MASTER 2 DATA SCIENCE

INSURANCE PROJECT

MODEL FOR DYNAMIC MULTIPLE OF CPPI STRATEGY WITH APPLICATION OF EXTREME VALUE THEORY

Étudiants :
DJIOMOU NGONGANG Cédric
LE Son Tung
NGUYEN Thi Ha Giang
TRAORE Babou

Encadré par :
Pr. BENEZET Cyril

February 8, 2021

Contents

1 Introduction

Portfolio insurance is a hedging strategy which allows investors to recover at maturity a given percentage of their initial investment. The Portfolio insurance strategy limits downside risk in falling markets, while allowing potential benefits in rising markets.

One of the most typical Portfolio insurance strategy is the Constant Proportion Portfolio Insurance (CPPI). This strategy is based on parameter setting with just only one parameter which is multiple (denoted by m) and operates as static strategies like stop-loss and buy-hold.

Multiple parameter m directly determines the risk exposure. In tradtitional CPPI, this parameter is fixed from the beginning, and doesn't changes with the market conditions. We reallocate the portfolio at the end of each period based on m. This strategy will not fail if re-allocation occurs continuously. But in practice, re-allocation cannot be carried out continuously, or even between a very short period because of transaction costs.

Our project focuses on the dynamic setting of the multiple m, which makes the CPPI strategy more stable with the changes of market but still benefit from growing market.

For building the model with multiple dynamic setting, we will use extreme value theory (EVT). Although extreme values scarcely happen in real data, we can use extreme value theory to obtain the entire shape of losses' tail, which give a better fit for the distribution of the loss.

In this report, we give some basic descriptions of the CPPI strategy and introduce the model for multiple dynamic setting of CPPI with an application of extreme value theory. Then, for empirical studies, we will apply this model for CAC40 weekly data and compare the results with the traditional CPPI setting. In the final section, we will resume the results we have obtained and give the conclusion.

2 CPPI strategy and its basis of modeling

2.1 Overview

The Constant Proportion Portfolio Insurance (CPPI) has been introduced by (Pernold, 1986) and further studied by (Perold Sharpe, 1988) for fixed-income instruments and (Black Jones, 1987) for equity instruments. This strategy is based on dynamic asset allocations over time. It mainly allocate between two asset classes: risk-free asset such as T-bills, obligation and risky asset such as equity, commondity.

An investor starts by setting a floor which equals to the lowest acceptable value of the portfolio. Then, he determines the cushion as the portion exceeds the portfolio's value relative to the floor. The amount allocated to the risky asset is equal to the cushion multiplied by a predetermined *multiple* which mesure the investor's exposure.

According to the work of Black and Jones, the CPPI strategy has the advantage of being more simple and flexible than other portfolio strategies. For instance, the risk exposure can be chosen according to the risk appetite of investors when there is no maturity. In addition, according to Kingston, CPPI strategy is optimal with the decreasing of absolute and relative risk aversion. Moreover, in comparison with portfolio insurance strategies based on option replication, which need complicated option pricing techniques, CPPI strategy is much easier in operation because it is only based on parameters setting.

2.2 *Gap risk* and multiple settings

The setting of the multiple m is significant in operation of CPPI strategy, as it is directly related to the risk profile. The bigger m is, the more the portfolio participates in market. However, this leads to a bigger risks; as the price of risky assets falls, the portfolio value will drop. Therefore, the research on CPPI strategy mainly focused on the setting of the multiple m. For example, Focusing on the parameter "Multiple" of CPPI strategy, (Guangyuan Xing et al.,2014)[3] proposes a dynamic setting model of multiple for gap risk management purpose.

In the traditionnal CPPI strategy, the multiple is considered as a fixed value m and stays constant even if the market goes up or crashes. Some researchs has shown that the CPPI strategy can be more profitable by using the dynamic setting principles. However, using frequently rebalancing will lead to some *gap risks*.

The CPPI *gap risk* is defined as the risk that the market crashed between two rebalancing dates (cf. lecture). The maximum crash allowed in a period of adjustment is $1/m$. Therefore the gap risk can also be defined as the probability of the risk asset loss more than or equal to $1/m$.

2.3 Mathematical notations

The first main portfolio insurance method was the option based portfolio insurance (OBPI) introduced in 1976 by Leland and Rubinstein. It consists of a portfolio invested in a risky asset (usually a financial index) covered by a listed put option written on it.

