Diversification in the Chinese Stock Market

Yexiao Xu *

School of Management
The University of Texas at Dallas
and
Shanghai Stock Exchange

Abstract

Modern finance theory suggests that individual investor should hold a well diversified portfolio instead of individual stocks. In practice, one only needs to hold limited number of stocks to achieve the effect of diversification. In this study, we revisit the same issue for the Chinese equity market in three different dimensions. They are (1) relative idiosyncratic risk and risk adjusted portfolio returns; (2) the likelihood of achieving diversification; and (3) turnover of a portfolio. Due to faster declining in the market volatility relative to the aggregate idiosyncratic volatility, one needs to hold 20 stocks in a portfolio in order to diversify away 90% of the total idiosyncratic volatility or equivalent to 95% of the market volatility nowadays. In addition, we have shown that holding one or two stocks will subject to huge negative risk adjusted returns. Therefore, Chinese investors can benefit greatly from diversification with a relatively long investment horizon.

Keywords: Correlation; Cross-sectional Dispersion; Diversification; Likelihood; Transactions Costs;

*I am grateful to ... I also acknowledge financial support from Shanghai Stock Exchange. The address of the corresponding author is: Yexiao Xu, School of Management; the University of Texas at Dallas; Richardson, TX 75083. Email: yexiaoxu@apache.utdallas.edu

Diversification in the Chinese Stock Market

Abstract

Modern finance theory suggests that individual investor should hold a well diversified portfolio instead of individual stocks. In practice, one only needs to hold limited number of stocks to achieve the effect of diversification. In this study, we revisit the same issue for the Chinese equity market in three different dimensions. They are (1) relative idiosyncratic risk and risk adjusted portfolio returns; (2) the likelihood of achieving diversification; and (3) turnover of a portfolio. Due to faster declining in the market volatility relative to the aggregate idiosyncratic volatility, one needs to hold 20 stocks in a portfolio in order to diversify away 90% of the total idiosyncratic volatility or equivalent to 95% of the market volatility nowadays. In addition, we have shown that holding one or two stocks will subject to huge negative risk adjusted returns. Therefore, Chinese investors can benefit greatly from diversification with a relatively long investment horizon.

Keywords: Correlation; Cross-sectional Dispersion; Diversification; Likelihood; Transactions Costs;

1 Introduction

In the past decade or so, both institutional investors and individual investors have experienced large swings in their investment returns. Investors are eagerly seeking advice to weather such volatile markets. What is more striking, as documented by Campbell, Lettau, Malkiel, and Xu (2001), is that the overall market is relatively calm while firm specific risks have gone up significantly. Nowadays, individual U.S. stocks are more than twice as volatile as those in the 1950s on average. This evidence alone bears no consequence on asset prices in the CAPM world, where investors are supposed to only invest in the market portfolio. In other words, the required return from individual investors (thus the cost of capital to the firm) remains the same if investors' holdings are well-diversified. Therefore, the prescription for long-term investors during volatile markets is simple–diversification!

However, survey suggests that 15% of the individual investors in U.S. only hold a single stock and an average investor holds three stocks. Under such a circumstance, an increase in the firm specific risks will have direct impact on the total risk that a typical investor is subject to. This could be more problematic in China. Most Chinese investors only care about short-term gains. In fact, individual investors turnover their investors very often with a very limited number of stocks. Although, there are only 54 close-end mutual funds and 28 open-end mutual funds that manage 81.7 million Yuan as of June, 2003 in China,[1] this only counts 4.58% of the total market capitalization. Majority investors have to hold individual stocks. In contrast, there are over 8,000 open-end mutual funds that manage over six trillion U.S. dollars in U.S. More than half of them are equity mutual funds. In fact, the majority wealth is invested in such a way. Therefore, it is relatively easy for individual investors to diversify their portfolio in the U.S. even though the number of stocks in the individual investment accounts is low. Therefore, compare to the U.S. practice, the level of diversification is far from adequate. This motivates us to study the benefit of diversification

[1]This is the total market capitalization for close-end funds only. Since stock holdings can not exceed 80% of the total portfolio value by government regulation, the equity value of these funds is less than 65.4 million Yuan.

in the Chinese equity markets.

A classical study on the diversification issue was conducted by Evans and Archer (1968). It is widely cited in textbooks as saying that the maximum benefit of naive diversification is achieved when holding about 15 stocks. Based on balancing reduction in the probability of loss and foregone gain opportunities, Jennings (1971) also found that it is optimal to hold a portfolio with 15 stocks.[2] However, the threshold for gain and loss is set arbitrarily. Most studies on the diversification issue use simulation approach. Elton and Gruber (1977) used an analytical approach based on certain distributional assumptions about returns. They concluded that simulation approach may have underestimated the number of stocks needed to achieve diversification. Using the market return as a benchmark for any portfolios with similar risk profiles, Statman (1987) compared the cost of holding a benchmark portfolio to the cost of a portfolio which is the difference between the return would be according to the capital market line and the benchmark return. Applying this methodology, he has shown that it is optimal to hold 30 stocks in a randomly selected portfolio.

This paper differs from the existing studies in the following three important ways. First, most studies focus on developed markets, where investors generally have long-term investment objectives. In contrast, this is the first study examining the diversification issue in the Chinese equity markets. Since most Chinese investors have short-term focus, this study touches the same issue from a different investment environment. Second, most diversification studies have focused on reducing idiosyncratic volatilities. In reality, investors are equally concerned with their returns. At the same time, since most Chinese stock returns are distributed with positive skewness, it is also useful to investigate the impact of diversification on portfolio returns. Moreover, it is unclear as to how often an investor should rebalance his or her portfolio in the presence of transactions costs. In other words, there is a tradeoff between unique return opportunities and increases in the transactions costs. Finally, we investigate the likelihood of achieve certain level of diversification. The traditional

[2]Gain is defined as the probability of a portfolio return exceeding a benchmark return, while loss is defined as the probability of a portfolio return that is less than 75% of the benchmark return.

textbook style diversification graph only tells us the number of stocks K needed to diversify away $\gamma\%$ of the idiosyncratic volatility on average. However, when forming a portfolio of K stocks for an individual investor, it is unlikely to achieve γ. Therefore, we provide a three dimensional diversification plane with the additional dimension showing the probability of achieving the corresponding level of idiosyncratic volatility. Due to lack of institutional investment, majority individual investors will continue to rely on buying individual stocks in China. This study will provide useful guideline for individual investment.

