Common Risk Factors in China's A Shares

Anthony Pan

August 12, 2015

Abstract

In this report, I investigate the presence of risk factors in China's A Shares. I give a brief overview of previous literature on portfolio risk, and then conduct my own factor analysis of size, value, price/earnings, yield, and inclusion based on returns from the MSCI China A Index from January 2008 to July 2015.

1 Introduction

Harry Markowitz is often called the "father of modern portfolio theory" for his 1952 research paper and 1959 book which established a mathematically based approach for portfolio selection and managing risk. He assumed a normal distribution for portfolio returns, and set up a framework to maximize expected returns for a given amount of risk, or minimize risk for a given amount of expected returns based on how each security co-moves with all other securities. Moving forward, further theoretical research has been conducted studying the expansion the portfolio optimizing problem over more than one period[1] as well as a continuous time period,[2] separation theorems which study the implications of the optimum portfolio of risky assets being independent of the investors taste for expected returns or variance,[3] and how portfolio rebalancing is affected by current holdings and transaction costs.

The development of mean variance portfolio theory set the foundation for predicting asset returns. Estimates of correlation coefficients and covariances were applied at first through a single index model, called the Capital Asset Pricing Model (CAPM). The model was created by Treynor (1961, 1962), Sharpe (1964), Lintner (1965), and Mossin (1966) and was popularized by Sharpe (1967). The CAPM predicts the expected return of a security based on beta (β), the volatility of the security in relation to the market.

Fama and French (1993) expanded the one-factor CAPM to a three-factor model, finding that size (market capitalization) and value (ratio of book equity to market equity) of an equity were also correlated to its returns in the United States: the three factor model accounted for over 90% of equity returns in sample, as opposed to CAPM's 70%. Much research has been done on the identification of other possible risk factors, including momentum, volatility, quality, and dividend yield, and on expanding these results to global markets.

In this report I investigate risk factors in the China A Share market, which is gaining more relevance internationally as China's government pushes toward greater market liberalization and internationalization. A Shares are mainland China-based companies that trade on Chinese stock exchanges with heavy restrictions for foreign investors. The China A market is unique in that 90% of the market is owned by retail investors, and government regulated quotas and lock-up periods have greatly affected volatility and foreign exposure. In the past, international investors were required to apply for quotas to purchase China A Shares as a "Qualified Foreign Institutional Investor" (QFII); now, the Shanghai-Hong Kong Stock Connect established in 2014 allows investors to trade shares between markets through local brokers and clearing houses. Furthermore, MSCI is strongly considering adding China A Shares to its Emerging Markets Index in the near future, which would be an opportunity for long-term and international investors to trade in the mainland Chinese market.

[1] See Fama (1970), Hakansson (1970, 1974), Mossin (1969)
[2] See Merton (1969)
[3] See Ross (1978)

I examine the presence of risk factors for asset returns in Chinese A Share market from January 2001 to July 2015 through the MSCI China A Index, which captures large and mid-cap representation among over 500 securities listed on the Shanghai and Shenzhen Stock Exchanges.

2 Literature Review

The introduction of the Fama and French (1993) three factor model significantly improved the CAPM explanation of empirically observed excess returns of equities by using the risk factors of size, and value in addition to β. Beta is defined as the volatility of the returns of the company or portfolio in relation to the returns of the stock market, size is measured as market capitalization, and value is measured by the ratio of book equity to market equity. In their paper, Fama and French find that returns of companies in the United States from 1962 to 1989 are negatively correlated with size and positively correlated to book equity to market equity in addition to being correlated with CAPM's β.

In recent years, the rapidly growing Asian stock markets have attracted much research interest, especially in the area of identifying risk factors that correspond to asset returns. Chui and Wei (1998) verify the validity of the Fama and French three factor model internationally in the stock markets of Hong Kong, South Korea, Malaysia, Taiwan, and Thailand from 1977 to 1993. They find that the relationship between average stock return and market beta is weak in all countries, and that book to market equity is significant in three countries and size is significant in four; this suggests that these risk factors also exist in emerging Asian markets.

China's stock market, however, operates differently than other countries. Shares of Chinese companies are split into A-shares, B-shares, H-shares, Red chips, P-chips, and N-shares. B shares are shares of mainland companies quoted in USD or HKD and open to foreign ownership; H shares are mainland companies traded in Hong Kong; Red chips are state owned companies incorporated outside the mainland and traded in Hong Kong; P-chips are non-state owned companies incorporated outside the mainland and traded in Hong Kong; N-shares are companies incorporated outside the mainland and listed in the U.S. on the NYSE or NASDAQ. Domestic investors, the vast majority of whom are retail investors, are most involved with the A-share market touched upon earlier in the introduction.

One of the earliest studies of market risk factors in China's A Share market was conducted by Chen and Qu (2000), whom tested the CAPM in the Shanghai Stock Exchange and found that β played a significant but unstable role in measuring market risk. However, literature in the following years disagree on the extent to which market factors are significant. Drew, Naughton, and Veeraraghavan (2002) expanded on Chen and Qu by testing the Fama and French three-factor model in Shanghai A Shares from 1993 to 2001, and found that the three factors were significant, but growth stocks (low book to market ratio) outperformed value stocks (high book to market ratio), contrary to Fama and French. In further investigations of the three factors, Jia and Chen (2003), who examined the price-earnings ratio, and Su and Mai (2004), who controlled for liquidity, both found significance in all factors and that small and value stocks consistently earn higher returns than large firms and growth stocks. In 2004, Wang and Xu also examined A shares from 1996-2002, but found that the book to market ratio was not significant and suggested using the floating ratio as a predictive factor instead.

More recently, studies have begun to agree on the validity of the three-factor model as well as explore other possible risk factors in China. Chen, Leong, and Anderson (2007), Narayan and Zheng (2010), and Gan, Hu, Liu, and Li (2013) all agree that returns on Chinese A shares are negatively correlated with size and positively correlated to book to market ratio. Guo (2013) finds that the three-factor Fama French model captures less variation in returns than the U.S., suggesting that government control, momentum, and "herd mentality" may provide additional information, and Li et al. (2010) finds momentum-based strategies unprofitable in Chinese stock returns from 1994-2007. Chueng, Hoguet, and Ng (2015) conducts the most recent study of risk factors in China's A share market, testing for the effect of size, value, yield, volatility, and momentum from 2002 through 2013 in the MSCI China A Index. They find that only the value and dividend yield factors are significant at the 5% level, while finding positive but insignificant premiums for the other factors.