To avoid the complexity and inconvenience of the OBPI strategy, (Black Jones, 1987)[2] proposed the constant proportion portfolio insurance strategy (CPPI). Investors refer to the difference between the present value of the insured portfolio and the current value of maturity floor as the expected loss. They choose the risk multiplier m according to the tolerance of the risk and use the simple dynamic formula to adjust the position of risky asset and riskless asset.

We denote:

- V_0: the initial investment volume,

- T: the maturity,
- N: the terminal guaranteed value,
- r_f: the risk free interest rate,
- V_t: the total portfolio value at time t,
- E_t: the amount invested in risky asset,
- M_t: the residual amount invested in risk free asset at time t,
- m: the risk multiplier(or the Multiple),
- P_t: the present value of the floor at time t,
- $C_t = V_t - P_t$: the cushion at time t.

At the beginning, the terminal guaranteed value N and the multiplier m are decided according to the investor's risk tolerance, which are generally fixed through the whole time period. The higher the multiplier, the more the investor will benefit from an increase in the risky asset price but more risk when market crashes.

The floor P_t is discounted value the terminal guaranteed value N with the risk-free rate r_f:

$$P_t = N \times (1 + r_f)^{-(T-t)}$$

There are two cases of the balancing of portfolio at time t:

- if $V_t > P_t$: The amount invested in risky asset is given by:

$$E_t = m_t = m \times (V_t - P_t)$$

- if $V_t \leq P_t$: All portfolio is invested in non risky asset

$$E_t = 0 \text{ and } M_t = V_t$$

In the second case, the value of the portfolio at the maturity V_T can not attain the guaranteed value N. Thus, this strategy fails.

2.4 Basis of modeling

Define the discrete time of asset allocation in the CPPI strategy as $\{t_0^n = 0 < t_1^n < ... < t_n^n = T, t_{k+1}^n - t_k^n = \frac{T}{n}\}$, where n denote the number of rebalancing.

Let $x_k = \dfrac{P_{t_k} - P_{t_{k+1}}}{P_{t_k}}$ be the loss of the risky asset in adjustment period (t_k, t_{k+1}).
As we see above, the CPPI strategy will fail if the value of portfolio falls below the value of floor $V_t < P_t$. Based on this, we can define the maximum loss allowed between 2 rebalancing dates as:

$$\begin{aligned} V_{t_{k+1}} &= m_{t_k}(1 - x_k) + (V_{t_k} - m_{t_k})(1 + r_f) \\ &= m \times (V_{t_k} - P_{t_k})(1 - x_k) + (V_{t_k} - m \times (V_{t_k} - P_{t_k}))(1 + r_f) \\ &\geq P_{t_{k+1}} \end{aligned}$$

For simplicity, we consider: r_f small and $P_{t_{k+1}} \approx P_{t_k}$ for $\frac{T}{n}$ small. We got the maximum loss allowed between 2 rebalancing: $x_k \leq \frac{1}{m}$.

3 Multiple dynamic setting model in CPPI strategy

For this project, we will consider x_k follows GARCH(1,1) which can capture some main characteristics of the returns of financial assets.

$$\begin{cases} x_k = \mu + \sigma_k z_k \\ \sigma_k^2 = \omega + \alpha_1 x_{k-1}^2 + \beta_1 \sigma_{k-1}^2 \end{cases}$$

where μ is the expected loss of the risky asset and residual terms z_k are independent and identically distributed white noise.

The measurement of gap risk is shown in the following formula

$$\Pr(x_k \geq \frac{1}{m_k}) = \alpha$$

where α is the probability of the loss of the risky asset exceeding $\frac{1}{m_k}$. As we needed to control the gap risk less than α, $\Pr(x_k < 1/m_k)$ must be equal to $(1-\alpha)$. This is equivalent to

$$\Pr(\frac{x_k - \mu}{\sigma_k} < \frac{\frac{1}{m_k} - \mu}{\sigma_k}) = 1 - \alpha$$

$$\Leftrightarrow \Pr(z_k < \frac{\frac{1}{m_k} - \mu}{\sigma_k}) = 1 - \alpha$$

Hence,

$$m_k = \left[\mu + \sigma_k F_Z^{-1}(1 - \alpha)\right]^{-1}$$

where F_Z is the distribution function of z_k. We can obtain μ and σ_k from estimation of the model GARCH(1,1). What is left is searching for an estimation of $F_Z^{-1}(1 - \alpha)$.