Using all stocks traded on both Shanghai and Shengzheng stock exchanges, we find that there are significant diversification benefit with respect to different measure. In particular, one needs to hold about 20 stocks in order to diversify away most of the idiosyncratic risks. This conclusion is very significant since the number is much higher than the average level in China. At the same, diversification benefit hinges on the relative magnitude of idiosyncratic risk. We have also found that both aggregate idiosyncratic volatilities and market volatility are declining over time. This is very different from those found in the developed markets. However, it also implies increasing difficult in achieving diversification in recent year in China.

The paper is organized as follows. In the next section, we set the theoretical foundation for studying diversification in our framework and discuss the data source. In Section 3, we first study the behavior of systematic versus idiosyncratic volatility in the Chinese equity markets. We then examine the diversification issue using the different approaches, including reducing idiosyncratic volatility, the likelihood of achieving diversification, diversification benefits from return perspective, and turnover in considering transactions costs in Section 4. Section 5 provides concluding comments.

2 Data and Methodology

The goal of any asset pricing theories is to establish a quantitative relationship between risk and return. In general, there are two types of risks for individual securities: the systematic risk, which is determined by common risk factors; and the idiosyncratic risk, which only affects a particular firm or hand full firms. Since idiosyncratic risks are uncorrelated across firms, they can be diversified away in a standard finance theory. Their role in asset pricing has been largely ignored because bearing such risks will not be rewarded with returns by the market. The paper by Campbell, Lettau, Malkiel, and Xu (2001) is the first one in recent year that has renewed the importance of idiosyncratic risk. Indeed, the whole issue of diversification is about reducing idiosyncratic risk. As a practical matter, it is interesting to know how easy it is to diversify away idiosyncratic risks. In order to do so, we begin by studying the dynamic behavior of idiosyncratic risk for the Chinese equity markets.

2.1 Constructing idiosyncratic volatility

In this study, idiosyncratic risk is measured by idiosyncratic volatility. Although total volatility is unobservable, it can be estimated using the standard deviation of returns. In contrast, since idiosyncratic volatility is only part of the total volatility, its decomposition usually depends on a particular asset pricing model. For an initial investigation, we apply the following market model to decompose the total return into systematic and idiosyncratic components,

$$R_{i,t} = \alpha_i + \beta_i R_{m,t} + \epsilon_{i,t} \tag{1}$$

where $R_{i,t}$, $R_{m,t}$, and $\epsilon_{i,t}$ are the individual stock i's return, the value-weighted market return, and stock i's idiosyncratic return, respectively. Note that we have ignored the risk-free rate in equation (1) since we are using daily returns. Moreover, the risk-free interest rate is very stable and is determined by the Chinese government. We measure monthly idiosyncratic volatility using the root mean square of residuals in the corresponding month. This is a more efficient approach to estimate monthly volatility than applying rolling monthly returns. In order to compute the aggregate idiosyncratic volatility, we can then value weight individual stocks idiosyncratic volatilities. Monthly market volatility is also computed from daily market returns.

The market model may be misspecified if we fail to measure the market return accurately, or if other risk factors exist as suggested by Fama and French (1993). Campbell, Lettau, Malkiel, and Xu (2001) have proposed a model-free decomposition procedure based on daily return data. This approach only applies to computing the total aggregate idiosyncratic volatility, which is the focus of this section. In particular, we first aggregate individual stocks' total volatilities. We then compute the aggregate idiosyncratic volatility as the difference between teh aggregate total volatility and the market volatility (see Xu and Malkiel, 2001).

2.2 Methodology

Most studies on the diversification issue have applied simulation approach. However, nothing prevents one from analyzing the problem using algebraic approach (see Elton and Gruber, 1977). In practice, returns are far from normally distributed, especially for individual stocks. Algebraic approach to diversification may not be practical since one has to tolerate restrictive distributional assumptions. Therefore, we will rely on simulation approach in this study. In particular, we first randomly select n stock from a pool of available individual stocks. Portfolio returns are then computed using equal weights. The characteristics of those portfolios can be studied by repeating this process many times.

The conventional approach on diversification starts from investigating the relationship between a portfolio's total (or idiosyncratic) volatility and the portfolio size (n). In this study, we propose two relative measures. In fact, Markowitz (1959) has suggested the following volatility relationship for an equally weighted portfolio with size n,

$$\sigma^2_{p,n} = \left(1 - \frac{1}{n}\right) \bar{Cov}_{(n)} + \frac{1}{n}\bar{\sigma}^2_{(n)} \tag{2}$$

where $\bar{\sigma}^2_{(n)}$ is the average total volatility for individual stocks and $\bar{Cov}_{(n)}$ is the average covariance between all n stocks. When n goes to infinity, we have the relation of $\bar{Cov}_{(n)} \rightarrow \sigma^2_m$ and $\bar{\sigma^2_{(n)}} \rightarrow \sigma^2$. At the same time, equation (2) can be rewritten as,

$$\frac{\sigma^2_{p,n} - \bar{Cov}_{(n)}}{\bar{\sigma}^2_{(n)} - \bar{Cov}_{(n)}} = \frac{1}{n} \tag{3}$$

