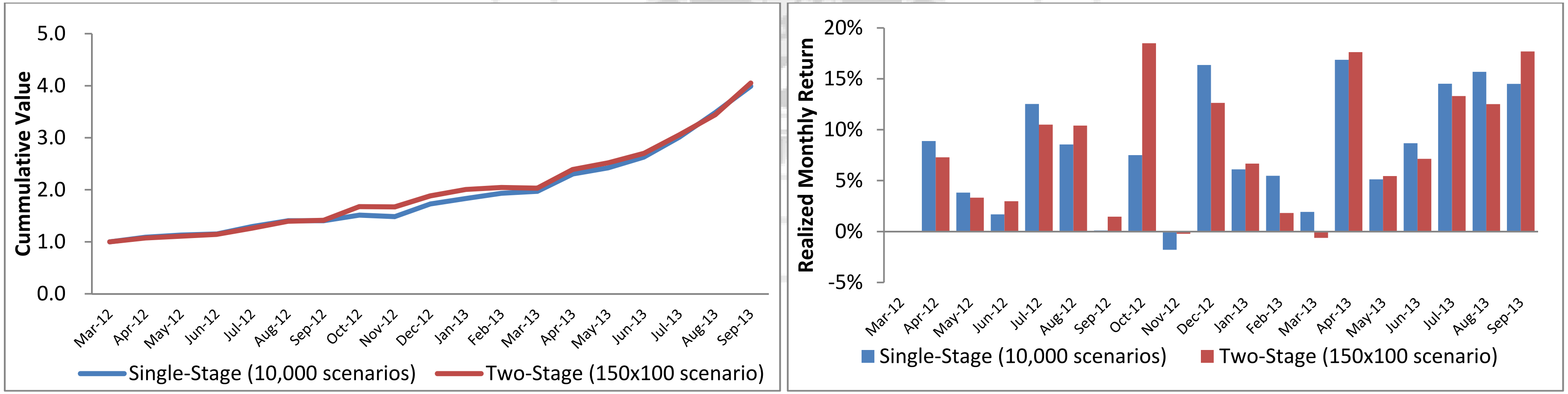


Figure 6. Realized performance of single- and two-stage models in $\mu = 0.07$.

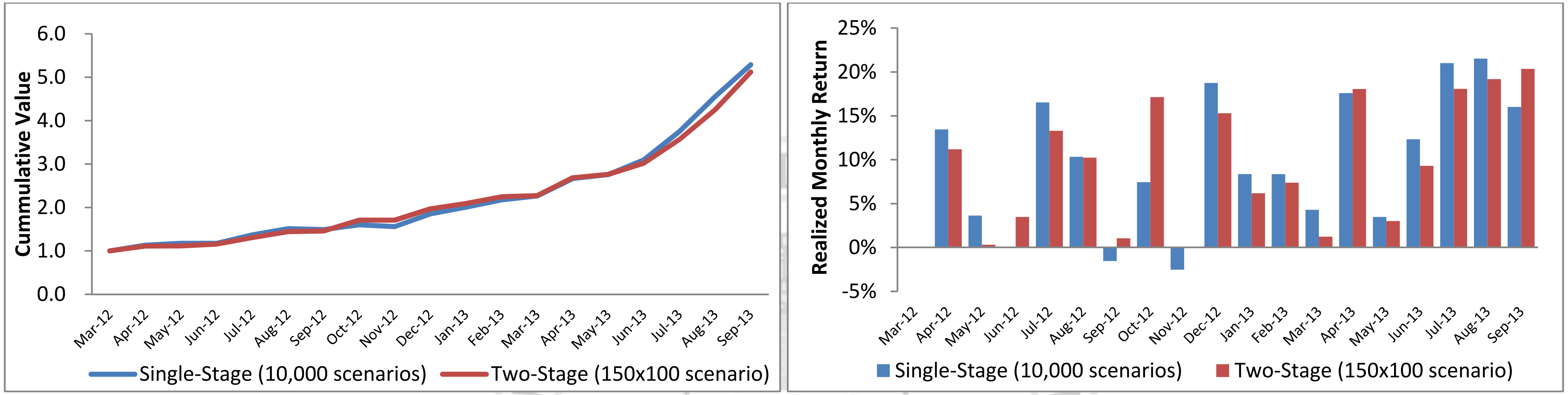
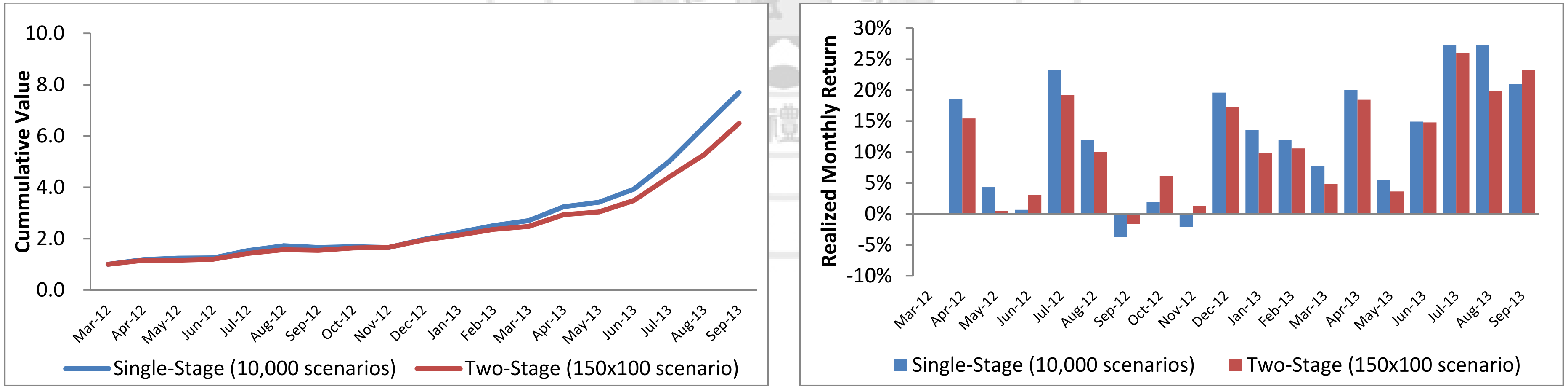