3 Data and Methodology

I use the procedure of Fama and French (1993, 2012) to investigate the cross section of average returns of the Chinese A Share market from January 2008 to June 2015, similar to Cheung, Hoguet, and Ng (2015). The specific market that I analyze is the MSCI China A Share Index, which consists of mid-cap and large-cap China A securities listed on the Shanghai and Shenzhen exchanges, and captures 65% of the free-float adjusted market capitalization in each GICS industry group on each exchange and the largest 25 securities (Cheung, et. al, 2015). MSCI China A is an extremely relevant Chinese index with much investor interest. Historical membership and company statistics are obtained from MSCI, and index returns and the risk free rate of return are obtained from Bloomberg. Data manipulation and analysis were completed in R.

The period I investigate consists of 90 months, and within that time period the securities included in the MSCI China A index vary considerably due to the quickly developing market. I exclude companies during the last month that they are included in the index, companies that have halted trading, and companies experiencing other extenuating circumstances that prevent the calculation of essential risk factors in order to have a consistent and complete sample. For my analysis, I examine 40,493 data points over the course of 90 months, testing size, value, price/earnings, yield, and inclusion as risk factors.

3.1 Factor Definitions

The factors I use are size, value, price/earnings, yield, inclusion, and the market factor from January 2008 to June 2015. Size is measured as a company's market capitalization at month t, defined as closing price multiplied by the closing number of shares at the end of the month. Value, or book equity to market equity, is measured as the book value of the company as calculated by MSCI divided by the closing price of the stock at the end of month t. Price/earnings is a company's closing price divided by MSCI calculated earnings; I change all negative P/E ratios to zero. Yield is monthly dividend yield, and the inclusion is MSCI's domestic inclusion factor at the end of month t. MSCI defines its domestic inclusion factor as the free float, or proportion of shares outstanding available for purchase in the public equity markets, that is available to domestic investors. The market factor is the monthly excess earnings of the index from month t to $t+1$ minus the risk free rate, which I define as the monthly returns of a one year Chinese government bond purchased at the end of month t.

By investigating these factors, I look for whether or not a company's size, value, price/earnings ratio, yield, inclusion factor, and future market returns at the end of month t can be used to explain its returns from t to $t+1$. I define the returns to be predicted in month t as the stock's price in month $t+1$ divided by the price in month t multiplied by 100.

To solve for factor premia, I used the methodology of Fama and French. I split the MSCI China A data into months, and for each month, the stocks are broken into two groups, small and big (S and B) based on the median size. Independently, the stocks are split into three groups for the other factors–value, price/earnings, yield, and inclusion–based on breakpoints for the bottom 30% (Low), middle 40% (Medium), and top 30% (High) of the values for the factor. The decision to sort the firms into three groups based on book to market equity is based on evidence in Fama and French (1992a) that book to market equity has a stronger role in average stock returns than size. However, sorting the other non-size factors into three groups is arbitrary, but consistent with previous studies of additional risk factors. The market factor premium is defined each month as the returns on the MSCI China A index from the current month to the next month.

Table 1: Portfolio sorts on size (Small/Big) against other factors (Low/Medium/High)

	Low (bottom 30%)	Medium (middle 40%)	High (top 30%)
Small (bottom 50%)	Small/Low (S/L)	Small/Medium (S/M)	Small/High (S/H)
Big (top 50%)	Big/Low (B/L)	Big/Medium (B/M)	Big/High (B/H)

Each factor, which has been sorted into three groups, is then intersected with the two sorts on size, forming factor tables with six levels each, as shown in the Table 1 above. This is done for size∗value, size∗price/earnings, size∗yield, and size∗inclusion. The value-weighted monthly returns of each of these portfolios are then used to solve for the effect of each risk factor on returns. Value-weighted returns are used to reduce variance, which is negatively related to size, and to mimic realistic investment opportunities.

The portfolio SMB, used to represent the risk factor in returns related to size, is calculated for each month in the time period by the difference between the simple average of the returns of the three small-stock portfolios and the simple average of the returns on the three big stock portfolios for value. In other words, SMB is the average value-weighted returns of all of the small portfolios (S/L, S/M, S/H) minus the average of value-weighted returns of the large portfolios (B/L, B/M, B/H). The other factor premia are calculated by taking the average of value-weighted returns for the high portfolios (S/H, B/H) minus the average of value-weighted returns for the low portfolios (S/L, B/L). In addition, I look at the factor premia for value, price/earnings, yield, and inclusion separated by size to observe the effect of size on the other factors.

I use the average of these factor returns over all of the months to conduct a one-sample t-test as a preliminary test to determine if the factor is significant in explaining cross sectional returns in the A share market.

3.2 Explaining Returns

I use the Fama French (1993) approach for investigating factor premia in the market by running the factors through a time-series regression to explain value weighted returns from sixteen independently double-sorted portfolios on size and another factor, consistent with much of the already published literature investigating factor returns.

The double sorted portfolios whose returns are used as returns to be explained are created similarly to the sub-portfolios used to calculate factor premia. The stocks in the MSCI A index are split into four equal quartiles independently by size, value, price/earnings, yield, and inclusion. The double sorted portfolios are then created by taking the intersection of each portfolio sorted on size with each of the other factors, creating four sets of sixteen double sorted portfolios, one set for each of the factors: size∗value, size∗price/earnings, size∗yield, and size∗inclusion. Quartiles are used because of the smaller number of stocks listed on MSCI China A compared to the stocks examined by Fama and French (1993). The sub-portfolios are rebalanced every month to reflect the most updated company statistics at a specific point in time.

To measure the significance of each of the risk factors, I test to see if each of the factors is significant in explaining returns from their respective sub-portfolio sorted on size. From the models shown below, we can examine how well the factors capture variability of stock returns on the A-share market.

$$R_p - R_f = \alpha_p + \beta_p(R_m - R_f) \tag{1}$$
$$R_p - R_f = \alpha_p + \beta_p(R_m - R_f) + s_p(SMB) + v_p(HML) \tag{2}$$
$$R_p - R_f = \alpha_p + \beta_p(R_m - R_f) + s_p(SMB) + p_p(P/E) \tag{3}$$
$$R_p - R_f = \alpha_p + \beta_p(R_m - R_f) + s_p(SMB) + y_p(Yield) \tag{4}$$
$$R_p - R_f = \alpha_p + \beta_p(R_m - R_f) + s_p(SMB) + n_p(Inclusion) \tag{5}$$

R_p is the excess value-weighted return of the 16 double sorted portfolios; R_f is the risk free rate, which I define as the monthly returns for a one year government bond; $R_m - R_f, SMB, HML, P/E, Yield, Inclusion$ are the monthly factor returns of the market premium, size, value, price/earnings, yield, and inclusion factors; s_p, v_p, p_p, y_p, n_p are the factor coefficients for each of the factors, and α_p is the error, or returns unexplained by the model. Equation (1) is the classic CAPM model, and equation (2) is the Fama and French three-factor model. I regress equation (1) and each of equations (2) through (5) onto its respective double sorted portfolio in order to test the explanatory power of each of the factors.