If we substitute in the limiting value σ^2_m for $\bar{Cov}_{(n)}$ and σ^2 for $\sigma^2_{(n)}$ in equation (3), we can conveniently define,

$$\eta_I(n) = \sqrt{\frac{\sigma^2_{p,n} - \sigma^2_m}{\bar{\sigma}^2 - \sigma^2_m}} \tag{4}$$

as a relative diversification measure. As suggested by Xu and Malkiel (2001), $\bar{\sigma}^2 - \sigma^2_m$ measures aggregate idiosyncratic volatility of individual stocks. $\eta_I(n)$ can thus be intuitively interpreted as a portfolio's idiosyncratic volatility relative to aggregate individual stocks' idiosyncratic volatilities. When both total volatilities of individual stocks and covariances among stocks are constant, $\eta_I(n)$ reduces to $\sqrt{\frac{1}{n}}$. Although this is a very restrictive assumption, we can use $\sqrt{\frac{1}{n}}$ as the theoretical reference line for the relative diversification measure $\eta_I(n)$.

Perhaps, it makes more sense for an investor to know the level of idiosyncratic risk relative to the market risk. The same one percent reduction in idiosyncratic volatility is more important when the market volatility is 5% than when the market volatility is 10%. We, therefore, propose an alternative measure of diversification as the following,

$$\eta_M(n) = \sqrt{\frac{\sigma^2_{p,n} - \sigma^2_m}{\sigma^2_m}} \tag{5}$$

The measure is especially useful when the market volatility changes over time.

Although we rely on simulation approach, there are three innovations including the above diversification measures. After constructing randomly selected portfolios with n stocks, the traditional approach uses the average total volatility from many times of replications. Since volatilities of individual portfolios may not have a normal distribution, such an average will not suggest that there are 50% chance to observe such a volatility when holding a portfolio with n stocks. More generally, investors may want to know the magnitude of a portfolio volatility at which it has certain level of confidence δ. Such confidence levels can be obtained simply by sorting portfolio volatilities from many replications.

As discussed in the next section, most Chinese stocks are distributed with positive skewness. A combination of limited number of stocks may have large probability of outperforming the market. It is thus useful to compute the likelihood for a size n portfolio to earn return exceeding 50%, 75%, 90%, and 100% of the market return over the same period. Theoretically, the likelihood should be decreasing when the portfolio size n increases. This information focuses on performance, which is complementary to the diversification diagram.

The above two approaches are either from volatility perspective or from return perspective. Modern portfolio theory suggests that return and risk are related. A less diversified portfolio is an inefficient portfolio in the sense that it contains too much idiosyncratic risk. Therefore, a portfolio's actual average return with a total volatility of σ_a should be compared to that of an efficient portfolio on the capital market line with the same total volatility. In other words, we propose the following Sharpe ratio adjusted excess return measure ER.

$$ER = \bar{R}^a - \bar{R}^e = \bar{R}^a - \frac{\sigma_a}{\sigma_m}\bar{R}^m, \tag{6}$$

where $\bar{R}^a$ is the actual average return, $\bar{R}^m$ is the market return, and σ_m is the market volatility. Naturally, ER will change overtime when the market return changes over time. In order to be comparable, we can scale ER measure by the absolute value of of the market return, which we define as the relative return measure θ.

$$\theta = \frac{ER}{|\bar{R}^m|} = \frac{\bar{R}^a}{|\bar{R}^m|} - \frac{\sigma_a}{\sigma_m}\frac{\bar{R}^m}{|\bar{R}^m|}, \tag{7}$$

Both the ER and θ measures should be close to zero when portfolio size n is large enough. Whether these measures go to zero fast enough when n increases is the key issue here.

2.3 Description of the data

China established its first stock exchange–Shanghai Stock Exchange in December 18, 1990. The second stock exchange–Shenzhen Stock Exchange was introduced in early 1991. The Securities Committee of the State Council (which was later merged into the China Securities Regulatory Commission) approves stock listing and decides which stock should be traded on which exchanges. At that time, there were only A-share stocks available for domestic investors using RMB denomination. The B-share markets were introduced in February 1992, which is only for foreign investors with U.S. dollar denomination. Historically, the B-share markets were very illiquid with large discounts relative to the A-share markets. The discounts have decreased substantially after allowing domestic investors to invest in the B-share markets. Since B-share markets are much smaller than the A-share markets with less than 10% of the total number of stocks outstanding, we are focusing on the A-share markets only in this study.

We use stock return data from the 2002 version of China Stock Market and Accounting Research Database (CSMAR). This is one of the most widely used security database in China. For the first year, there are only eight traded stocks. The total number of stocks increased to 14 and 53 at the beginning of year 1992 and 1993, respectively. Since then we have experienced a rapid increase in the number of stocks traded in the two exchanges. Therefore, our sample covers all individual stocks from the beginning of 1994 to the end of 2002. In order to estimate volatility accurately, we use daily stock returns. Due to the speculative nature of the Chinese stock markets, there is virtually no inactively traded stock except for special circumstances. Therefore, market micro-structure effect such as non-synchronous trading will not be a problem. Table 1 reports the summary statistics for our data set.

