Predicting Chinese Stock Market Crashes

Sébastien Lleo[*] and William T. Ziemba[†]

Tuesday 1st December, 2015

[*]Finance Department, NEOMA Business School - Campus Reims, 59 rue Pierre Taittinger, 51100 Reims, France, Email: sebastien.lleo@NEOMA-bs.fr

[†]Alumni Professor of Financial Modeling and Stochastic Optimization (Emeritus), University of British Columbia, Vancouver, BC and Distinguished Visiting Associate, Systemic Risk Centre, London School of Economics, Email: wtzimi@mac.com

Predicting Chinese Stock Market Crashes

Abstract

Predicting stock market crashes has been a focus of interest for both researchers and practitioners. Over the years, several prediction models have been developed, mostly for use on mature financial markets. In this paper, we investigate whether traditional crash predictors, the price-to-earnings ratio and the Bond-Stock Earnings Yield Differential model, work for the Shanghai Stock Exchange Composite Index.

Keywords: stock market crashes, Shanghai Stock Exchange, Bond-Stock Earnings Yield Differential (BSEYD), price-earnings-ratio, Cyclically-Adjusted Price Earnings ratio (CAPE).

JEL classification: C12, C13, C15, G14, G15.

Through the summer of 2015, the gyrations of the Shanghai stock exchange captured the headlines of the financial press. In fact, what has been labeled the "2015 Chinese stock market crash" is just the latest in a series of eighteen major downturns in the twenty-five years of the Chinese stock market history. Headlines aside, the Chinese stock market is certainly one of the most interesting equity markets in the world by its size, scope, structure and recency. These features have a deep influence on the behavior and returns of the Chinese stock market.

First, we discuss four key stylized facts on the return distribution of the Shanghai Stock Exchange Composite Index (SHCOMP). Then, we explain how equity downturn and crash prediction models work, and how to test their accuracy. The construction process for the signal and hit sequence is crucial to ensure that the crash prediction models produce out of sample predictions free from look-ahead bias. It also eliminates data snooping by setting the parameters *ex ante*, with no possibilities of changing them during the analysis. More importantly, the construction process removes the effect of autocorrelation, making it possible to test the accuracy of the measures using standard statistical techniques. We also conduct a Monte Carlo study to address small sample bias.

In this paper, we test whether the price-to-earnings ratio (P/E) based on current earnings, the Bond-Stocks Earnings Yield Differential model (BSEYD) and the Cyclically Adjusted Price-to-Earnings ratio (CAPE), accurately predicts the downturns of the SHCOMP. We find that the logarithm of the P/E has successfully predicted crashes over the entire length of the study (1990-2015). On a shorter 9-year period (2006-2015), we find mixed evidence of the predictive ability of the BSEYD models and CAPE. Overall, this study provides supporting evidence for the application of crash prediction models to the Chinese market.

I. A Brief Overview of the Chinese Stock Market

Mainland China has two stock exchanges, the Shanghai Stock Exchange (SSE, 上海证券交易所) and the Shenzhen Stock Exchange (深圳证券交易所). The Shanghai Stock Exchange is the larger of the two. It is also the fifth largest stock market in the world by market capitalization. The modern Shanghai Stock Exchange officially came into being on November 26, 1990 and started trading on December 19, 1990. The Shenzhen Stock Exchange was formally founded on December 1, 1990, and it started trading on July 3, 1991. While the largest and most established companies usually trade on the Shanghai Stock Exchange, the Shenzhen Stock Exchange is home to smaller and privately-owned companies. Taken together, the Shanghai and Shenzhen Stock Exchanges represent the second largest stock market in the world after the New York Stock Exchange.

On November 17, 2014, the Chinese government launched the Shanghai-Hong Kong Stock Connect (沪港通) to enable investors in either market to trade shares on the other market. This initiative heralds closer integration between securities markets in China.

Chinese companies may list their shares under various schemes, either domestically or abroad. Domestically, companies may issue:

- *A-shares*: common stocks denominated in Chinese Reminbi and listed on the Shanghai or Shenzhen stock exchanges.

- *B-shares*: special purpose shares denominated in foreign currencies but listed on the domestic stock exchange. Until 2001, only foreign investors had access to B-shares.

In addition to B-shares, foreign investors interested in the Chinese equity market may also buy:

- *H-shares*: shares denominated in Hong Kong Dollars and traded on the Hong Kong Stock Exchange.
- *L-chips, N-chips and S-chips*: shares of companies with significant operations in China, but incorporated respectively in London, New York and Singapore.
- *American Depository Receipts (ADRs)*: an ADR is a negotiable certificate issued by a U.S. bank representing a specified number of shares in a foreign stock traded on an American exchange. As of October 2015, there were around 110 Chinese ADRs listed on American exchanges and another 200 Chinese ADRs on American over-the-counter markets.

The diversity of investment schemes available shows that the Shanghai and Shenzhen Stock Exchange are a large, crucial part of the Chinese equity market, but do not represent the whole market. For example, there are also *red chips* (shares of companies incorporated outside mainland China but owned or substantially controlled by Chinese state-owned companies) and *P-chips* (shares of companies owned by private individuals and traded outside mainland China, for example on the Hong Kong stock exchange). Our study focuses on equity market downturns on the Shanghai Stock Exchange.

II. Four Key Stylized Facts

The Shanghai Stock Exchange Composite Index (上海证券交易所综合股价指), or SHCOMP (上证综指), is the main Chinese stock index. In October 2015, the SHCOMP consisted of the shares of 1,070 Chinese companies.

We observe and discuss four key stylized facts on the historical distribution of daily log returns on the SHCOMP. Undoubtedly, various aspects of the index are also interesting and warrant a thorough analysis similar to Cont's analysis of the S&P500 (Cont, 2001), but this is beyond the scope of this paper.

A. Stylized Fact 1: The Return distribution is highly volatile, right skewed and has very fat tails

Figure 1 displays the evolution of the SHCOMP since its launch on December 19, 1990, as well as the distribution of daily log returns on the index.

Table 1 shows that over the entire period, the daily log return on the SSE averaged 0.06%, with a median return of 0.07%. The lowest and highest daily returns were respectively -17.91% and +71.92%. The exhibit also gives the corresponding statistics at a weekly and monthly frequency.