4 Analysis

I test for the existence of factor premiums through descriptive statistics, average monthly factor returns, and regression model fit. The descriptive statistics of the double-sorted portfolios are an intuitive way to observe the effect of different factors on returns, and the characteristics of companies that make up each of these portfolios. Calculating and testing for significance of average monthly factor returns reveal trends in the overall market regarding the returns associated with the risk factors. The explanatory power of each factor for explaining the cross section of asset returns is then tested in the regression models.

4.1 Descriptive Statistics

Tables 2 and 3 below display the average value-weighted monthly returns and the standard deviation for those returns for each portfolio. Portfolios containing smaller firms performed better in all cases and value firms with greater book to market ratios also performed better, supporting Fama and French's conclusion that smaller firms and value firms exhibit greater returns. Returns seem to be slightly positively correlated with yield, and slightly negatively correlated with inclusion and the price/earnings ratio. The portfolios that don't seem to follow the trend are small firms with a high price to earnings ratio which perform better than expected, and big firms with a high inclusion factor, which also perform better than expected but not enough to overcome negative average returns. The portfolios formed on small-cap stocks are very slightly more volatile than those formed on large-cap stocks; aside from size, the standard deviations of all of the portfolios do not seem to exhibit any other clear patterns in other factors.

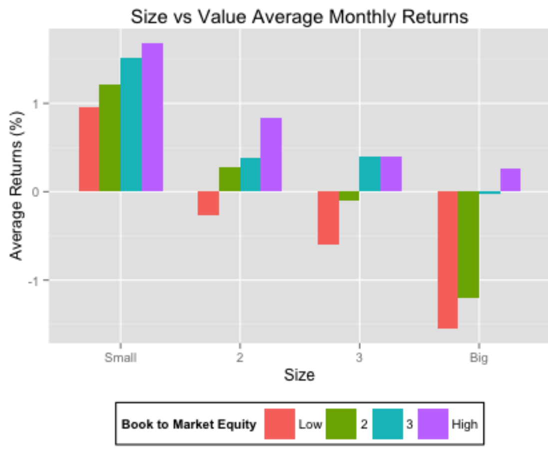
Table 2: Average value-weighted monthly returns (%) for portfolios double sorted by factor

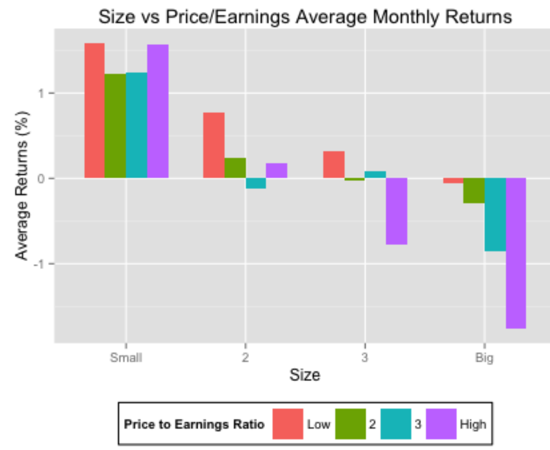
Size Quartile	Non-size Factor Quartile									
	Low	2	3	High	Average	Low	2	3	High	Average
	Size vs Value					*Size vs P/E*				
Small	0.959	1.215	1.518	1.683	1.344	1.587	1.228	1.241	1.568	1.406
2	-0.265	0.270	0.379	0.828	.303	0.774	0.236	-0.124	0.176	0.265
3	-0.591	-0.103	0.398	0.393	.024	0.324	-0.024	0.087	-0.777	-0.097
Big	-1.551	-1.207	-0.028	0.269	-.629	-0.055	-0.299	-0.848	-1.763	-0.741
Average	-.362	.044	.567	.793		0.657	.285	0.089	-0.199	
	Size vs Yield					*Size vs Inclusion*				
Small	1.0642	1.661	1.450	1.880	1.514	1.817	1.765	1.506	1.183	1.568
2	0.405	-0.083	0.283	0.479	0.271	0.470	0.111	0.186	0.056	0.205
3	-1.015	0.019	0.282	0.287	-0.107	-0.013	0.008	-0.240	-0.392	-0.159
Big	-0.693	-1.043	-0.328	-0.334	-0.600	-0.395	-0.527	-1.413	-0.320	-0.664
Average	-0.060	0.139	0.422	0.578		0.470	0.339	0.010	0.132	

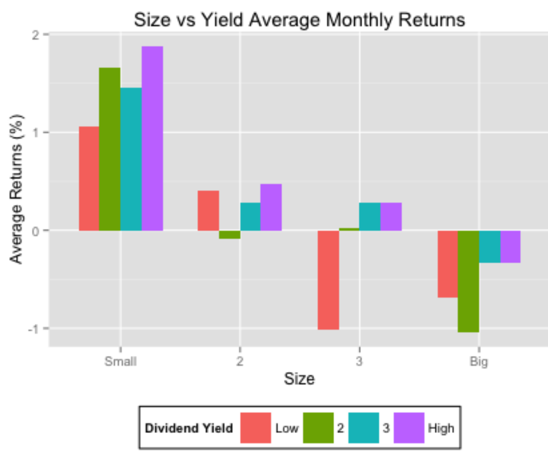
Table 3: Standard deviation of value-weighted monthly returns (%) for doubled sorted portfolios

Size Quartile	Non-size Factor Quartile									
	Low	2	3	High	Average	Low	2	3	High	Average
	Size vs Value					*Size vs P/E*				
Small	10.50	10.52	10.98	10.48	10.62	10.81	10.26	10.05	11.12	10.56
2	9.74	10.49	10.51	10.65	10.35	10.86	10.22	9.43	10.51	10.26
3	10.02	10.65	10.50	10.62	10.45	11.56	10.04	9.36	10.57	10.38
Big	8.73	9.64	9.86	9.15	9.34	9.55	9.10	9.34	10.07	9.51
Average	9.75	10.32	10.46	10.22		10.69	9.90	9.55	10.57	
	Size vs Yield					*Size vs Inclusion*				
Small	11.01	10.68	10.17	10.29	10.54	11.33	10.54	10.77	10.16	10.70
2	10.34	9.92	10.00	10.03	10.08	10.55	9.74	10.53	9.71	10.14
3	11.28	10.33	9.69	9.99	10.32	10.25	10.23	9.94	10.52	10.23
Big	10.07	9.32	10.48	8.34	9.55	8.16	9.65	10.30	10.58	9.67
Average	10.68	10.06	10.09	9.66		10.07	10.04	10.39	10.24	

The graphs below give a quick visual representation of each of the double sorted portfolio returns, with returns for the non-size factor grouped by size on the x-axis. We see again that average monthly returns decrease as size increases and generally, returns are positively correlated with book to market equity and yield and negatively correlated with price/earnings and inclusion.