Insert Table 1 Approximately Here

The number of stocks has gradually increased from 252 to 1150 over the nine-year sample period from 1994 to 2002, which has been doubled twice. After 2001, new shares are only traded in Shanghai Stock Exchange. The total market capitalization has increased from 40 (about \$4.82 billion U.S. dollars) billion RMB to 400 billion (about \$48.19 billion U.S. dollars). Although relatively small (about the size of a large U.S. company such as McDonald), it plays an important role in the overall Chinese economy. There are several indices available but lack of representativeness. Since we study stocks traded on both exchanges, we have constructed both value weight an equal wieght indices in this study. Since the shares owned by the state are prohibited from trading, we use tradable shares to compute the total market capitalization from asset pricing perspective. Over the nine year period, the value-weighted index has returned an arithmetic average annual return of 38.82% with a standard deviation of 55.28%. This level of volatility far exceeds that of the U.S. stocks. Since we are concerned about individual stocks, we have also reported distributions for individual stocks over time in Table 1. The arithmetic average return is 15.76% with an average volatility of 38.67% across individual stocks. Therefore, return differences are substantial at any given time. This is a necessary condition to achieve diversification by holding a portfolio. Returns are also positively skewed since the mean always exceed the corresponding median. (see Figure 1) In other words, it is more likely to observe large return than small return. This may seem to provide motivation for investors to hold individual stocks instead of portfolios. We will study this issue further in the next section.

Insert Figure 1 Approximately Here

As documented by Morck, Yeung, and Yu (2000), stocks tends to have high coefficient of determinations (R^2) using a market model in the emerging markets. This is also confirmed in Table 1 using daily returns. Although average R^2 across individual stocks fluctuated from 27% to 75% over time, the average level exceeds 50%. Therefore, it is interesting to ask if it is much easier to achieve diversification in the Chinese equity markets than that in the U.S. markets.

If the Chinese stock market is very speculative, the turnover ratio should be relatively high. We have also reported the annual turnover defined as the ratio between total trading volume and the average market capitalization in Table 1. Clearly, turnover exceeds 10 in the early years which is extremely high. Although it has come down gradually, it was still around 5 at the end of 90's and further decreased to 2 in 2002. This trend suggests that the average holding period has increased in recent years. Another unique feature about the Chinese stock market is the state ownership and the legal person shares. Those shares are prohibited from trading. Fore example, in 1994 only 16% shares are tradable on average with a huge variation from stock to stock. The situation has improved greatly since 1996. About 31% of the total shares were tradable. This number increased to 35% in 2002.

Since diversification is mainly about reducing volatility, we now take a brief look at the level of volatilities over time. The return volatilities for the value weighted index were very large from 1994 to 1997 which varied from 33% to 70%. The average total volatilities for individual stocks are also large from 48% to 84%. Since then, the market volatility has decreased substantially, fluctuating between 19% to 25%. This may be partly attributed to the 10% price limit implemented towards the end of 1996. Although both the market volatility and the aggregate total volatilities for individual stocks have come down greatly, idiosyncratic volatility takes a greater part of the total volatility nowadays. This suggests that diversification is increasingly important nowadays.

3 The Volatility and Correlation Structure

If a market model of equation (1) describes individual stock returns, the total volatility of a size n portfolio can be expressed as,

$$\sigma_{p,n}^2 = \beta_p^2 \sigma_m^2 + \frac{1}{n} \bar{\sigma}_I^2 \tag{8}$$

where $\bar{\sigma}_I^2$ is the average idiosyncratic volatility. Therefore, a portfolio's idiosyncratic volatility $\sigma_{p,n}$ decreases with the portfolio size (n) at a speed of $1/\sqrt{n}$. This conclusion is based on the assumption of independent CAPM residuals across individual stocks. As shown by Fama and French (1993), a multi-factor model serves better to capture return variations. In other words, the CAPM residuals are correlated to some degree. Therefore, a simulation approach is needed to access the actual speed of diversification. Equation (8) also suggests that a portfolio's total volatility depends on the level of idiosyncratic volatility. From the U.S. experience, we have learned that the volatility structure, especially idiosyncratic volatility has changed in the past decade. Therefore, it is important to study the dynamic behavior of idiosyncratic volatility v.s. market volatility first.

The market volatility was very high before 1996 as shown in the first panel of Figure 2. For example, there were two huge spikes in October 1994 and May 1995, which corresponded to the events of introducing new IPO trading mechanism and stopping trading government bond futures. After implementing the price limit in December 1996, the market volatilities were stabilized. In contrast, the monthly idiosyncratic volatilities behave somewhat differently from that of the market volatility. As discussed in the last section, there are different ways to construct idiosyncratic volatilities. Surprisingly, the two methods produce very similar idiosyncratic volatility estimates. Therefore, we only plot the monthly aggregate idiosyncratic volatility estimates using the CAPM residuals in the second panel of Figure 2. In particular, we first compute residuals by fitting a market model to daily returns over the sample period of a particular stock. Monthly idiosyncratic volatilities are then computed from daily residuals. Finally, we value weight idiosyncratic volatility in any given month. The aggregate idiosyncratic volatility fluctuated between 1% to 4%. Such fluctuations existed in both pre and post 1996 sample period. In other words, price limit has less impact to the idiosyncratic volatility than to the market volatility.

Insert Figure 2 Approximately Here

Changes in the level of volatilities will affect the degree of diversification for the same portfolio size. In order to visualize any possible trends, we have also plotted a twelve moving average in Figure 2. Clearly, as the solid line shows, both the market volatility and the idiosyncratic volatility have exhibited downward trends. This is in contrast to the U.S. experience documented by Campbell, Lettau, Malkiel, and Xu (2001), which found an increasing trend in idiosyncratic volatility only with the stable market volatility.