The returns are highly volatile: the standard deviation of daily returns is 2.40%, equivalent to around 40 times the mean daily return. The distribution of daily returns is positively skewed (skewness = 5.26) with very fat tails (kurtosis = 149). As a result, the Jarque-Bera statistic reaches 5,419,808, rejecting normality at any level of significance. The Jarque-Bera statistic also leads to a strong rejection of normality for weekly and monthly data. The aggregational gaussianity, the phenomenon in which the empirical distribution of log-returns tends to normality as the time scale Δt over which the returns are calculated increases, observed by Cont (2001) on the S&P500, has a much weaker effect on the SHCOMP.

[Place Figure 1 here]

[Place Table 1 here]

B. Stylized Fact 2: Log returns on the SHCOMP do not exhibit significant autocorrelation

Figure 2 shows that the autocorrelation of daily log returns is low. The autocorrelation does not appear statistically meaningful and partial autocorrelation up to lag 20 stays within the interval $[-0.03, 0.06]$. This implies that the SHCOMP does not have a short-term memory: today's return does not help to forecast tomorrow's return.

[Place Figure 2 here]

C. Stylized Fact 3: No fewer than six states are necessary to capture the evolution of the returns with a Gaussian Hidden Markov Chain

Hidden Markov Models (HMMs) are a useful way to model the behavior of a physical or economic system when we suspect that this behavior is determined by the transition between a finite number of underlying but unobservable "regimes" or "states." We refer the reader to Rabiner (1989) and Rabiner and Juang (1993) for an excellent tutorial on HMMs.

The simplest, and often the best, HMM models are Gaussian Hidden Markov Chains. In these models, the returns in each state are normally distributed, but the parameters of the distribution depend on the state. As the state changes over time, the returns will be drawn from different distributions, resulting in an aggregate distribution that bears little resemblance to a normal one. Gaussian HMMs are estimated via the Baum-Welch algorithm (Baum, Petrie, Soules, and Weiss, 1970), an application of the well-known EM algorithm (see Dempster, Laird, and Rubin, 1977).

One of the difficulties is to find the optimal number of states for the model. To that end, it is customary to use an information criterion such as the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC) to discriminate between model formulations. The optimal model will minimize the criterion.

Table 2 presents the LogLikelihood, AIC and BIC for a fitted HMM with one to seven states. Contrary to the LogLikelihood, the AIC and BIC penalize the model for the number of parameters used. For the SHCOMP, we find that the optimal model specification, the specification that minimizes the AIC and BIC, is a six state model.

[Place Table 2 here]

The transition probability matrix P equals

$$\begin{pmatrix} 8.7520e-01 & 3.9815e-03 & 1.1249e-01 & 1.5765e-27 & 7.1677e-03 & 1.1572e-03 \\ 8.6346e-51 & 2.3576e-01 & 1.0069e-04 & 7.5280e-01 & 2.1882e-193 & 1.1332e-02 \\ 3.5656e-02 & 2.2829e-02 & 9.4090e-01 & 4.6334e-06 & 6.1205e-04 & 6.8722e-51 \\ 1.0453e-45 & 4.4398e-01 & 2.5787e-02 & 5.3023e-01 & 3.4880e-236 & 1.7463e-20 \\ 2.6672e-01 & 1.9219e-228 & 7.0054e-187 & 4.8388e-298 & 7.3328e-01 & 1.5823e-238 \\ 4.2126e-108 & 1.3715e-01 & 3.2580e-83 & 1.1485e-30 & 2.3927e-275 & 8.6285e-01 \end{pmatrix}$$

The initial probability and the parameters of the normal distribution for each state are given in Table 3.

[Place Table 3 here]

Finally, we use the Vitterbi (1967) aalgorithm to backtrack the evolution of the HMM for the SHCOMP. The most likely historical path of the HMM across the six states is shown in Figure 3.

[Place Figure 3 here]

D. Stylized Fact 4: Downturns and large market movements occur frequently

The return distribution of the SHCOMP has fat tails, which indicates that extreme events are more likely to occur than a Normal distribution would predict. In fact, we counted 26 market movements with cumulative returns of 10% or more and 24 market movements with losses of 10% or more in the 25 years since the SHCOMP started trading.

Earlier studies, such as Lleo and Ziemba (2015a,b), defined an equity market downturn or crash as a decline of at least 10% from peak to trough based on the closing prices for the day, over a period of at most one year. Given the frequency of large market movements in the SHCOMP, we redefine an equity market downturn or crash as a decline of at least 20% in the level of the SHCOMP from peak to trough based on closing prices for the day, over a period of at most one year (252 trading days).

We identify a correction on the day when the daily closing price crosses the 20% threshold. The identification algorithm is as follows:

1. *Identify all the local troughs in the data set.* Today is a local trough if there is no lower closing price within $\pm d$ business days.
2. *Identify the crashes.* Today is a crash identification day if all of the following conditions hold:
   (a) The closing level of the SHCOMP is down today at least 20% from its highest level within the past year, and the loss was less than 20% yesterday;

(b) This highest level reached by the SHCOMP prior to the present crash differs from the highest level corresponding to a previous crash; and

(c) This highest level occurred after the local trough that followed the last crash.

The objective of these rules is to guarantee that the downturns we identify are distinct. Two downturns are not distinct if they occur within the same larger market decline. Although these rules might be argued with, they have the advantage of being unambiguous, robust and easy to apply.

Table 4 presents the 18 downturns that occurred between December 19, 1990 and October 31, 2015. On average, a downturn lasted 199 days and caused a 35.1% decline in the value of the SHCOMP.

[Place Table 4 here]

III. How Equity Downturn Prediction Models Work

From these stylized facts, it is clear that the SHCOMP behaves differently from the mature markets in Europe and North America. In order to apply the equity downturn prediction models to the SHCOMP, we need to examine the inner workings of these models: how they are constructed, how to convert them into a testable model, and how to test the accuracy of their predictions.

The construction process for the signal and hit sequence is crucial to ensure that the crash prediction models produce out of sample predictions free from look-ahead bias. It also eliminates data snooping by setting the parameters *ex ante* during the signal construction, with no possibilities of changing them when we construct the hit sequence. More importantly, the construction of the hit sequence removes the effect of autocorrelation, making it possible to test the accuracy of the measures using a standard likelihood ratio test. In addition to the standard likelihood ratio test using the asymptotic χ^2 distribution, we conduct a Monte Carlo study on the empirical distribution to address small sample bias.