Tables 4, 5, 6, and 7 give a more detailed description of each of the portfolios formed on size against value, yield, price/earnings, and inclusion. According to the book to market quartiles among different sizes are about the same, but the average size of the largest quartile of firms greatly increases as book to market equity increases. Portfolios that include companies with higher book to market equities or larger sizes make up a larger portion of total market value. A book to market equity of around .150 is considered low, and a book to market equity of around .755 is considered high.

Tables 5 and 6 reveal that the largest quartile of firms tend to give higher dividend yields and have a low price/earnings ratio, while there is no apparent trend in yield or price/earnings within companies that are not in the largest quartile. Monthly yield increases slightly as size increases, and price/earnings decreases slightly as size increases. The greatest percent of market value comes from portfolios formed on large firms with high yield, and large firms with low price/earnings. A highly monthly yield is around 2.9, and a low price/earnings ratio is around 9. From Table 7, we can see that larger firms display a lower inclusion factor–in other words, less of their shares are available for purchase by domestic investors. The large firms with a low inclusion factor make up the greatest percent of market value. An inclusion factor of around .3 is considered low, and a factor of around .73 is in the highest quartile.

Table 4: Descriptive statistics for portfolios formed on size and value: 2008-2015

Size Quartile	Book to Market Equity Quartiles							
	Low	2	3	High	Low	2	3	High
	Average monthly size (billions)				*Average monthly book to market*			
Small	7.075	6.821	6.646	6.383	0.169	0.302	0.443	0.752
2	10.79	10.69	10.72	10.83	0.157	0.295	0.444	0.765
3	18.47	18.29	18.40	18.59	0.153	0.294	0.444	0.759
Big	55.28	85.27	125.7	145.8	0.142	0.297	0.446	0.750
	Average monthly number of firms				*Average monthly percent of market value*			
Small	19.933	29.867	32.578	30.433	1.482	2.272	2.365	2.071
2	32.000	31.333	25.556	23.478	3.436	3.231	2.590	2.287
3	33.667	28.656	25.467	24.333	5.499	4.446	4.053	3.577
Big	27.233	22.578	28.556	34.256	12.184	10.816	15.843	20.134

Table 5: Descriptive statistics for portfolios formed on size and yield: 2008-2015

Size Quartile	Dividend Yield Quartiles							
	Low	2	3	High	Low	2	3	High
	Average monthly size (billions)				*Average monthly yield*			
Small	6.556	6.890	6.711	6.649	0.041	0.556	1.215	2.871
2	10.715	10.686	10.797	10.818	0.067	0.551	1.213	2.856
3	18.076	18.449	18.580	18.592	0.071	0.559	1.225	2.948
Big	50.382	67.270	89.042	166.259	0.086	0.559	1.237	3.076
	Average monthly number of firms				*Average monthly percent of market value*			
Small	40.01	24.64	26.70	21.46	2.959	1.778	1.978	1.474
2	31.63	31.21	26.51	23.01	3.405	3.208	2.717	2.215
3	23.49	31.71	31.24	25.68	3.671	5.093	4.879	3.932
Big	17.79	24.76	27.68	42.40	6.318	12.542	15.272	24.844

Table 6: Descriptive statistics for portfolios formed on size and price/earnings: 2008-2015

Size Quartile	Price/earnings Quartiles							
	Low	2	3	High	Low	2	3	High
	Average monthly size (billions)				*Average monthly price/earnings*			
Small	6.549	6.634	6.818	6.689	6.368	21.574	35.075	194.967
2	10.751	10.828	10.680	10.744	9.168	21.455	35.905	180.856
3	18.707	18.625	18.169	18.334	10.139	21.274	35.800	166.076
Big	151.311	102.042	73.657	49.525	10.210	20.807	35.608	166.350
	Average monthly number of firms				*Average monthly percent of market value*			
Small	25.61	25.86	29.80	31.54	1.847	1.793	2.155	2.394
2	21.93	25.30	31.47	33.67	2.099	2.536	3.215	3.694
3	23.87	28.70	29.90	29.66	3.780	4.531	4.750	4.514
Big	41.40	32.51	20.96	17.76	27.414	15.771	9.404	6.387

Table 7: Descriptive statistics for portfolios formed on size and inclusion: 2008-2015

Size Quartile	Inclusion Quartiles							
	Low	2	3	High	Low	2	3	High
	Average monthly size (billions)				*Average monthly inclusion factor*			
Small	7.377	6.901	6.674	6.410	0.344	0.451	0.560	0.729
2	10.984	10.716	10.666	10.611	0.326	0.449	0.559	0.724
3	18.511	18.850	18.121	17.940	0.300	0.449	0.554	0.724
Big	152.863	55.839	73.943	70.594	0.238	0.445	0.553	0.761
	Average monthly number of firms				*Average monthly percent of market value*			
Small	11.19	24.26	30.34	47.02	0.541	1.374	2.166	4.108
2	29.79	29.48	24.53	28.57	2.004	2.668	2.750	4.122
3	44.40	30.89	21.33	15.50	4.613	5.034	4.089	3.838
Big	55.29	30.82	15.16	11.36	21.321	14.981	11.222	11.452

4.2 Factor Premia

I test the factor effects for market, size, value, price/earnings, yield, and inclusion in this section. The factor effects are recalculated every month, and will be used in the next section as explanatory variables in a time-series regression to explain the returns of the double-sorted portfolios in the previous section. As discussed earlier, the market factor $R_m - R_f$ is the excess returns of the MSCI China A index over the risk free rate of return, and SMB is the difference between returns of small and big firms accounting for value, and the other factors are the difference between the average returns of the top 30% of firms subtracted by the average returns of the bottom 30% of firms, accounting for size.