Diversification can also be viewed as reducing the the relative importance of the idiosyncratic volatility, not the total portfolio volatility $\sigma^2_{p,n}$. More preciously, when the level of volatility changes over time, one would like to know how easy it is to get ride of idiosyncratic volatility. This question can only be answered when the portfolio idiosyncratic volatility is measured relative to the level of market volatility. This can be seen from rewriting equation (8) as,

$$\frac{\sigma^2_{p,n}}{\sigma^2_m} = \beta^2_p + \frac{1}{n}\frac{\bar{\sigma}^2_I}{\sigma^2_m} \tag{9}$$

In other words, whether it is easier or more difficult to achieve diversification nowadays depends on the relative magnitude of the time trends in both the idiosyncratic volatility and the market volatility. Therefore, we perform a unit root test with a time trend in Table 2. Since volatilities are positive, we use log volatility. In addition, we have allowed six lags in the following testing equation to account for the persistence in the volatility.

$$\ln(\sigma_t) = \mu + \gamma t + \rho \ln(\sigma_{t-1}) + \alpha_1 \Delta \ln(\sigma_{t-1}) + \cdots + \alpha_6 \Delta \ln(\sigma_{t-6}) + \epsilon_t. \tag{10}$$

As shown in Table 2, for both the market volatility and the idiosyncratic volatility, no matter whether it is equally weighted or value weighted, we have failed to reject the unit root hypothesis using the Dicky-Fuller t test statistics. In other words, there is no stochastic trend for any of the volatility series considered. Therefore, we can now test the hypothesis of deterministic trend by applying the conventional t statistics. Clearly, the t ratios are all significant at a 1% level. The downward trends in both the market and the idiosyncratic volatilities are confirmed. Since we have used log volatilities in our estimation, the coefficient γ can be directly interpreted as the percentage change in the volatility. In particular, the γ estimate is about 0.1% per month no matter how we estimate the idiosyncratic volatility. The trend coefficient γ is about 0.28% per month. In other words, the decreasing trend in the market volatility is more than twice as large as that in the aggregate idiosyncratic volatility.

Insert Table 2 Approximately Here

Since the residual returns from the CAPM can be correlated to some degree, an alternative way to study the impact of changing volatilities on diversification is to investigate the correlation structure. Since correlations among individual stocks are largely due to common risk factors, residual risks are difficult to get ride of when two stocks do not share too many common factors. In this case, correlations among stocks are low. Therefore, there should be negative relationship between the degree of diversification and the level of correlation structure. We have computed all pairwise correlations among individual stocks using daily returns. The average correlations are plotted in the first panel of Figure 3.

Insert Figure 3 Approximately Here

Average correlations are very high (over 65%) during 1994 and 1995. It has steadily declined to about 35% during the four year period from 1998 to 2001. The correlation has been creeping up in 2002. This suggests that, on average, it becomes more difficult to diversify away the idiosyncratic risk in recent year than in the early period. This conclusion is similar to the U.S. case despite the differences in the volatility structures in the two countries (see Campbell, Lettau, Malkiel, and Xu, 2001). It is also interesting to know how much correlation is due to the single most important market factor. Therefore, we have also computed the pairwise average correlations using residual returns from a market model. The second panel of Figure 3 shows that such correlations are very small overall. It also shares a decreasing trend. The average residual correlations are about 3% and 1% for the period from 1994 to 1997 and the period from 1998 to 2002, respectively.

4 The Diversification Benefit for the Chinese Stock Market

At a first glance, diversification is a strict forward problem. According to the CAPM theory, one should hold a "very" diversified portfolio and be rewarded with the market return. However, it is not obvious when we try to quantify the word "very." It is important to know how much idiosyncratic risk can be reduced when holding a size n portfolio. As shown in Figure 1 that an individual stock returns are distributed with heavy tails and positive skewness, it is more likely to observe large positive returns than that under a normal return distribution. When the degree of diversification increases, such unique distribution properties will likely vanish. If we consider idiosyncratic volatility as a cost, it is also important to study diversification from return benefit perspective. In particular, we will study in this section the question whether the likelihood of observing certain level of return increases when the degree of diversification increases. Turnover is another key issue facing an investor. It is interesting to know how turnover affects the likelihood.

4.1 Diversification By Reducing Idiosyncratic Volatility

At the beginning of each year from 1994 to 2002, we randomly select n stocks to form a portfolio with equal weights, where $n = 2, 3, \cdots, 30$. We then compute the total standard deviation for a portfolio using daily returns. For the monthly volatility, we multiply the standard deviation by a factor of $\sqrt{21}$, where 21 is the average trading days in a month. In order to compute annual portfolio returns, we first compound individual stocks' daily return. The compounded returns are then equally weighted to form portfolio returns. This process is repeated for 1250 times. As discussed in the last section, the volatility structure has changed over time, we have also separated our sample period into two subsample periods of 1994–1997 and 1998–2002. For the purpose of computing the relative diversification measure $\eta_I(n)$, we have also calculated the aggregate total volatility σ by equally weighting individual stocks' volatilities. We have plot the diversification measure $\eta_I(n)$ in Figure 4.

Insert Figure 4 Approximately Here

For the whole sample period, the solid line in the first panel of Figure 4 suggests that the speed of diversification is very fast at the beginning. About 70% of the idiosyncratic volatility can be diversified away when there are four stocks in a portfolio. An additional 10% idiosyncratic volatility can be reduced when holding eight stocks in a portfolio. The reduction in idiosyncratic volatility is very slow thereafter. For the two subsample periods, the diversification line confirms our finding on the volatility trend. Since the broken line (for the early sample period) is above the dotted line (for the recent sample period), it is indeed more difficult to diversify away idiosyncratic volatility nowadays.

Diversification effect can also be measured with respect to the market volatility. We have shown the relationship between $\eta_M(n)$ and portfolio size n in the second panel of Figure 4. Despite the fact of decreasing idiosyncratic volatility, the relative idiosyncratic volatility with respect to the market volatility is twice as large in recent subsample period as that in the early subsample period. Such a pattern continues to hold when the portfolio size increases. This confirms our conclusion of increasing diversification benefit using $\eta_I(n)$ measure. Moreover, the relative idiosyncratic volatility decreases very fast. For example, it gets reduced to 20% when there are four stocks in a portfolio in the recent subsample period.