A. Signal Construction

Equity market crash prediction models such as the BSEYD (Ziemba and Schwartz, 1991; Lleo and Ziemba, 2012, 2015b), the high P/E model (Lleo and Ziemba, 2015b), the variations on Warren Buffett's market value-to-GNP measure (Lleo and Ziemba, 2015a), or the continuous time disorder detection model (Shiryaev, Zitlukhin, and Ziemba, 2014, 2015) generate a signal to indicate a downturn in the equity market at a given horizon h. This signal occurs whenever the value of a crash measure crosses a threshold. Given a prediction measure $M(t)$, a crash signal occurs whenever

$$SIGNAL(t) = M(t) - K(t) > 0 \tag{1}$$

where $K(t)$ is a time-varying threshold for the signal.

Three key parameters define the signal: (i) the choice of measure M(t); (ii) the definition of threshold $K(t)$; and (iii) the specification of a time interval H between the occurrence of the signal and that of an equity market downturn.

We test the measures using two time-varying thresholds: (1) a dynamic confidence interval based on a Normal distribution; and (ii) a dynamic confidence interval using Cantelli' s inequality - see Problem 7.11.9 in Grimmett and Stirzaker (2001) for a statement of the mathematical result, and Lleo and Ziemba (2012, 2015b) for applications to crash predictions.

To construct the confidence intervals, we compute the sample mean and standard deviation of the distribution of the measures as a moving average and a rolling horizon standard deviation respectively. Using rolling horizon means and standard deviations has the advantage of providing data consistency. Importantly, this construction is purely in-sample. The h-day moving average at time t, denoted by μ_t^h, and the corresponding rolling horizon standard deviation σ_t^h are

$$\mu_t^h = \frac{1}{h}\sum_{i=0}^{h-1} x_{t-i}, \qquad \sigma_t^h = \sqrt{\frac{1}{h-1}\sum_{i=0}^{h-1}(x_{t-i} - \mu_t^h)^2}.$$

We establish the one-tailed confidence interval at a 95% level. This corresponds to 1.645 standard deviations above the mean in the Normal distribution.

Cantelli's inequality relates the probability that the distance between a random variable X and its mean μ exceeds a number $k > 0$ of standard deviations σ to provide a robust confidence interval:

$$P\left[X - \mu \geq k\sigma\right] \leq \frac{1}{1+k^2}.$$

Setting $\alpha = \frac{1}{1+k^2}$ yields $P\left[X - \mu \geq \sigma\sqrt{\frac{1}{\alpha} - 1}\right] \leq \alpha$. The parameter α provides an upper bound for a one-tailed confidence level on any distribution. In our analysis, the horizon for the rolling statistics is h = 252 days. We select $\alpha = 25\%$ which produces a slightly higher threshold than the standard confidence interval. In a Normal distribution, we expect 5% of the observations to lie in the right tail, whereas Cantelli's inequality implies that the percentage of outliers in a distribution may reach up to 25%.

The last parameter we need to specify is the horizon H. Recall that the crash identification time is the date by which the SHCOMP has declined by at least 20% in the last year (252 trading days). We define the local market peak as the highest level reached by the market index within 252 trading days before the crash. We set the horizon H to a maximum of 252 trading days prior to the crash identification date.

B. Construction of the Hit Sequence X

Crash prediction models have two components: (1) a signal, which takes the value 1 or 0 depending on whether the measure has crossed the confidence level, and (2) a crash indicator, which takes the value 1 when an equity market correction occurs and 0 otherwise.

From a probabilistic perspective, these components are Bernoulli random variables, but they exhibit a high degree of autocorrelation, that is, a value of 1 (0) for the crash signal is more likely to be followed by another value of 1 (0) on the next day. This autocorrelation makes it difficult to test the accuracy of the model.

To remove the autocorrelation effect, we define a signal indicator sequence $S = \{S_t, t = 1, \ldots, T\}$. This sequence records as the signal date the first day in a series of positive signals, and it only counts distinct signal dates. Two signals are distinct if a new signal occurs more than 30 days after the previous signal. The objective is to have enough time between two series of signals to identify them as distinct. The signal indicator S_t takes the value 1 if date t is the starting date of a distinct signal, and 0 otherwise. Thus, the event "a distinct signal starts on day t" is represented as $\{S_t = 1\}$. We express the signal indicator sequence as the vector $s = (S_1, \ldots, S_t, \ldots, S_T)$. This construction effectively removes the effect of autocorrelation.

For the crash indicator, we denote by $C_{t,H}$ the indicator function returning 1 if the crash identification date of at least one equity market correction occurs between time t and time $t + H$. We identify the vector C_H with the sequence $C_H := \{C_{t,H}, t = 1, \ldots, T - H\}$ and define the vector $c_H := (C_{1,H}, \ldots, C_{t,H}, \ldots C_{T-H,H})$. The number of correct predictions n is defined as

$$n = \sum_{t=1}^{T} C_{t,H} = \mathbf{1}' c_H.$$

The accuracy of the crash prediction model is the conditional probability $P(C_{t,H} = 1 | S_t = 1)$ of a crash being identified between time t and time $t + H$, given that we observed a signal at time t. The higher the probability, the more accurate the model.

We use maximum likelihood to estimate this probability and to test whether it is significantly higher than a random guess. We obtain a simple analytical solution because the conditional random variable $\{C_{t,H} = 1 | S_t = 1\}$ is a Bernoulli trial with probability $p = P(C_{t,H} = 1 | S_t = 1)$.

To estimate the probability p, we change the indexing to consider only events along the sequence $\{S_t | S_t = 1, t = 1, \dots T\}$ and denote by $X := \{X_i, i = 1, \dots, N\}$ the "hit sequence" where $x_i = 1$ if the ith signal is followed by a crash and 0 otherwise. Here N denotes the total number of signals, that is

$$N = \sum_{t=1}^{T} S_t = \mathbf{1}' s$$

where $\mathbf{1}$ is a vector with all entries set to 1 and v' denotes the transpose of vector v. The sequence X can be expressed in vector notation as $x = (X_1, X_2, \dots, X_N)$. The empirical probability p is the ratio n/N.