Figure 7. Realized performance of single- and two-stage models in $\mu = 0.08$.



return was sucessfully achieved, an averaged monthly return of 6.8% is realized in the single model, while the figure for two-stage model is 6.9%. The portfolio values were increased steadily during the planning horizon by applying both models, the wealthes showed slightly declines in depressed periods.

However, the downside risk was well diversified by the solution of our models. In two-stage case, the portfolio value dropped down by only -0.7% in Mar 2013, which was the only month that saw declines in two-stage model. Meanwhile, in single-stage model, the portfolio value experienced depresses in Sep 2012 and Nov 2012. Both models generated approximately same averaged montly return and standard devaitions of return (4.8%) of the portfolio value, The two-stage model outperformed his single-stage counterpart by only showing one negative return month.

The performance of two models were not significant different under minimum expected target return $\mu = 0.05$, we furher investigated those results generated by single- and two-stage models with higher expected return μ. Monthly expected target returns of $\mu = 6\%, 7\%$ and 8% were applied in the simulation resepectively. The same scenario trees that used in previous models were also employed in the solution of higher-μ models. Graphs in Figures 5-7 show the expermental results of the three cases. As the target return increased, the weight of commodities with higher expected returns in the portfolio increased substantially and the portfolio was exposed in a higher risky environment. Surprisingly, despite the risk was increased, both single- and two-stage models recorded steady portfolio value growth during the planning horizon. The averaged monthly expected returns was well achieved again, while portfolio recessions happened again in Sep 12 and Nov 12 and the decreasing percentages were amplified with the increase of expected target minimum return. For the $\mu = 7\%$, the porfolio value in single-stage model decreased by 1.5% and 2.5% in Sep and Nov respectively, while slightly wealth rise of 1.0% and 0.1% is recorded in two-stage model. μ was further increased to 8% in the backtesting simulation, and same situation was identified. In Sep 2012, the portoflio value dropped by 3.7% and

1.6% respectively in the single- and two-stage models. When the trading activities moved to Nov 2012, single-stage model generated a 2.7% decrease, but the portfolio value in his two-stage counterpart enjoyed a 1.3% growth. These results showed that two-stage model is apparently considered as a better solusion. It consistently achived target returns while maintained greater stability of returns. The simulation results implied that the increased information contents in the multi-stage model brings incremental benefits into dynamic portfolio managments.

3) Comparison of stochastic programming approcah and traditional trend following strategies.

Fundamentally, our stochastic programming approach could be viewed as a kind of trend following strategies. To decide an optimal porfolio, it also considers past returns of the assets in the portfolio. Shortfalls are unavoidable in the bad times when the price movements are in contrast with the patterns that indicated by the historical data. The targets of our models are set to avoid excess shortfall while achieve minimum target returns during the planning hoizon. As we presented in the previous 2 sections, the stochastic programming approach successfully achived the target returns and generated much greater stability of return series. The downside risk in bad times are well diversified as well, with the best performance contributed by the multi-stage model, which implicitly includes additional future informaton as longer time period data are projected.