I also look at factor returns for small and big sized stocks for value, price/earnings, yield, and inclusion; this takes into account the existence of factor premia within smaller and larger companies. For example, while HML is calculated by the average of returns for the S/H and B/H portfolios minus the average returns for S/L and B/L portfolios, HML_{small} is calculated by average returns of the S/H portfolio minus average returns of the S/L portfolio and HML_{big} is calculated by average returns of B/H minus average returns of B/L. This procedure is repeated for the other factors.

Below, Table 8 conducts a one-sample t-test with 89 degrees of freedom for each of the factor returns. Normal Q-Q plots and scatterplots for each of the factors (not pictured) show that over the 90-month period of January 2008 to July 2015, the factor premia are slightly heavy-tailed but relatively normal otherwise, and variance seems to be constant, fulfilling the requirements of the t-test.

Table 8: Significance tests for monthly factor premia

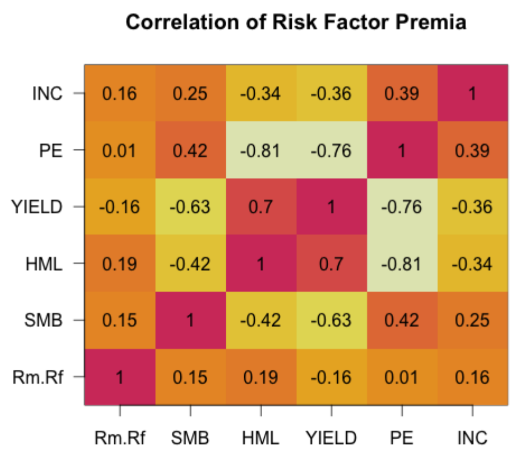
	Mean	Std. Dev.	t-statistic	p-value
$R_m - R_f$	.335	9.05	.35	.727
SMB	1.203	4.54	2.51	.0139*
HML	1.352	5.65	2.27	.025*
HML_{small}	.981	5.79	1.61	.111
HML_{big}	1.722	6.45	2.53	.013*
$Price/Earnings$	-.787	4.56	-1.64	.105
P/E_{small}	-.346	4.26	-.7703	.443
P/E_{big}	-1.23	6.28	-1.86	.067
$Yield$	.312	3.76	.787	.433
$Yield_{small}$	.136	3.43	.376	.708
$Yield_{big}$	.487	5.38	.861	.392
$Inclusion$	-.183	2.42	-.717	.475
$Inclusion_{small}$	-.137	2.63	-.493	.623
$Inclusion_{big}$	-.228	4.06	-.535	.594

HML and SMB are the only factors that display significant excess returns at the .05 level. In other words, this means that value-weighted returns for stocks in MSCI China A that are in the bottom half of stocks sorted by market capitalization are significantly different than value-weighted returns for stocks that are in the upper half, and value-weighted returns for stocks in the top 30% of stocks sorted on book to market equity are significantly different than value-weighted returns for stocks in the bottom 30%.

The market factor is not significant, so we cannot conclude that overall returns in the MSCI China A index are significantly greater than buying risk-free one-year government bonds. The value premium for larger stocks has higher mean returns and a lower p-value than the value premium for small stocks, signifying that the value effect is more evident in large-cap firms. The other factors display similar results: the larger firms exhibit a stronger effect in returns and a higher significance level than smaller firms for price/earnings, yield, as well as inclusion. The difference in returns for the top 30% and bottom 30% of firms sorted on these factors is greater for larger firms than it is for smaller firms, implying that the risk factors make a bigger difference in returns of large stocks than in small stocks.

As observed in the double sorted portfolios, value and yield have a positive premium while price/earnings and inclusion have a negative premium. The inclusion factor seems to have the least significant results as well as the lowest standard deviation, which is intuitive because out of all of the risk factors examined it is the least informative, providing only the percent of shares available to domestic investors with no further financial information. That said, it is interesting that stocks with higher inclusion factors generally provide lower returns than stocks with lower inclusion factors–this may have to do with company or government regulations.

A visualization of the correlation between returns from risk factors is shown in the figure to the right, where red signifies perfectly positive correlation, and white signifies a perfectly negative correlation. The market factor is not strongly correlated with any of the other risk factors. However, value (HML) and yield are moderately correlated with each other, and both are correlated negatively with the price/earnings ratio. The relationship between value and price/earnings intuitively makes sense, as returns for firms whose price is undervalued in the market are expected to be higher, while returns for firms with a greater price/earnings ratio are expected to be lower; thus, they are negatively correlated. The size factor is weakly positively correlated with price/earnings, and weakly negatively correlated with value and yield. Inclusion is also weakly correlated positively with size and price/earnings, and negatively with value and yield.



4.3 Regression Analysis

I run time-series regressions of the factor premia against the returns of the double-sorted portfolios on size and other factors to test if the risk factors can be used to explain the variability of stock returns in China's A-Share market. First, I regress the market factor onto the portfolio returns to test the effectiveness of the basic CAPM model in explaining returns using the returns of the index, shown below in Table 9. I find that the CAPM market beta already explains much of the portfolio returns, accounting for an average of $R^2 \approx$ 80% of the variation.

Table 9: CAPM Adjusted R^2 on sorted portfolio returns: 2008-2015

$R_p - R_m = \alpha_p + \beta_p(R_m - R_f)$								
Size Quartile	Factor Quartiles							
	Low	2	3	High	Low	2	3	High
	Size vs Value ($\mu = .80$)				*Size vs P/E* ($\mu = .81$)			
Small	0.72	0.74	0.81	0.85	0.84	0.77	0.77	0.77
2	0.63	0.76	0.84	0.86	0.79	0.81	0.75	0.75
3	0.73	0.85	0.90	0.91	0.89	0.91	0.85	0.80
Big	0.77	0.87	0.85	0.79	0.83	0.85	0.80	0.83
	Size vs Yield ($\mu = .83$)				*Size vs Inclusion* ($\mu = .81$)			
Small	0.78	0.76	0.80	0.83	0.70	0.72	0.81	0.80
2	0.68	0.80	0.85	0.86	0.83	0.80	0.76	0.74
3	0.85	0.86	0.91	0.93	0.91	0.87	0.86	0.84
Big	0.88	0.88	0.83	0.83	0.87	0.92	0.81	0.75

The results of Table 9 show that the CAPM is more successful in predicting returns for stocks in the MSCI China A index than predicting returns for stocks in the United States. Fama and French (1993) found that the CAPM model accounted for about $R^2 = 75\%$ of returns in U.S. stock returns from 1963-1991, and more studies have been done since regarding the inadequacy of the CAPM beta to fully explain stock returns. The slightly higher R^2 for stock returns in MSCI China A may be because of the smaller sample size that only includes mid to large-cap Chinese companies, unlike Fama and French, who included all stocks trading on the NYSE, NASDAQ, and Amex for a period almost four times as long as the one that I am examining.