Originally in a GARCH(1,1) model, z_k are i.i.d $\sim \mathcal{N}ormal$. However, as the loss of risky assets is almost never normally distributed, we will use another assumption for z_k.

Assume that z_k follows t-distribution. In addition, we will use extreme value theory to describe z_k, in which we assume its tail follows pareto distribution (GPD). With these assumption, we can obtain an estimation of distribution of z_k which have the characteristics of Leptokurtic, skewness, etc. of the loss.

Now, for estimate F_Z^{-1}, we will investigate the tail area of z_k distribution.[1]

Denote u the threshold. Let N be the number of sample observations and N_u be the number of sample observations exceeding u. Let $y = z - u$, we have

$$F_u(y) = \Pr(Z - u \leq y | Z > u) = \frac{F_Z(z) - F(u)}{1 - F(u)}$$

As $u \to z_+$, we have $F_u(y) \to G_{\epsilon,u}(y)$, where z_+ is the extreme point of F_Z and

$$G_{\epsilon,u}(y) = \begin{cases} 1 - (1 + \frac{y}{u})^{-\epsilon} & \text{for } y \geq 0 \\ 0 & \text{for } y < 0 \end{cases}$$

Therefore,

$$\begin{aligned} F_Z(z) &= [1 - F(u)] \times G_{\epsilon,u}(y) + F(u) \\ &= \frac{N_u}{N} \times \left[1 - \left(1 + \frac{y}{u}\right)^{-\epsilon}\right] + \left(1 - \frac{N_u}{N}\right) \\ &= 1 - \frac{N_u}{N} \times \left(1 + \frac{z - u}{u}\right)^{-\epsilon} \end{aligned}$$

Consider $F_Z(z) = 1 - \alpha$. This implies that

$$\frac{z}{u} = \left(\alpha \frac{N}{N_u}\right)^{-1/\epsilon}$$

Hence, we obtain

$$F_Z^{-1}(1 - \alpha) = u\left(\frac{N}{N_u}\alpha\right)^{-1/\epsilon}$$

and

$$m_k = \left[\mu + \sigma_k u\left(\frac{N}{N_u}\alpha\right)^{-1/\epsilon}\right]^{-1}$$

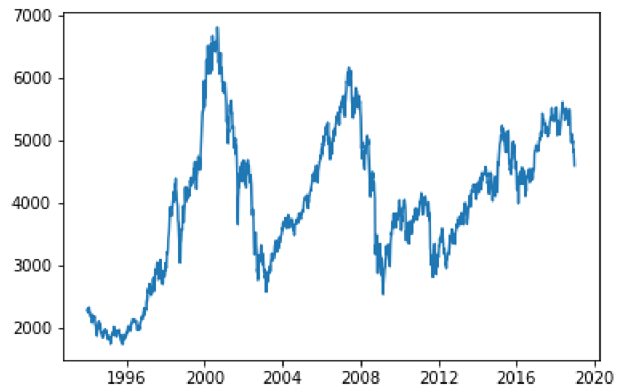
.

4 Empirical studies

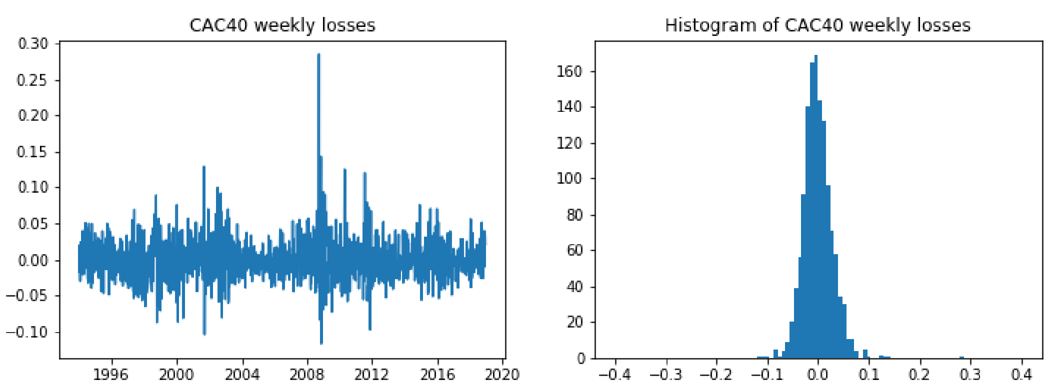
For this project, we used the CAC40 weekly data from 1993 to 2020. We split this data into 2 parts: data in the period 1993 - 2018 for estimating the parameters and data in 2018 - 2020 to see the results in reality.