Overal performances of stochastic approaches and trend following strategies that tested are summairzed in Table 3. These results show that the traditional trend following trading strategies in commodity future market are dominated by stochastic programming approaches. Although remarkable high geometric average returns were generated by most of the trend following strategies, high standard devaitions ranging from 10% to 30% were shown. Monthly losses were occurred frequently in the return series. During the 18-month planning hoirzon, the trend following strategies reported at least 5 monthly losses. The two-stage model definitely exhibited the best meaures. Comparing to its single-stage counterpart, it generally showed lower standard deviations while achieved the target returns. Number of months that experienced portfolio losses is also much less. The two-stage model showed only totally 4 monthly losses in all simulations. In an aggressive case of $\mu = 7\%$, the two-stage model

Table 5. Statistics of realized monthly returns

KPIs of monthly realized return		
	Single stage	**Two stage**
μ =5%		
Geometric Mean	6.79%	6.89%
Standard Deviation	4.81%	4.75%
Monthly Loss	2	1
μ =6%		
Geometric Mean	7.99%	8.08%
Standard Deviation	5.87%	6.22%
Monthly Loss	1	2
μ =7%		
Geometric Mean	9.69%	9.50%
Standard Deviation	7.67%	7.17%
Monthly Loss	2	0
μ =8%		
Geometric Mean	12.01%	10.95%
Standard Deviation	9.83%	8.37%
Monthly Loss	2	1

KPIs of monthly realized return		
Momentum	**3-month formulation**	**6-month formulation**
Geometric Mean	8.03%	4.61%
Standard Deviation	10.72%	10.95%
Monthly Loss	5	6
DMAC	**ST=1 LT=3**	**ST=1 LT=6**
B=0.025		
Geometric Mean	6.67%	5.93%
Standard Deviation	20.36%	12.77%
Monthly Loss	6	5
B=0.05		
Geometric Mean	-4.78%	7.47%
Standard Deviation	30.78%	19.57%
Monthly Loss	8	6

enjoyed a perfect performance during the full trading period as positive monthly return was recored in all trading months. These observations suggest that stochastic programing approach considering the return joint distributions in the past is a superior tool in the dynamic sequential decision making problems in Chinese commodity future markets

We show the computational complexity in generating a scenario tree and solving model in multi-stage case finally. The model was solved in a regular personal laptop with Intel i7 CPU and 8 GB ram. Alghough all procedures were automated in SAS and AMPL, it took 45 miniutes to generate a two-stage tree with 15,000 (150x100) scenarios in SAS and 3 miniutes to solve the model in AMPL. The action was repeated 18 times for a simulation. The soluton time will be increared proportional with the increaing number of scenarios. To increase the computational efficiency, the solution of the model can be moved to paralled computin system, which generally provided by professional academic and commecial institutes.

VI. Conclusions

We presented a stochastic programming model for dynamic commodity future portfolio managements and demonstrated its feasibility and flexibility by applying the solution in the Chinese commodity future market. Through extensive simulations in the real market data, we compared the performance of stochastic programming model with the results generated by two traditional trend following strategies. The stochastic programming approach consistently achieved better performances in terms of high growth of the portfolio value and stability of returns, while the multi-stage case achieved best results as additional information was captured in the simulation of future outcomes.

Our results also demonstrated that the CVaR is an ideal risk measurement for commodity future contracts. The CVaR objective in our model minimize the excess shortfall beyond VaR in the end of planning horizon and suitable for asymmetric distributions which were shown in the commodity return distributions.

Scenario trees representing the uncertainties of future outcomes were served as a key input in the solution of stochastic programming model. We made an improvement in a scenario generation method proposed by Kaut and Wallace (2003). The improved moment matching algorithm made it possible to generate scenario tree with arbitrary outcomes that well matched the empirical distributions of historical data. Mean, standard deviation, skewness and kurtosis as well as the correlation matrix of random variables were captured as the approximation of empirical distribution.

A contribution of this study is to develop automated computer programs to solve the complex scenario generation process and stochastic programming models. Our automated computer programs enable non-specialist to build their own models and solve problems using stochastic programming approach.

We used monthly trading period in our model. while profit and losses are settled on a daily basis in future exchanges. However, our scenario generation and modeling approach could be extended to a daily trading period case. The monthly model was served as a pioneer to explore the efficacy of stochastic programming approach in commodity future markets. The daily model could be easily implemented in a platform with higher computational efforts.