I regress each three factor model onto its respective sorted portfolios to see if the other factors are significant in predicting the cross section of returns on MSCI China A shares, and if adding additional factors help explain more of the variation in returns. Results are shown in tables 10-13.

It is evident from all of the regression results that the β and size coefficients are very significant in explaining variability in returns. β displays extremely high t-statistics in all cases, meaning that market returns play a significant part in explaining portfolio returns. The high significance of the market factor may be due to the analysis being completed using realized returns from the past. The t-statistics for size are smaller, but all except one instance are significant at the .05 level, with a vast majority also significant at the .01 level. Size seems to be the most consistently significant factor other than the market factor across the different portfolios. It is interesting to note that in the cases where size is not significant at the .01 level, the other factor is significant in explaining returns–for example, this occurs in the big size/high book to market portfolio where the size factor is relatively less significant compared to the other portfolios, but the value factor is extremely significant.

For the regression of the size and value model onto the size/value sorted portfolios in Table 10, the value factor was significant at the .05 level in 13 out of 16. Thus, book to market equity does seem to be significant in explaining return variability. The portfolios for which the value factor was most significant were the portfolios formed from the smallest and largest quantiles of book to market value. The average R^2 among the size and value three factor model was about .93, a clear improvement from the CAPM's .80. The three factor model explains returns for portfolios based on smaller portfolios marginally better than returns for larger portfolios: R^2 decreases and residual standard error increases slightly as portfolio size increases. Although the model explains much of the data, α, or abnormal returns, is still significantly different than zero at the .05 level for 9 out of the 16 portfolios. This signifies that there may be other unidentified factors in returns.

Table 11 examines the three factor model with size and price/earnings. The price/earnings factor is significant at the .05 level for 14 out of 16 portfolios, so we can conclude that it is important in explaining returns. The size factor seems to be more significant for explaining returns for smaller sized portfolios, and the price/earnings factor seems to be more significant in explaining the negative returns for lower price/earnings ratio portfolios. Average R^2 increases from .81 to .92 when adding the size and price/earnings factor, and the average standard error of residuals is 2.79. 10 out of 16 values of α are significant, which suggests that there may be other factors in explaining returns. Again, the model seems to be slightly better at explaining returns for smaller portfolios, since residual standard error tends to increase slightly for the larger sized portfolios.

The yield factor, as shown in Table 12, is significant at the .05 level for 13/16 of the portfolios, and is more significant for those portfolios in the lowest quartile and highest quartile of dividend yield. Size is also more significant for the smaller sized portfolios. 9 out of 16 values of α are significant, a similar number to the other double sorted portfolios. The R^2 (.92) and residual standard error (2.66) values are also similar to the other portfolios. The model has a slightly higher R^2 and lower standard error for portfolios of smaller sizes, and portfolios of larger yields.

The inclusion factor in Table 13 shows the least significance out of all of the factors, with only 6 out of 16 regression coefficients significant at the .05 level. Abnormal returns (α) are significant in 10 out of 16 portfolios. There doesn't seem to be any patterns in results among the portfolios. However, the three factor model does improve R^2 with an average value of .91, though this may be due in large part to the size factor.

Table 10: Regression results for size and value three factor model onto size/value sorted portfolios

$$R_p - R_m = \alpha_p + \beta_p(R_m - R_f) + s_p(SMB) + v_p(HML)$$

Size Quartile	Book to Market Equity Quartiles											
	Low	2	3	High	Low	2	3	High	Low	2	3	High
	β				s_p				v_p			
Small	0.97	0.95	0.99	0.96	0.80	0.93	1.03	0.82	-0.37	-0.14	0.22	0.35
2	0.86	0.97	1.00	0.97	0.75	0.84	0.69	0.74	-0.54	-0.20	0.11	0.53
3	1.00	1.05	1.07	1.04	0.32	0.60	0.32	0.29	-0.64	-0.11	0.07	0.46
Big	0.93	1.05	1.02	0.85	0.28	-0.44	-0.45	-0.15	-0.54	-0.18	0.18	0.54
	β *t*-statistic				s_p *t*-statistic				v_p *t*-statistic			
Small	29.52*	27.89*	35.76*	32.05*	11.26*	12.62*	17.38*	12.73*	-6.37*	-2.30^{+}	4.59*	6.61*
2	28.95*	28.79*	26.25*	34.44*	11.60*	11.51*	8.45*	12.18*	-10.32*	-3.42^{+}	1.66	10.70*
3	28.47*	31.05*	28.36*	35.60*	4.21*	8.26*	3.89*	4.64*	-10.41*	-1.77^{-}	1.10	8.98*
Big	22.79*	27.52*	29.02*	28.05*	-3.20*	-5.32*	-5.93*	-2.36^{+}	-7.50*	-2.77*	2.89*	10.18*
	α *t*-statistic				*Adjusted* R^2 ($\mu = .93$)				*Residual Std. Err.* ($\mu = 2.71$)			
Small	0.28	-0.35	-1.67^{-}	-0.64	0.94	0.93	0.96	0.95	2.66	2.76	2.24	2.43
2	-2.85*	-2.72*	-2.82*	-4.42*	0.94	0.93	0.91	0.95	2.41	2.73	3.08	2.29
3	-1.59	-3.49*	-1.45	-3.64*	0.92	0.93	0.92	0.95	2.84	2.75	3.05	2.37
Big	-2.27^{+}	-2.40^{+}	-0.45	-2.23^{+}	0.86	0.90	0.92	0.93	3.31	3.08	2.84	2.45

Note: $^{-}$, $^{+}$, and * denote significance at the .10, .05, and .01 levels respectively.