C. Maximum Likelihood Estimate of $p = P(C_{t,H}|S_t)$ and Likelihood Ratio Test

The likelihood function L associated with the observations sequence X is

$$L(p|X) := \prod_{i=1}^{N} p^{X_i}(1-p)^{1-X_i}$$

and the log likelihood function $\mathcal{L}$ is

$$\mathcal{L}(p|X) := \ln L(p|X) = \sum_{i=1}^{N} X_i \ln p + \left(N - \sum_{i=1}^{N} X_i\right) \ln(1-p)$$

This function is maximized for $\hat{p} := \frac{\sum_{i=1}^{N} X_i}{N}$ so the maximum likelihood estimate of the probability $p = P(C_{t,H}|S_t)$, is in fact the historical proportion of correct predictions.

We apply a likelihood ratio test to test the null hypothesis $H_0 : p = p_0$ against the alternative hypothesis $H_A : p \neq p_0$. The null hypothesis reflects the idea that the probability of a random, uninformed signal correctly predicting crashes is p_0. A significant departure above this level indicates that the measure we are considering contains some information about future equity market corrections. The likelihood ratio test is:

$$\Lambda = \frac{L(p = p_0|X)}{\max_{p\in(0,1)} L(p|X)} = \frac{L(p = p_0|X)}{L(p = \hat{p}|X)}. \tag{2}$$

The statistic $Y := -2\ln\Lambda$ is asymptotically χ^2-distributed with $\nu = 1$ degree of freedom. We reject the null hypothesis $H_0 : p = p_0$ and accept that the model has some predictive power if $Y > c$, where c is the critical value chosen for the test.

We perform the test for the three critical values 2.71, 3.84, and 6.63 corresponding respectively to a 90%, 95% and 99% confidence level.

The probability p_0 is the probability to identify an equity market downturn within 252 days of a randomly selected period. To compute p_0 empirically, we tally the number of days that are at most 252 days before a crash identification date and divide by the total number of days in the sample.

D. Monte Carlo Study for Small Sample Bias

A limitation of this likelihood ratio test is that the χ^2 distribution is only valid asymptotically. In our case, the number of correct predictions follows a binomial distribution with an estimated probability of success $\hat{p}$ and N trials. However, "only" 18 crashes occurred during the period considered in this study: the continuous χ^2 distribution might not provide an adequate approximation for this discrete distribution. This difficulty is an example of small sample bias. We use Monte Carlo methods to obtain the empirical distribution of test statistics and address this bias.

The Monte Carlo algorithm is as follows. Generate $K = 10,000$ paths. For each path $k = 1, \ldots, K$, simulate N Bernoulli random variables with probability p_0 of obtaining a "success."

Denote by $X_k := \left\{X_i^k, i = 1, \ldots, N\right\}$ the realization sequence where $x_i^k = 1$ if the ith Bernoulli variable produces a "success" and 0 otherwise.

Next, compute the maximum likelihood estimate for the probability of success given the realization sequence X_k as $\hat{p} := \frac{\sum_{i=1}^N X_i^k}{N}$, and the test statistic for the path as

$$Y_k = -2 \ln \Lambda_k = -2 \ln \frac{L(p = p_0 | X_k)}{\max_{p \in (0,1)} L(p_k | X_k)} = -2 \ln \frac{L(p = p_0 | X_k)}{L(p = \hat{p}_k | X_k)}.$$

Once all the paths have been simulated, we use all K test statistics $Y_k, k = 1, \ldots, K$ to produce an empirical distributions for the test statistic Y.

From the empirical distribution, we obtain critical values at a 90%, 95% and 99% confidence level, against which we assess the crash prediction test statistic Y. The empirical distribution also enables us to compute a p-value for the crash prediction test statistics. Finally, we compare the results obtained with the empirical distribution to those derived using the asymptotic χ^2 distribution.

We are ready to implement this approach on a first predictive measure: the price-to-earnings ratio.

IV. The Price-to-Earnings Ratio

Practitioners have used the price-to-earnings (P/E) ratio to gauge the relative valuation of stocks and stock markets since at least the 1930s (for example, Graham and Dodd, 1934, discuss the use of the P/E ratio in securities analysis and valuation). In this section, we analyze the predictive ability of the P/E ratio calculated using current earnings. The advantage of this definition for the SHCOMP is that it is available over the entire period.

Table 5 shows that the P/E and logarithm of the P/E generated a total of 18 signals (based on normally distributed confidence intervals) and 19 signals (based on Cantelli's inequality). The number of correct predictions across models ranges from 15 to 18 and the accuracy of the models ranges from 83.33% for a signal based on the P/E ratio, to 94.74% for a signal computed using the logarithm of the P/E ratio. The type of confidence interval - normal distribution or Cantelli's inequality - only have a minor influence on the end result.

Next, we test the accuracy of the prediction statistically. Before applying the likelihood ratio test, we need to compute the uninformed prior probability p_0 that a day picked at random will precede a crash identification date by 252 days or less. We find that this probability is very high, at $p_0 = 71.17\%$. This finding is consistent with the stylized facts discussed in Section 2. The Likelihood ratio test then shows that the logarithm of the P/E ratio is a significant predictor of bear markets at (or near) a 99% confidence level. Although the accuracy of the P/E ratio is markedly higher than p_0, we cannot conclude that this measure is in itself significant.

[**Place Table 5 here**]

We continue our analysis with a Monte Carlo test for small sample bias, presented in Table 6. We compute the critical values at a 90%, 95% and 99% confidence level for the empirical distribution. Because we only have a limited number of signals, the distribution is lumpy, making it difficult to obtain meaningful p-values. Still, we find that the Monte Carlo analysis confirms our earlier conclusions about significance of the logarithm of the P/E ratio.

[**Place Table 6 here**]

V. The Cyclically-Adjusted Price-to-Earnings Ratio and the Bond-Stocks Earnings Yield Differential Model

The drawback of the P/E ratio calculated using current earnings is that it might be might be overly sensitive to current economic and market conditions. Graham and Dodd (1934) warned against this risk and advocated the use of a P/E ratio based on average earnings over ten years. In their landmark survey, Campbell and Shiller (1988) found that the R^2 of a regression of log returns on the S&P 500 over a 10-year period against the log of the price-earnings ratio computed using average earnings over the previous 10 and 30 years is significant (see Lleo and Ziemba, 2015b, for a review of the literature and a discussion of the key results.). This led Shiller to suggest the use of a Cyclically Adjusted Price-to-Earnings ratio (CAPE), or a price-to-earnings ratio using 10-year average earnings, to forecast the evolution of the equity risk premium (see Shiller, 2015).