4.1 Data descriptions

Firstly, we will investgate characteristics of the chosen data. The graph below shows the weekly prices of CAC40 in 25 years, from 1993 to 2018.



We focus on studying the statistical characteristics of the weekly losses.



After visualizing and plot the histogram of weekly losses, we calculate some statistical measure of the losses in this 25-year period and obtain the below table. In this table also includes the result from the normality tests LJungBox and Adf.

Statistical indicator	Value
Mean	-0.0001165
Maximum	0.2846727
Standard deviation	0.0294074
Skewness	1.0316023
Kurtosis	7.8337146
LJungBox test	27.236(0.002)
Adf test	-9.323(0.000)

We see that the losses has a heavy tail distribution, since the Kurtosis>3. The results of the two tests for normality shows that we can reject the hypothesis H_0: the weekly losses follow normal distribution.

These characteristics allow us to use GARCH(1,1) for modeling these losses in the next part.

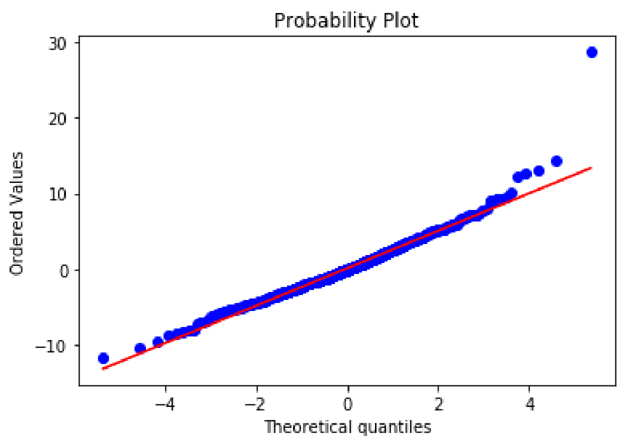
4.2 Model estimations

4.2.1 Estimation of GARCH(1,1)

We assume that the innovation term z_k follows Student's t-distribution corresponding to the heavy-tailed distribution. We obtain the estimated parameters for μ, ω, α_1, β_1 of GARCH(1,1).

```
                      Constant Mean - GARCH Model Results
==============================================================================
Dep. Variable:                  Adj Close   R-squared:                  -0.002
Mean Model:                 Constant Mean   Adj. R-squared:             -0.002
Vol Model:                          GARCH   Log-Likelihood:           -3109.20
Distribution:      Standardized Student's t   AIC:                     6228.41
Method:                Maximum Likelihood   BIC:                       6254.27
                                            No. Observations:             1304
Date:                    Mon, Feb 08 2021   Df Residuals:                 1299
Time:                            11:49:21   Df Model:                        5
                               Mean Model
=====================================================================
                 coef    std err          t      P>|t|     95.0% Conf. Int.
---------------------------------------------------------------------
mu            -0.1501  6.682e-02     -2.247  2.465e-02 [ -0.281,-1.917e-02]
                             Volatility Model
====================================================================
                 coef    std err          t      P>|t|    95.0% Conf. Int.
--------------------------------------------------------------------
omega          0.2041  8.031e-02      2.541  1.106e-02 [4.665e-02,  0.361]
alpha[1]       0.0989  2.112e-02      4.685  2.801e-06 [5.755e-02,  0.140]
beta[1]        0.8780  2.430e-02     36.132 7.038e-286   [  0.830,  0.926]
                              Distribution
==================================================================
                 coef    std err          t      P>|t|  95.0% Conf. Int.
------------------------------------------------------------------
nu            10.5175      3.588      2.932  3.372e-03 [  3.486, 17.549]
==================================================================
```

From this estimation, we also obtain z_k. Then, we use the qqplot to assess our assumption for the t-distribution of z_k.