In order to conclude on the portfolio size needed to reach reasonable level of diversification, we report both the absolute level and the relative idiosyncratic volatilities in Table 3. For the absolute value of idiosyncratic volatility, it is interesting to see that there is not much difference between the two subsample periods with respect to different portfolio size. Using relative measure $\eta_I(n)$, we conclude that one need to hold at least 20 stocks in oder to diversify away 90% of the total idiosyncratic volatility in the recent year. One only need to hold 13 to achieve the same level of diversification in the early years. If our benchmark is relative to the market volatility, with 20 stocks in a portfolio, the undiversified idiosyncratic risk only counts 6% of the market volatility. If one can only tolerate equivalent to 5% of the market volatility, a portfolio should contain 26 stocks.

Insert Table 3 Approximately Here

The above analysis has only examined the average level of idiosyncratic volatility over many times of replications. In practice, however, investors only have chance to form a portfolio once at any give time. Therefore, from a practical perspective, it is equally important to know the level of diversification at a certain confidence level. For this purpose, we present a three dimensional diversification graph showing the relationship between relative idiosyncratic volatility to market volatility at each confidence level in Figure 5 over the whole sample period. For example, for the top 20% of the most volatile stocks, the average volatility is about 1.65 times as large as large as that of the market volatility. In contrast, for the least volatile 20% stocks, the average volatility is just 1.3 times as large as that of the market volatility. Furthermore, such idiosyncratic volatilities decrease almost linearly from the highest percentile to the lowest percentile. It is also interesting to note that the speed of diversification varies with the percentile. It is much slower to diversify away 90% of the idiosyncratic volatility at the top 20% level than that at the lower 20% level. For example, at the top 20% level, it takes more than 30 stocks to diversify away 90% of idiosyncratic volatility. In contrast, one only need to hold 15 to achieve the same level of diversification at the lower 20% volatility group.

Insert Figure 5 Approximately Here

4.2 How Does Diversification Affect Portfolio Returns?

For an average investor, when holding a "well diversified" portfolio, she or he will be rewarded for the market return. Since individual Chinese stock returns are positively skewed with heavy tails, less diversified portfolios may sometime provide above market returns. However, this should be put into a risk-return perspective as discussed in the second section. In general, a less diversified portfolio will subject to high total volatility. Those excess returns may not even sufficient for the high level of volatility. Therefore, we have also plotted risk adjusted excess return in panel A of Figure 6.

Insert Figure 6 Approximately Here

First of all, for any portfolio size, the excess returns are negative. On average, there is an average negative excess return of 16% when holding a single stock. This is huge when compared to the 1% transactions costs on average. Such a negative return goes to zero very fast at the beginning. For example, for a four stock portfolio, the negative excess return goes up to -5%. When holding 14 stocks, it further increases to -2%. For different sample periods, it seems that the excess returns decrease to zero faster in the early subsample period than in the recent subsample period.

Returns fluctuate widely from time to time. It might be more realistic to measure the excess returns relative to the absolute market level at that point of time. Using equation (6), we have also plotted the relative excess return in the second panel of Figure 6. When holding a single stock, the relative excess return is equivalent to 32% of the market return. For a four stock portfolio, it is about 12% of the market return. Again, one needs to hold 20 stock in order to make the negative excess return greater than 2%.

From risk control perspective, investors would also like to know the probability (likelihood) of maintaining the principle amount of investment. We have computed such probability each year for different portfolio sizes. In particular, we have plotted these average probabilities over different subsample periods in Panel A of Figure 7. It is interesting to see that there is 56% chance to see a positive return every year in the early subsample period. When the portfolio size increases, such a probability rises to a maximum of 62% for a portfolio size of 4. Most increases in the likelihood vanish when holding a portfolio with 30 stocks. For the recent subsample period, however, the likelihood curve looks very different from that in the early period. The probability is gradually increasing from 53% for individual stocks to 60% for a portfolio of 20. For the whole sample period, the likelihood curve is again look like that in the first subsample period. Since the average return is positive, it makes perfect sense that the likelihood of observing positive return is greater than 50%. However, the hump shape likelihood can only occur with a non-symmetric distribution, which is the case here.

Insert Figure 7 Approximately Here

Zero percent is just one particular number. One might also want to know the likelihood of observing returns that are greater than certain percentage of the market return. This is plotted in a three dimensional graph in the second panel of Figure 7. The likelihood of observing market return is always less than but approaching 50%. Such a monotonic likelihood curve is increasing with decreases in reference level. When the reference level is relatively low, such as 0% or 20% of the market return, the likelihood curve is actually hump shaped. In general, diversification improves the likelihood of observing certain level of return too. Therefore, it also pays to diversify even from the return perspective.

4.3 Turnover And Diversification

Another practical issue facing a diversified investor is the holding period. Frequently rebalancing a portfolio may increase the chance of capturing new investment opportunities. However, it could also induce high transactions costs. Therefore, we study the diversification issue with the consideration of the holding period. In particular, we simulate portfolio returns over a single year, a two-year period, and a four-year period. Since turnover is our focus here, we use one year holding period as the benchmark for comparison. In particular, at the beginning of each year, we randomly form a portfolio with size n. We then compute the portfolio returns using the following two-year or four-year daily returns of the same stocks. We then compute the likelihood of observing a certain level of cumulative returns over a four-year period. The average volatility should be the same under different holding periods as plotted in Figure 4. In China, there is an average 1% transactions costs. We can recompute the cumulative returns by charging the transactions costs. For example, when the holding period is one year, we can subtract 1% transactions costs from each of the annual return before compounding to the four-year return. Similarly, when the holding period is two years, we only subtract 1% transactions costs from each of the two-year returns before compounding to the four-year return. For an average cumulative annual return of 5% over a four-year period, we plot the corresponding likelihood for one, two, and four year holding periods in Figure 8. Panel A and B of Figure 8 shows the likelihood without and with transactions costs, respectively.