Appendix: Algorithm of improved moment-matching technique:

Our moment-matching solution generally follows the procedures provided Høyland and Kaut (2002), but we modified the cubic transformation $\tilde{Y} = a + b\tilde{X} + c\tilde{X}^2 + d\tilde{X}^3$ as follows.

The purpose of the modification is to generate arbitrary n observations of random variable $\tilde{Y}$ with specified first four moments $E\left[\tilde{Y}^k\right], k = 1...4,$ given a sample of random variables $\tilde{X}$ with known first12 moments.

Without loss of generality, we use $x_k, k = 1...12$ to represent the first 12 moments of $\tilde{X}$, and $y_k, k = 1...4$ denote the first 4 moments of $\tilde{Y}$.

The problem is to find a, b, c and d in below equations.

1. $y_1 = a + bx_1 + cx_2 + dx_3$

2. $\frac{n-1}{n} y_2 = d^2 x_6 + 2cdx_5 + \left(2bd + c^2\right) x_4 + \left(2ad + 2bc\right) x_3 + \left(2ac + b^2\right) x_2 + 2abx_1 + a^2$

3. $$\frac{(n-1)(n-2)}{n^2} y_3 = d^3 x_9 + 3cd^2 x_8 + \left(3bd^2 + 3c^2 d\right) x_7 + \left(3ad^2 + 6bcd + c^3\right) x_6 + \left(6acd + 3b^2 d + 3bc^2\right) x_5 + \left(6abd + 3ac^2 + 3b^2 c\right) x_4 + \left(3a^2 d + 6abc + b^3\right) x_3 + \left(3a^2 c + 3ab^2\right) x_2 + 3a^2 bx_1 + a^3$$

4. $$\left(y_4 + \frac{3(n-1)^2}{(n-2)(n-3)}\right) \times \frac{(n-1)(n-2)(n-3)}{n \times n \times (n+1)} = d^4 x_{12} + 4cd^3 x_{11} + \left(4bd^3 + 6c^2 d^2\right) x_{10} + \left(4ad^3 + 12bcd^2 + 4c^3 d\right) x_9 + \left(12acd^2 + 6b^2 d^2 + 12bc^2 d + c^4\right) x_8 + \left(12abd^2 + 12ac^2 d + 12b^2 cd + 4bc^3\right) x_7 + \left(6a^2 d^2 + 24abcd + 4ac^3 + 4b^3 d + 6b^2 c^2\right) x_6 + \left(12a^2 cd + 12ab^2 d + 12abc^2 + 4b^3 c\right) x_5 + \left(12a^2 bd + 6a^2 c^2 + 12ab^2 c + b^4\right) x_4 + \left(4a^3 d + 12a^2 bc + 4ab^3\right) x_3 + \left(4a^3 c + 6a^2 b^2\right) x_2 + 4a^3 bx_1 + a^4$$

Again, we rely on statistical software to solve the unknown parameters a, b, c and d in above equations.

VII. Reference

William T. Ziemba, 2003. “The stochastic Programming Approach to Asset, Liability, and Wealth Management”, The Research Foundation of AIMR

Nikolas Topaloglou, Hercules Vladimirou, Stavros A. Zenios, 2006. “A dynamic stochastic programming model for international portfolio management”. European Journal of Operational Research 185(2008) 1501-1524

Roy Kouwenberg, 2001, “Scenario generation and stochastic programming models for asset liability management”, European Journal of Operational Research 134(2001) 279-292

Libo Yin, Liyan Han, 2013 “International Assets Allocation with Risk Management via Multi-Stage Stochastic Programming”, Computational Economics DOI 10.1007/s10614-013-9365-z

Andrew C. Szakmary, Qian Shen, Subhash C. Sharma, 2009 “Trend-following trading strategies in commodity futures: A re-examination”, Journal of Banking & Finance 34(2010) 409-426

Michal Kaut, Stein W. Wallace, 2003. “Evaluation of scenario-generation methods for stochastic programming”. Stochastic Programming E-Print Series ref. no. 14-2003

Kjetil Høyland, Michal Kaut, 2003. “A Heuristic for Moment-matching Scenario Generation”. Computational Optimization and Applications, vol.24, 169-185.

R. Tyrrell Rockafellar, Stanislav Uryasev, 2002. “Conditional value-at-risk for general loss distributions”, Journal of Banking & Finance 26 (2002) 1443-1471

Robert Fourer, David M. Gay, Brian W. Kernighan, 1990. “A Modeling Language for Mathematical Programming”. Management Science 36 (1990) 519-554