Table 11: Regression results for size and PE three factor model onto size/PE sorted portfolios

$$R_p - R_m = \alpha_p + \beta_p(R_m - R_f) + s_p(SMB) + p_p(Price/Earnings)$$

Size Quartile	Price to Earnings Ratio Quartiles											
	Low	2	3	High	Low	2	3	High	Low	2	3	High
	β				s_p				p_p			
Small	1.03	0.92	0.91	1.00	0.79	1.03	0.86	1.00	-0.27	-0.27	0.14	0.07
2	1.01	0.95	0.85	0.95	0.85	0.88	0.63	0.77	-0.60	-0.23	0.34	0.46
3	1.19	1.03	0.92	1.02	0.32	0.42	0.37	0.36	-0.49	-0.13	0.29	0.59
Big	0.98	0.96	0.96	1.04	-0.16	-0.45	-0.51	-0.32	-0.60	-0.05	0.22	0.67
	β *t*-statistic				s_p *t*-statistic				p_p *t*-statistic			
Small	29.52*	31.87*	33.26*	30.72*	10.30*	16.33*	14.27*	13.92*	-3.55*	-4.31*	2.28^{+}	1.00
2	24.79*	29.89*	24.65*	37.95*	9.59*	12.60*	8.21*	14.06*	-6.86*	-3.37*	4.49*	8.52*
3	31.61*	35.90*	28.78*	29.78*	3.87*	6.74*	5.23*	4.83*	-6.05*	-2.10^{+}	4.15*	8.01*
Big	33.36*	29.23*	22.40*	28.95*	-2.46^{+}	-6.24*	-5.34*	-4.05*	-9.55*	-0.76	2.38^{+}	8.66*
	α *t*-statistic				*Adjusted* R^2 ($\mu = .92$)				*Residual Std. Err.* ($\mu = 2.79$)			
Small	-0.02	-2.23^{+}	-0.28	0.03	0.93	0.94	0.95	0.94	2.95	2.43	2.31	2.75
2	-2.94*	-4.60*	-2.94*	-3.28*	0.90	0.93	0.90	0.96	3.42	2.68	2.92	2.11
3	-2.57^{+}	-3.86*	-1.69	-3.56*	0.93	0.94	0.92	0.93	3.16	2.42	2.70	2.88
Big	-2.67^{+}	-0.64	-1.14	-3.71*	0.93	0.91	0.85	0.91	2.46	2.78	3.63	3.03

Note: $^{-}$, $^{+}$, and * denote significance at the .10, .05, and .01 levels respectively.

Table 12: Regression results for size and PE three factor model onto size/yield sorted portfolios

$$R_p - R_m = \alpha_p + \beta_p(R_m - R_f) + s_p(SMB) + y_p(Yield)$$

Size Quartile	Price to Earnings Ratio Quartiles											
	Low	2	3	High	Low	2	3	High	Low	2	3	High
	β				s_p				y_p			
Small	0.97	0.95	0.99	0.96	0.80	0.93	1.03	0.82	-0.37	-0.14	0.22	0.35
2	0.86	0.97	1.00	0.97	0.75	0.84	0.69	0.74	-0.54	-0.20	0.11	0.53
3	1.00	1.05	1.07	1.04	0.32	0.60	0.32	0.29	-0.64	-0.11	0.07	0.46
Big	0.93	1.05	1.02	0.85	0.28	-0.44	-0.45	-0.15	-0.54	-0.18	0.18	0.54
	β t-statistic				s_p t-statistic				y_p t-statistic			
Small	29.52*	27.89*	35.76*	32.05*	11.26*	12.62*	17.38*	12.73*	-6.37*	-2.30^{+}	4.59*	6.61*
2	28.95*	28.79*	26.25*	34.44*	11.60*	11.51*	8.45*	12.18*	-10.32*	-3.42^{+}	1.66	10.70*
3	28.47*	31.05*	28.36*	35.60*	4.21*	8.26*	3.89*	4.64*	-10.41*	-1.77^{-}	1.10	8.98*
Big	22.79*	27.52*	29.02*	28.05*	-3.20*	-5.32*	-5.93*	-2.36^{+}	-7.50*	-2.77*	2.89*	10.18*
	α t-statistic				*Adjusted* R^2 ($\mu = .92$)				*Residual Std. Err.* ($\mu = 2.66$)			
Small	0.28	-0.35	-1.67^{-}	-0.64	0.94	0.93	0.96	0.95	2.66	2.76	2.24	2.43
2	-2.85*	-2.72*	-2.82*	-4.42*	0.94	0.93	0.91	0.95	2.41	2.73	3.08	2.29
3	-1.59	-3.49*	-1.45	-3.64*	0.92	0.93	0.92	0.95	2.84	2.75	3.05	2.37
Big	-2.27^{+}	-2.40^{+}	-0.45	-2.23^{+}	0.86	0.90	0.92	0.93	3.31	3.08	2.84	2.45

Note: $^{-}$, $^{+}$, and * denote significance at the .10, .05, and .01 levels respectively.

Table 13: Regression results for size and PE three factor model onto size/inclusion sorted portfolios

$$R_p - R_m = \alpha_p + \beta_p(R_m - R_f) + s_p(SMB) + n_p(Inclusion)$$

Size Quartile	Price to Earnings Ratio Quartiles											
	Low	2	3	High	Low	2	3	High	Low	2	3	High
	β				s_p				n_p			
Small	0.99	0.93	1.01	0.94	1.04	0.98	0.86	0.88	-0.35	-0.33	-0.05	0.08
2	1.03	0.91	0.95	0.85	0.79	0.79	0.95	0.83	-0.56	-0.09	-0.06	0.26
3	1.07	1.02	0.98	1.02	0.30	0.48	0.53	0.40	-0.18	-0.01	0.03	0.46
Big	0.89	1.04	1.01	1.02	-0.47	-0.23	-0.15	-0.70	-0.33	0.09	0.70	1.25
	β t-statistic				s_p t-statistic				n_p t-statistic			
Small	19.67*	22.19*	30.88*	38.69*	10.24*	11.43*	12.97*	17.86*	-1.81^{-}	-2.03^{+}	-0.41	0.85
2	34.23*	30.13*	26.51*	22.88*	12.95*	12.81*	13.08*	11.01*	-4.85*	-0.75	-0.42	1.80^{-}
3	33.00*	28.67*	28.92*	24.10*	4.60*	6.62*	7.78*	4.66*	-1.43	-0.04	0.24	2.88^{+}
Big	46.73*	33.81*	19.86*	23.28*	-12.09*	-3.75*	-1.46	-7.89*	-4.52*	0.77	3.61*	7.47*
	α t-statistic				*Adjusted* R^2 ($\mu = .91$)				*Residual Std. Err.* ($\mu = 2.98$)			
Small	0.19	0.35	0.15	-1.14	0.86	0.89	0.94	0.96	4.20	3.52	2.73	2.03
2	-3.64*	-4.43*	-4.14*	-3.68*	0.94	0.93	0.92	0.90	2.52	2.53	2.99	3.11
3	-2.84^{+}	-3.02*	-4.12*	-3.11*	0.93	0.92	0.92	0.89	2.71	2.98	2.83	3.53
Big	-1.55	-2.33^{+}	-3.26*	0.81	0.96	0.93	0.83	0.88	1.60	2.57	4.24	3.65

Note: $^{-}$, $^{+}$, and * denote significance at the .10, .05, and .01 levels respectively.