Insert Figure 8 Approximately Here

When considering a four-year investment horizon, it is relatively easy to achieve an average cumulative return of 5%. For example, with a 94% confidence, one only need to hold 4 stocks when there are no transactions costs and 5 stocks with transactions costs. In general, no matter whether there are transactions costs, the likelihood of observing a certain level of returns for the same portfolio size improves when the holding period increases. Such an improvement is much larger when imposing transactions costs than that under no transactions costs. Similarly, for the same probability of achieving 5% annual returns, one needs to hold a larger portfolio if rebalancing a portfolio often. By comparing different curves in Panel A or in Panel B of Figure 8, we can also learn that the improvement in the likelihood is relatively small from the two-year holding period to the four-year holding period. Therefore, when there is more than 8 stocks in a portfolio, rebalancing a portfolio every two years is almost as good as rebalancing it every four years.

We can also study the relationship between portfolio size and portfolio turnover for different average returns. These results are summarized in Table 4 for different confidence levels. When investors only require a 2% annual return, or 8.24% over a four-year period, the portfolio size needed with 95% confidence level is not very different under different turnover schedules. It is three stocks in this case with or without transactions costs.

However, when investors require an annual return of 8% (i.e., 36% over a four-year period), they need to hold at least 16 stocks in a portfolio if the portfolio compositions change every year and are subject to 1% transactions costs. When rebalancing the portfolio once every two years, the portfolio size can be reduced to 9 under the same scenarios. The portfolio size increases dramatically if we increase the confidence level to 98%. Therefore, it pays to rebalance a portfolio less often

Insert Table 4 Approximately Here

5 Conclusions

The benefit of holding a diversified portfolio has been well understood in the developed capital markets. Although the Chinese capital markets have over twelve years of history, they are still premature. Most investors focus on short-term gains rather than pursuing long-term investment objectives. At the same time, the overall market is very volatile. In such an environment, it is important to know if diversification benefit carries over. As a first study of its kind, we have examined the diversification issue not only from reducing the idiosyncratic risks perspective, but also from portfolio return perspective. We have also proposed two ratios to measure the degree of diversification.

Diversification is about reducing the unnecessary idiosyncratic risks facing an investor. Contrast to the U.S. experience documented by Campbell, Lettau, Malkiel, and Xu (2001), we have found decreasing trends in both idiosyncratic volatility and in the market volatility. Since such decreases in the market volatility are larger than those in the aggregate idiosyncratic volatility, it still implies an increasing benefit in the diversification in recent period, a similar conclusion found in the U.S. markets.

Quantitatively, we have shown that one needs to hold 20 stocks in a portfolio in order to diversify away 90% of the total idiosyncratic volatility or equivalent to 95% of the market volatility. We have also shown that holding one or two stocks will subject to huge negative risk adjusted returns. Therefore, individual investors should make every effort avoiding holding too few stocks. In addition, it is a good idea to hold a portfolio over a relatively long period even in China.

This study also has policy implications. Currently, there is no specific risk control requirement for a public investment company, such as a mutual fund, except the two ten percent requirement. The total market capitalization of a single stock can not exceed ten percent of the total portfolio value. In other words, the minimum diversification requirement is to hold eight stocks.[3] Clearly, this is not sufficient using any measure discussed in this paper. We suggest to change the holding requirement to 5%, which corresponds to a minimum diversification requirement of 16 stocks.

[3]Current law also requires a fund to hold at least 20% of government bonds.

References

[1] Blume, M. E., and I. Friend (1975), 'The Asset Structure of Individual Portfolios and Some Implications for Utility Functions,' *Journal of Finance,* Vol. 30, pp. 585-603.

[2] Campbell, John Y., Martin Lettau, Burton G. Malkiel, and Yexiao Xu (2000), 'Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk,' *The Journal of Finance,* Vol. 56, pp. 1-43.

[3] Elton, E. J., and M. J. Gruber (1977), 'Risk Reduction and Portfolios Size: An Analytical Solution,' *Journal of Business,* Vol. 50, pp. 415-437.

[4] Evans, J. L., and S. H. Archer (1968), 'Diversification and Reduction of Dispersion: An Empirical Analysis,' *Journal of Finance,* Vol. 23, pp. 761-767.

[5] Fama, Eugene F., and Kenneth R. French (1993), 'Common Risk Factors in the Returns on Stock and Bonds,' *Journal of Financial Economics,* Vol.33, pp.3-56.

[6] Markowitz, H. (1959), *Portfolio Selection: Efficient Diversification of Investments,* New York: John Wiley & Sons.

[7] Morck, R., B. Yeung, and W. Yu (2000), 'The Inofrmation Content of Stock Markets: Why Do Emerging Markets Have Synchronous Stock Price Movement,' *Journal of Financial Economics,* Vol. 58, 215-260.

[8] McDonald, J. B. and Y. Xu (1995), 'A Generalization of the Beta Distribution with Applications,' *Journal of Econometrics* Vol. 66, pp. 133-152.

[9] Sharpe, William F. (1964), 'Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk,' *Journal of Finance* Vol. 19, pp. 425-442.

[10] Statman, M. (1987), 'How Many Stocks Make a Diversified Portfolio?' *Journal of Financial and Quantitative Analysis* Vol. 22, pp. 353-363.

Table 1: **Summary Statistics for A-share Stocks Traded on Both Stock Exchanges** This table reports summary statistics for A-share stocks. Individual stocks' annual returns are compounded returns, while annualized volatilities are computed from daily stock returns. Both equal weighted index returns (R_{EW}) and value weighted index returns (R_{VW}) are computed from all tradable shares. The reported distributions for return and volatility are cross-sectional distribution. R^2 is the coefficient of determinant from fitting a market model to daily returns. Turnover is the ratio between total annual trading volume and the average market capitalization. "Float" refers the ratio between tradable shares and total outstanding share including state owned share and legal person share. All numbers are in percentage except for number of stocks and turnover.