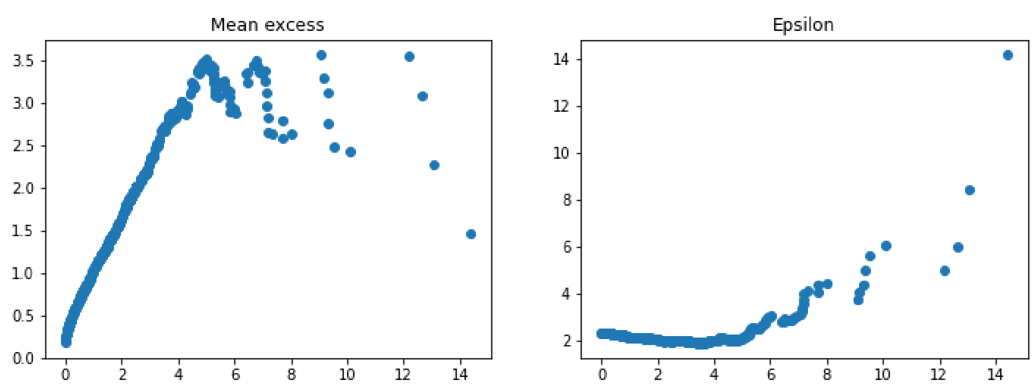
We see that the t-distribution is not good enough for describing the distribution of large values of innovation terms. So we will consider the extreme value theory.

4.2.2 Estimation of the distribution of large innovation terms

We have the Pareto distribution function:

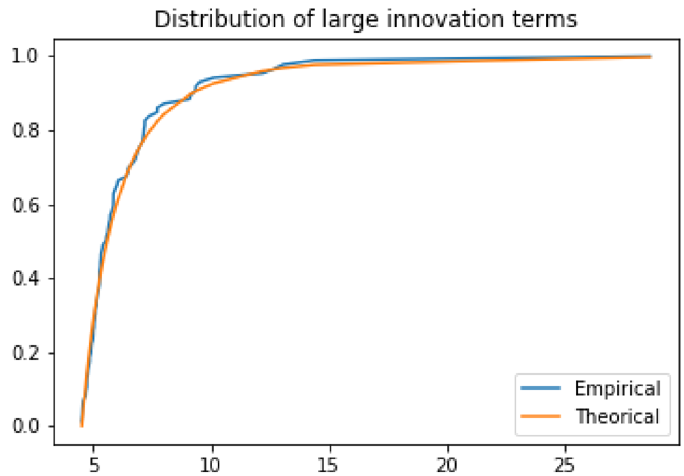
$$G_{\epsilon,u}(y) = \begin{cases} 1 - (1 + \frac{y}{u})^{-\epsilon} & \text{for } y \geq 0 \\ 0 & \text{for } y < 0 \end{cases}$$

For estimation of this distribution, we choose a threshold u and calculate the correspond value of parameter ϵ



We choose $u \approx 4.5$ and $\epsilon \approx 3.226$. This choice can be justified by the mean-excess plot and the plot of corresponding ϵ above.

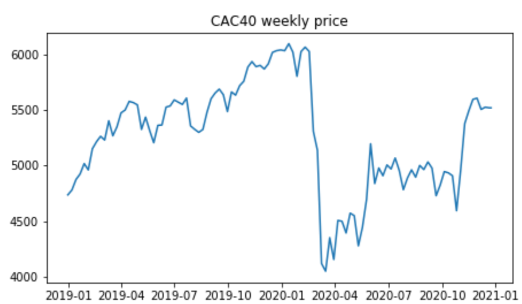
Then we plot this pareto distribution against the empirical distribution of z_k:

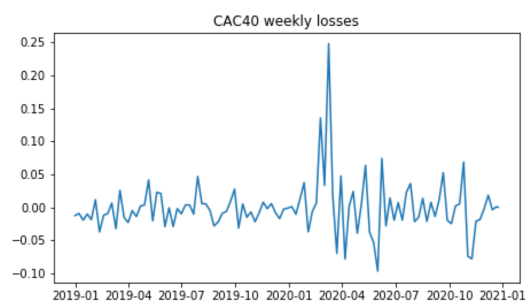


This graph shows that our choice of parameters are quite good.