The BSEYD, the second model we test, relates the yield on stocks (measured by the earnings yield, which is also the inverse of the P/E ratio) to that on nominal Government bonds.

$$BSEYD(t) = r(t) - \rho(t) = r(t) - \frac{E(t)}{P(t)}, \tag{3}$$

where $\rho(t)$ is the earnings yield at time t and $r(t)$ is the current 10-year government bond yield $r(t)$. The BSEYD was initially developed for the Japanese market shortly before the crash of 1990 (Ziemba and Schwartz, 1991), and it has since been used successfully on a number of international markets (see the review article Lleo and Ziemba, 2015c), including the 2007-2008 SHCOMP meltdown (Lleo and Ziemba, 2012).

In this section, we test the forecasting ability of four measures:

1. **PE0**: P/E ratio based on current earnings. This is the measure we tested in Section IV;
2. **CAPE10**: CAPE, which is a P/E ratio computed using average earnings over the previous 10-years;
3. **BSEYD0**: BSEYD based on current earnings;
4. **BSEYD10**: BSEYD using average earnings over the previous 10-years.

We also test the logarithm of these measures: **logPE0**, **logCAPE10**, **logBSEYD0** and **logBSEYD10**.

Because the CAPE10 and BSEYD10 require 10 years of earnings data, and the Bloomberg data series for 10-year government bonds only starts on October 31, 2006, we cannot use the full range of stock market data. The analysis in this section covers the period between October 31, 2006 and September 30, 2015. Over this period, the SHCOMP experienced six declines of more than 20% of its value.

Table 7 displays the results for the eight measures, calculated with a confidence interval based on a normal distribution. The results for a confidence interval based on Cantelli's inequality are identical and have been omitted. The measures in our study generated between three signals for CAPE10 and logCAPE10 and six signals for BSEYD0. The accuracy of the measures reaches a low of 50% for logBSEYD0 and a high of 100% for CAPE10 and logCAPE10. Only four of the eight measures are 75% accurate or better.

By comparison, the uninformed prior probability that a day picked at random will precede a crash identification date by 252 days or less is $p_0 = 70.99\%$. Because of the relatively short period and small number of downturns, only CAPE10 and logCAPE10 appear significant. However, these two models only predicted three of the six crashes.

Overall, the BSEYD-based models do not perform as well as the P/E-based models. This is surprising, because the BSEYD model contains additional information that is not in the P/E, namely government bond yields. In addition, the BSEYD and logBSEYD models have been shown to perform better than the P/E ratio and CAPE on the American market (Lleo and Ziemba, 2015b). One possible explanation for this paradox is that the BSEYD-based measures tend to produce a signal earlier than the P/E ratio. For example, if we double the lead-time of the measures from 252 days to 504 days, the accuracy of the logBSEYD10 model improves to 100% with all 6 crashes predicted. However, the same adjustment does not produce a similarly spectacular improvement for the BSEYD0 and logBSEYD0 models.

[**Place Table 7 here**]

The Monte Carlo study for small sample bias, presented in Table 8 confirms these findings.

[**Place Table 8 here**]

VI. Conclusion

The Chinese stock market is certainly one of the most interesting equity markets in the world. Its size, scope, structure and recency make it unique. These characteristics inevitably affect its behavior and returns. Although the Shanghai Stock Exchange is currently world's fifth largest stock exchange by market capitalization, its behavior is much more volatile and extreme than that of more mature equity markets in Europe, North America or East Asia.

Our investigation of traditional crash predictors reveals that the logarithm of the P/E calculated using current earnings has successfully predicted crashes over the entire length of the study. We also found that the P/E ratio outperformed an uninformed benchmark, but that this outcome was not statistically significant.

We found mixed evidence for the BSEYD models and CAPE over a shorter 9-year period. Although five of the eight models tested were 75% accurate or better, the historical sample was too short to confirm whether this advantage was statistically significant. We also found that the BSEYD measures performed less well than expected. The BSEYD provides an accurate signal typically more than a year before a crash is identified, outside of the range use for our test.

Overall, this study shows clearly that crash prediction models can be applied directly to the Chinese market, and reveals potential areas for further research both on the behavior of Chinese equity markets and on crash prediction models.

References

Baum, L, T Petrie, G Soules, and N Weiss, 1970, A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains, *The Annals of Mathematical Statistics* 41, 164–171.

Campbell, J. Y., and R. J. Shiller, 1988, Stock prices, earnings, and expected dividends, *The Journal of Finance* 43, 661–676, Papers and Proceedings of the Forty-Seventh Annual Meeting of the American Finance Association, Chicago, Illinois, December 28-30, 1987.

Cont, R., 2001, Empirical properties of asset returns: stylized facts and statistical issues, *Quantitative Finance* 1, 223–236.

Dempster, A.P., N.M. Laird, and D.B. Rubin, 1977, Maximum-likelihood from incomplete data via the em algorithm, *Journal of the Royal Statistical Society, Series B* 39, 1–38.

Graham, B., and D. L. Dodd, 1934, *Security Analysis*, first edition (Whittlesey House).

Grimmett, G., and D. Stirzaker, 2001, *Probability and Random Processes* (Oxford University Press).

Lleo, S., and W. T. Ziemba, 2012, Stock market crashes in 2007–2009: were we able to predict them?, *Quantitative Finance* 12, 1161–1187.

Lleo, S., and W.T. Ziemba, 2015a, Can warren buffett also predict equity market downturns?, Technical report.

Lleo, S., and W.T. Ziemba, 2015b, Does the bond-stock earnings yield differential model predict equity market corrections better than high p/e models?, Technical report.

Lleo, S., and W.T. Ziemba, 2015c, Some historical perspectives on the bond-stock yield model for crash prediction around the world, *International Journal of Forecasting* 31, 399–425.

Rabiner, L., 1989, A tutorial in Hidden Markov Models and selected applications in speech recognition, *Proceedings of the IEEE* 77, 257–286.

Rabiner, Lawrence, and Biing-Hwang Juang, 1993, *Fundamentals of Speech Recognition*, Prentice Hall Signal Processing Series (Prentice Hall).