Out of all of the non-market risk factors examined, size displayed the greatest predictive power, while inclusion displayed the least predictive power. Value, price/earnings, and yield were all fairly significant in explaining returns, while inclusion was moderately significant. The model with value marginally provided the highest R^2 value of .92, and the yield model provided the lowest residual standard error value of 2.66. Overall, the three factor models provide an improvement of about .10 to the CAPM R^2 of $\approx .81$

5 Conclusion

I find that in the MSCI China A index from January 2008 to July 2015, the size factor plays a significant role, the value, price/earnings, and dividend yield factors play a fairly significant role, and the inclusion factor plays a moderately significant role in explaining the cross section of stock returns. The factor models tend to explain more of the return variation in smaller portfolios than in larger portfolios. There exists a significant return premium in the size and the value factors, a non-significant positive premium in yield, and a non-significant negative premium in inclusion and price/earnings. The historical return premium is more significant for larger sized firms than smaller firms.

The significance of results in this case may partly be due to the smaller sample of stocks examined: the MSCI China A index only includes mid to large-cap A shares with a member count that fluctuates around 500. Currently, the index includes 577 constituents. Nevertheless, there is evidence that tilting portfolios toward firms with smaller size, higher book to market, lower price/earnings, higher dividend yields, or a lower inclusion factor will increase expected returns.

Further research may include optimization of risk factor calculation, examination of the interactions of risk factors, and the predictive power of the risk factors in individual stocks. In this paper, I used the standard Fama and French procedure of splitting size into two groups and other factors into three groups to solve for factor premia, but different splits may result in more explanatory power. In addition, I did not look explicitly at factor interactions in this report, but it may be worth researching especially considering the strong correlations between the value, price/earnings, and yield factors. Lastly, this report uses double-sorted portfolios as a substitute for individual stocks, but an examination of risk factors in individual stock returns may also provide valuable information about the application of these risk factor models.

References

[1] Chen, Jianguo, Kwong Leong Kan, and Hamish Anderson. "Size, book/market ratio and risk factor returns: evidence from China A-share market." *Managerial Finance* 33.8 (2007): 574-594.

[2] Christopher Cheung, George Hoguet, and Sunny Ng, *The Journal of Portfolio Management Special China Issue 2014.* "Value, Size, Momentum, Dividend Yield, and Volatility in Chinas A-Share Market."

[3] Chui, Andy CW, and KC John Wei. "Book-to-market, firm size, and the turn-of-the-year effect: Evidence from Pacific-Basin emerging markets." *Pacific-Basin finance journal* 6.3 (1998): 275-293.

[4] Dongwei, Su, and Mai Yuanxun. "Liquidity and Asset Pricing: An Empirical Exploration of Turnover and Expected Returns on Chinese Stock Markets [J]." *Economic Research Journal* 2 (2004): 95-105.

[5] Drew, Michael E., Tony Naughton, and Madhu Veeraraghavan. "Firm size, book-to-market equity and security returns: Evidence from the Shanghai Stock Exchange." *Australian Journal of Management* 28.2 (2003): 119-139.

[6] Elton, Edwin J., and Martin J. Gruber. "Modern portfolio theory, 1950 to date." *Journal of Banking & Finance* 21.11 (1997): 1743-1759.

[7] Fama, Eugene F., and Kenneth R. French. "Common risk factors in the returns on stocks and bonds." *Journal of financial economics* 33.1 (1993): 3-56.

[8] Fama, Eugene F. "Efficient capital markets: A review of theory and empirical work*." *The journal of Finance* 25.2 (1970): 383-417.

[9] Fama, Eugene F., and Kenneth R. French. "Size, value, and momentum in international stock returns." *Journal of financial economics* 105.3 (2012): 457-472.

[10] Gan, Christopher, et al. An empirical cross-section analysis of stock returns on the Chinese A-share stock market. Working Paper, 2009.

[11] Guo, Mingqiu MQ. *The Comparative Effectiveness of Pricing Variables in the Chinese Stock Market.* Diss. Stern School of Business New York, 2013.

[12] Hakansson, Nils H. "Convergence to isoelastic utility and policy in multiperiod portfolio choice." *Journal of Financial Economics* 1.3 (1974): 201-224.

[13] Hakansson, Nils H. "Optimal investment and consumption strategies under risk for a class of utility functions." Econometrica: Journal of the Econometric Society (1970): 587-607.

[14] Langnan, Chen, and Qu WenZhou. "An empirical study on capital asset pricing model [J]." *Journal of Economic Research* 4 (2000).

[15] Lintner, John. "The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets." *The review of economics and statistics* (1965): 13-37.

[16] Markowitz, Harry. "Portfolio selection." *The journal of finance* 7.1 (1952): 77-91.

[17] Markowitz, Henry. "Portfolio selection: efcient diversication of investments." *Cowies Foundation Monograph* 16 (1959).

[18] Merton, R.C., 1990. Continuous Time Finance, Basil Blackwell, Oxford.

[19] Mossin, Jan. "Equilibrium in a capital asset market." *Econometrica: Journal of the econometric society* (1966): 768-783.

[20] Mossin, Jan. "Security pricing and investment criteria in competitive markets." *The American Economic Review* (1969): 749-756.

[21] Narayan, Paresh Kumar, and Xinwei Zheng. "Market liquidity risk factor and financial market anomalies: Evidence from the Chinese stock market." *Pacific-Basin finance journal* 18.5 (2010): 509-520.

[22] Ross, Stephen A. "A simple approach to the valuation of risky streams." *Journal of business* (1978): 453-475.

[23] Sharpe, William F. "A linear programming algorithm for mutual fund portfolio selection." *Management Science* 13.7 (1967): 499-510.

[24] Sharpe, William F. "Capital asset prices: A theory of market equilibrium under conditions of risk*." *The journal of finance* 19.3 (1964): 425-442.

[25] Treynor, J. L. "Towards a Theory of Market Value of Risky Assets, Unpublished manuscript." 1), Asset Pricing and Portfolio Performance: Models, Strategy and Performance Metrics, London: Risk Books (1962).

[26] Wang, Fenghua, and Yexiao Xu. "What determines Chinese stock returns?." *Financial Analysts Journal* 60.6 (2004): 65-77.

[27] Zhangwu, Jia Quan Chen. "An Empirical Analysis of Efficiency in China Stock Market [J]." *Journal of Finance* 7 (2003): 010.