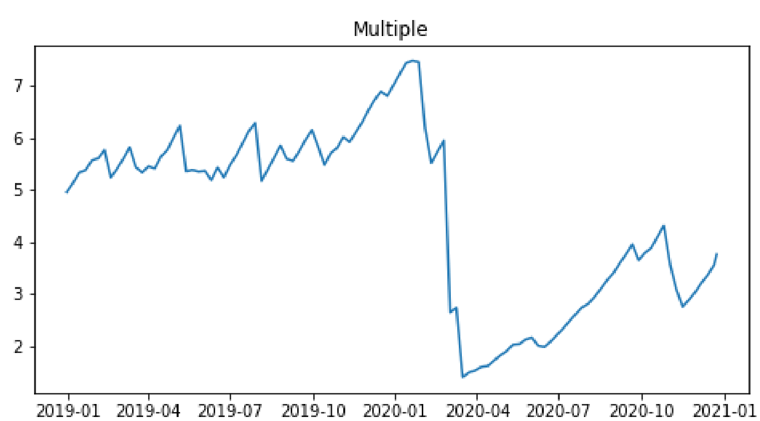
4.3 Dynamic multiple CPPI of CAC40 from 2018 to 2020

We perform this strategy for CAC40 in the period 2018 - 2020:



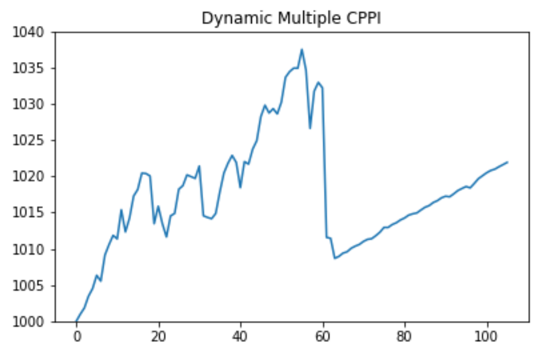


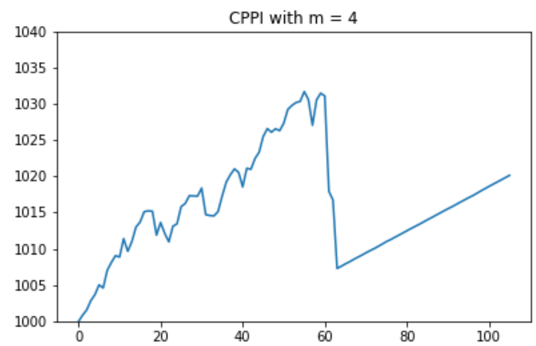
Choose $\alpha = 1\%$ to calculate the multiple m_k



We can see that in the early period, the value of multiple m_k is between 5 and 8, which give us benefits from the upward market. When the market went down, the multiple decrease fast to reduce the associated risk.

With $V_0 = 1000$, $N = 1020$ and the riskfree rate $rf = 0.03\%$, we calculate the performance of the portfolio in two setting: dynamic multiple CPPI with the above calculation and the traditional CPPI with the constant multiple $m = 4$.Here are the graphs of the value of portfolio in these two settings:





The dynamic setting gives us more return in the upward period of the market. This setting also keeps the strategy from failing when the market crashed, whereas the constant m leads to a failure of the strategy.

5 Conclusion

This project focuses on the model for dynamic setting of the multiple of CPPI strategy. By using GARCH(1,1) - EVT approach, we find the value of dynamic multiple as a function of time-varying volatility expected loss and the probability of extreme events of risky asset. The investors can control the associated risk by the value of *alpha*. In the empirical implementation on CAC40, the dynamic multiple model gives us a better results in both two aspects: returns and risk management.

As proposed in (Guangyuan Xing et al.,2014)[3], SV model is a better choice in comparison to GARCH(1,1) for modeling the dynamic movement of financial loss. In addition, for EVT approach, Generalized Pareto distribution can be considered instead of Pareto distribution. However, these assumption demand a much more complex estimation methods such as Markov chain Monté Carlo.

References

[1] Alexander J. McNeil Rudiger Frey. *Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: an Extreme Value Approach*, 2000.

[2] F. Black R. Jones. *Simplifying portfolio insurance for corporate pension plans*, 1988.

[3] Guangyuan Xing Yong Xue Zongxian Feng Xiaokang Wu. *Model for Dynamic Multiple of CPPI Strategy*, 2014.