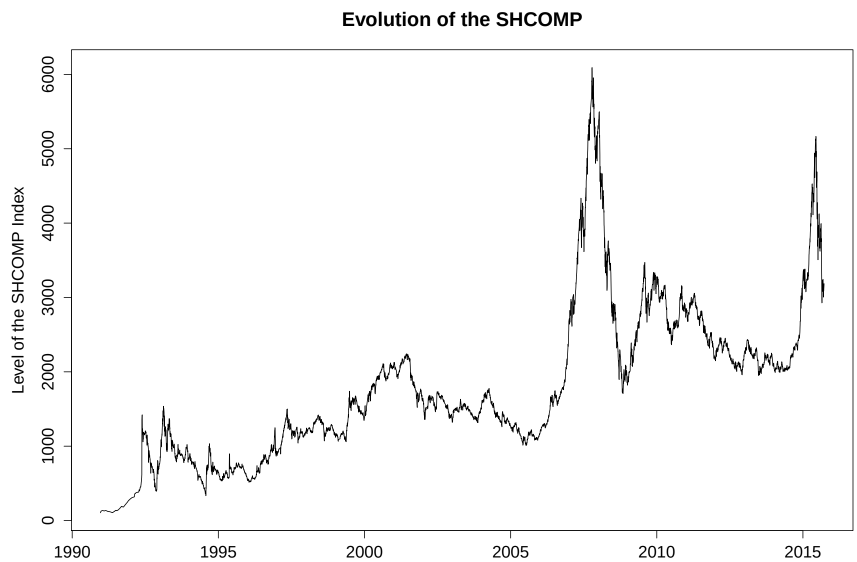
Shiller, R. J., 2015, *Irrational Exuberance*, third edition (Princeton University Press).

Shiryaev, A. N., M. Zitlukhin, and W. T. Ziemba, 2014, When to sell Apple and the Nasdaq 100? trading bubbles with a stochastic disorder model, *Journal of Portfolio Management* 40, 1–10.

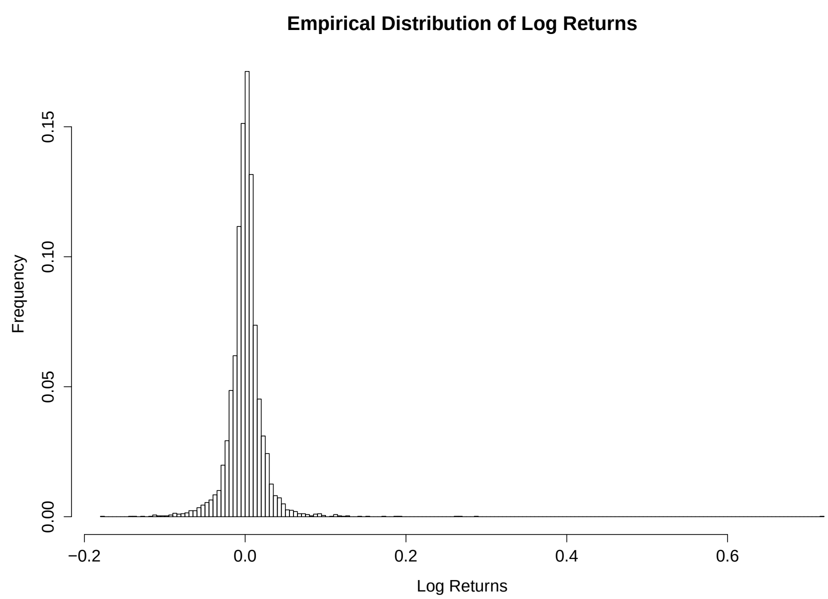
Shiryaev, A. N., M. Zitlukhin, and W. T. Ziemba, 2015, Land and stock bubbles, crashes and exit strategies in japan, circa 1990 and in 2013, *Quantitative Finance (forthcoming)* .

Vitterbi, Andrew, 1967, Error bounds for convolutional codes and an asymptotically optimum decoding algorithm, *IEEE Transactions on Information Theory* 13, 260–269.

Ziemba, W.T., and S.L. Schwartz, 1991, *Invest Japan* (Probus, Chicago).



(a)

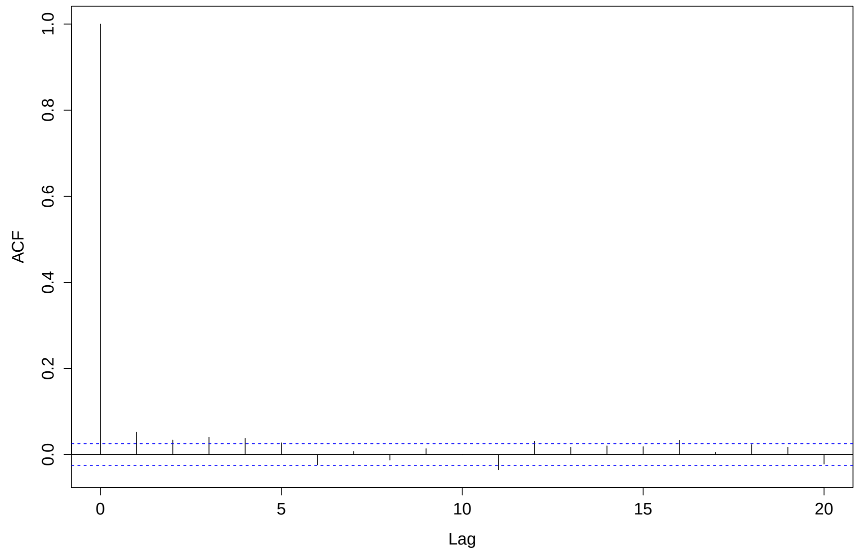


(b)

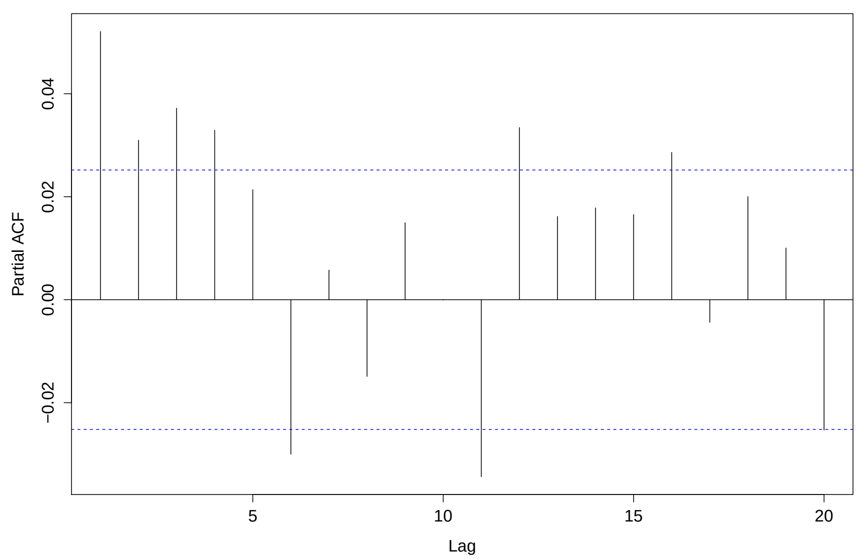
Figure 1: Evolution of the SHCOMP Index and empirical distribution of the daily log return (December 19, 1990 - September 30, 2015).

Descriptive Statistics	Frequency		
	Daily	Weekly	Monthly
Number of observations	6,055	1,286	299
Mean	0.0572%	0.2714%	1.1152%
Median	0.0682%	0.0741%	0.7164%
Minimum	-17.9051%	-22.6293%	-37.3283%
Maximum	71.9152%	90.0825%	101.9664%
Standard deviation	2.3999%	5.6249%	12.9767%
Variance	0.0006	0.0032	0.0168
Skewness	5.2594	5.3774	2.3729
Kurtosis	149.1905	78.3044	20.7136
Jarque-Bera statistics	5,419,808	310,056	4,190
(p-value)	(< 2.2e-16)	(< 2.2e-16)	(< 2.2e-16)

Table 1: Descriptive statistics for daily, weekly and monthly log returns on the SHCOMP



(a) Sample autocorrelation up to lag 20



(b) Sample partial autocorrelation up to lag 20

Figure 2: Autocorrelation and partial autocorrelation of the daily log returns on the SHCOMP

Criterion	1	2	3	4	4	6	7
LogLikelihood	13,992.53	16,015.10	16,329.13	16,383.62	16,600.10	16,677.31	16,685.14
AIC	-27,981.06	-32,016.20	-32,630.26	-32,721.25	-33,132.21	-33,260.61	-33,246.27
BIC	-27,967.64	-31,969.24	-32,536.34	-32,566.95	-32,904.11	-32,945.31	-32,830.34

Table 2: Hidden Markov Model fitting for the daily log returns on the SHCOMP

State	Initial Probability	Mean	Standard Deviation
1	0.00	-0.2909%	4.5990%
2	0.00	0.0580%	0.5096%
3	0.00	0.0973%	1.9328%
4	0.00	-0.0435%	1.2217%
5	1.00	9.4089%	17.1468%
6	0.00	-0.9600%	0.0519%

Table 3: Initial probability and parameters of the Gaussian distributions for each state of the HMM

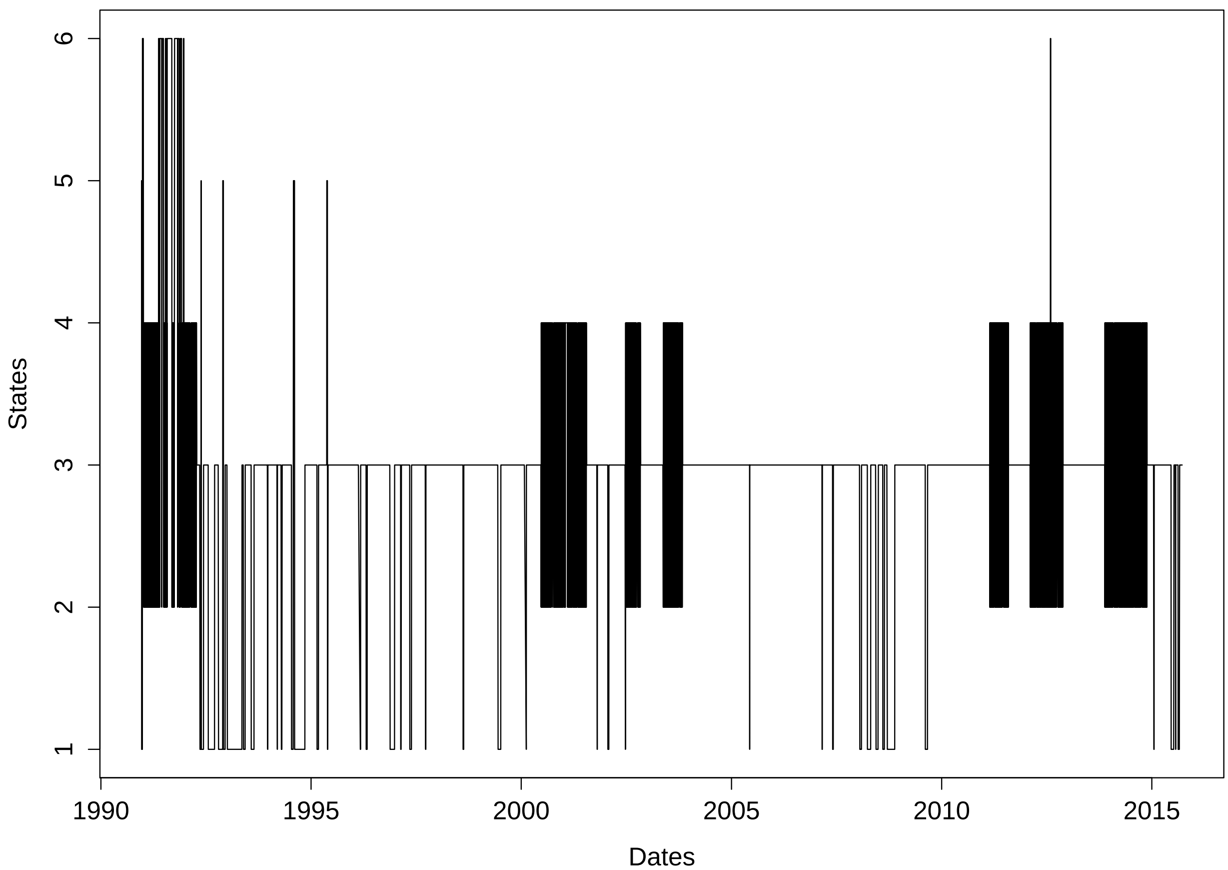


Figure 3: Historical path of the HMM across the six states.

	Crash Identification Date	Peak Date	SHCOMP Index at Peak	SHCOMP Level at 20% Cutoff	Trough date	SHCOMP Level at trough	Peak-to-trough decline (%)	Peak-to-trough duration (in days)
1	1992-06-03	1992-05-25	1421.57	1137.26	1992-11-17	393.52	72.3%	176
2	1993-03-01	1993-02-15	1536.82	1229.46	1993-03-31	925.91	39.8%	44
3	1994-09-30	1994-09-13	1033.47	826.78	1995-02-07	532.49	48.5%	147
4	1995-10-04	1995-05-22	897.42	717.94	1996-01-22	516.46	42.5%	245
5	1996-12-17	1996-12-09	1247.66	998.13	1996-12-24	865.58	30.6%	15
6	1997-07-02	1997-05-12	1500.40	1200.32	1997-09-23	1041.97	30.6%	134
7	1998-08-17	1998-06-03	1420 .00	1136.00	1998-08-17	1070.41	24.6%	75
8	1999-12-22	1999-06-29	1739.21	1391.36	1999-12-27	1345.35	22.6%	181
9	2001-09-17	2001-06-13	2242.42	1793.94	2002-01-22	1358.69	39.4%	223
10	2003-01-10	2002-07-08	1732.93	1386.35	2003-03-26	1384.86	20.1%	261
11	2004-06-25	2004-04-06	1777.52	1422.01	2004-09-13	1260.32	29.1%	160
12	2005-08-12	2004-09-23	1464.78	1171.82	2005-12-05	1079.19	26.3%	438
13	2007-11-27	2007-10-16	6092.06	4873.65	2008-11-04	1706.70	72.0%	385
14	2009-08-31	2009-08-04	3471.44	2777.15	2009-08-31	2667.75	23.2%	27
15	2010-08-18	2009-11-23	3338.66	2670.93	2010-09-20	2588.71	22.5%	301
16	2011-08-08	2010-11-08	3159.51	2527.61	2012-01-05	2148.45	32.0%	423
17	2012-12-03	2012-03-02	2460.69	1968.55	2012-12-03	1959.77	20.4%	276
18	2015-06-29	2015-06-12	5166.35	4133.08	2015-08-26	2927.29	43.3%	75

Table 4: The SHCOMP Index experienced 18 crashes between December 19, 1990 and October 31, 2015.

Signal Model	Total number of signals	Number of correct predictions	ML Estimate $\hat{p}$	$L(\hat{p})$	Likelihood ratio Λ	Test statistics $-2\ln\Lambda$	p-value
PE (confidence)	18	15	83.33%	3.0000e-04	0.4854	1.4455	22.93%
PE (Cantelli)	19	16	84.21%	2.5175e-04	0.4854	1.7717	18.32%
logPE (confidence)	19	18	94.74%	1.9888e-02	0.0318	6.8962^{**}	0.86%
logPE (Cantelli)	18	17	94.44%	2.1025e-02	0.0423	6.3271^{*}	1.19%

* significant at the 5% level;
** significant at the 1% level;
*** significant at the 0.5% level.

Table 5: Maximum likelihood estimate and likelihood ratio test for the PE and logPE

Signal Model	Total number of signals	ML Estimate $\hat{p}$	Critical Value: 90% confidence	Critical Value: 95% confidence	Critical Value: 99% confidence	Test statistics $-2\ln\Lambda(p_0)$
PE (confidence)	18	83.33%	3.30	3.56	6.33	1.4455
PE (Cantelli)	19	84.21%	3.30	3.56	6.33	1.7717
logPE (confidence)	19	94.74%	2.90	3.75	6.90	6.8962^{*}
logPE (Cantelli)	18	94.44%	2.90	3.75	6.90	6.3271^{**}

* significant at the 5% level;
** significant at the 1% level;
*** significant at the 0.5% level.

Table 6: Monte Carlo likelihood ratio test for the PE and logPE

Signal Model	Total number of signals	Number of correct predictions	ML Estimate $\hat{p}$	$L(\hat{p})$	Likelihood ratio Λ	Test statistics $-2\ln\Lambda$	p-value
BSEYD0	4	3	75.00%	1.0547e-01	0.9840	0.0322	85.76%
logBSEYD0	6	3	50.00%	1.5625e-02	0.5590	1.1631	28.08%
PE0	4	3	75.00%	1.0547e-01	0.9840	0.0322	85.76%
logPE0	4	3	75.00%	1.0547e-01	0.9840	0.0322	85.76%
BSEYD10	3	2	66.67%	1.4815e-01	0.9868	0.0265	87.07%
logBSEYD10	5	3	60.00%	3.4560e-02	0.8712	0.2758	59.95%
CAPE10	3	3	100.00%	∞	-	-	-
logCAPE10	3	3	100.00%	∞	-	-	-

* significant at the 5% level;
** significant at the 1% level;
*** significant at the 0.5% level.

Table 7: Maximum likelihood estimate and likelihood ratio test for the BSEYD0, PE0, BSEYD10 and CAPE10 and their logarithm

Signal Model	Total number of signals	ML Estimate $\hat{p}$	Critical Value: 90% confidence	Critical Value: 95% confidence	Critical Value: 99% confidence	Test statistics $-2\ln\Lambda(p_0)$
BSEYD0	4	75.00%	2.74	3.61	3.61	0.0322
logBSEYD0	6	50.00%	4.11	4.11	4.11	1.1631
PE0	4	75.00%	2.74	3.61	3.61	0.0322
logPEE	4	75.00%	2.74	3.61	3.61	0.0322
BSEYD10	3	66.67%	2.06	2.06	7.43	0.0265
logBSEYD10	5	60.00%	3.43	3.43	5.58	0.2758
CAPE10	3	100.00%	2.06	2.06	7.43	∞
logCAPE10	3	100.00%	2.06	2.06	7.43	∞

* significant at the 5% level;
** significant at the 1% level;
*** significant at the 0.5% level.

Table 8: Monte Carlo likelihood ratio test for the BSEYD0, PE0, BSEYD10 and CAPE10 and their logarithm