	Individual Stock Return					# of	Market		R^2	
Year	10%	50%	90%	Mean	C.Std	Stocks	R_{EW}	σ_{EW}	Mean	C.Std
1994	-43.84	-17.99	31.11	-11.07	30.57	252	-5.95	73.16	74.96	9.39
1995	-29.23	-10.73	21.70	-5.36	24.95	272	-6.43	45.00	70.85	15.22
1996	0.51	64.50	195.6	87.71	90.79	364	93.21	39.22	46.56	8.69
1997	-19.07	19.46	92.56	29.92	47.39	613	30.31	32.79	46.87	9.38
1998	-22.19	10.43	57.96	16.32	39.68	764	16.87	19.92	31.16	10.71
1999	-8.89	18.99	70.08	26.36	38.05	861	28.97	24.49	38.95	14.27
2000	20.94	56.46	119.4	66.55	46.46	961	71.09	19.07	27.33	11.18
2001	-35.06	-20.61	-2.17	-18.96	15.46	1090	-18.47	19.96	50.20	16.80
2002	-30.09	-17.31	4.37	-14.45	15.69	1150	-13.88	24.06	60.03	16.87
	Individual Stock Volatility						Turnover		Float	
Year	10%	50%	90%	Mean	C.Std	σ_{VW}	Mean	C.Std	Mean	C.Std
1994	69.34	84.63	96.41	83.87	10.50	69.88	10.49	5.00	16.15	21.00
1995	43.45	52.93	63.48	53.09	8.08	45.31	4.552	2.74	16.60	20.39
1996	47.50	57.19	71.02	58.51	9.45	39.01	11.84	5.55	31.46	17.81
1997	40.83	47.40	54.80	47.57	5.44	33.20	7.834	2.04	29.92	14.78
1998	28.81	35.67	44.62	36.38	6.35	18.85	4.873	2.08	30.32	13.41
1999	33.14	39.10	46.47	39.68	5.49	25.21	4.502	1.73	30.95	13.12
2000	29.95	37.04	44.69	37.28	6.03	19.58	5.279	1.61	33.29	13.59
2001	23.05	27.93	34.46	28.45	4.68	19.01	2.339	1.18	35.10	14.08
2002	23.79	30.88	38.90	31.22	6.19	22.64	2.165	1.47	35.25	13.82

Table 2: **Testing Volatility Trend**

This table provides the significant tests for stochastic trend vs. time trend in both market volatility and idiosyncratic volatility over the entire sample period from 1994-2002. All monthly volatilities are computed using daily returns. Two measures of idiosyncratic volatilities are used. They are indirect measure according to Xu and Malkiel (2001) and the root mean square of the CAPM residuals. In addition, we use the following model to test trend in each volatility series,

$$Model: \quad \ln(\sigma_t) = \mu + \gamma t + \rho \ln(\sigma_{t-1}) + \alpha_1 \Delta \ln(\sigma_{t-1}) + \cdots + \alpha_6 \Delta \ln(\sigma_{t-6}) + \epsilon_t.$$

	Using Equal Weighting				Using Value Weighting			
	μ	γ	ρ	R^2	μ	γ	ρ	R^2
	Market Index							
Estimate	-0.5540	-0.0028	0.3785	0.400	-0.5521	-0.0029	0.3814	0.412
(St.D.)	0.1486	0.0009	0.1597		0.1485	0.0009	0.1584	
t	-3.7268	-2.9173	2.3701		-3.7171	-2.9956	2.4082	
DF-t			-3.8905				-3.9045	
	Idio. Volt. Constructed Using Xu's Method							
Estimate	-0.8610	-0.0011	0.4620	0.489	-0.8373	-0.0010	0.4864	0.488
(St.D.)	0.2061	0.0004	0.1278		0.2036	0.0004	0.1240	
t	-4.1768	-2.6563	3.6151		-4.1120	-2.4275	3.9218	
DF-t			-4.2082				-4.1409	
	Idio. Volt. Constructed Using the CAPM Residuals							
Estimate	-0.7952	-0.0011	0.4981	0.517	-0.8497	-0.0012	0.4705	0.499
(St.D.)	0.2025	0.0004	0.1269		0.2118	0.0004	0.1310	
t	-3.9260	-2.7104	3.9242		-4.0103	-2.7602	3.5911	
DF-t			-3.9528				-4.0408	

Table 3: **Diversification and Idiosyncratic Volatility**

This table reports absolute and relative measures of portfolio idiosyncratic volatilities for different portfolio size. Portfolios are formed by randomly select n stocks and using equal weights. Simulations are done with 1250 replications. Portfolio idiosyncratic volatility is defined as the difference between portfolio total volatility and volatility of market index return over the same time period. "Relative to Stk. Idio. Volt." stands for portfolio idiosyncratic volatility relative to individual stocks' idiosyncratic volatility. "Relative to Market Volt." stands for portfolio idiosyncratic volatility relative to market volatility.

Pfl.	Portfolio Idio. Volt.			Relative to Stk. Idio. Volt.			Relative to Market Volt.		
Size	94-02	94-97	98-02	94-02	94-97	98-02	94-02	94-97	98-02
1	4.01	4.02	4.00	100.	100.	100.	50.8	32.4	65.6
2	2.28	2.24	2.30	56.3	54.2	57.9	29.0	17.9	37.9
3	1.63	1.60	1.65	40.1	38.0	41.7	20.7	12.7	27.2
4	1.28	1.25	1.30	31.4	29.2	33.1	16.3	9.82	21.4
5	1.06	1.04	1.07	25.8	23.7	27.6	13.5	8.05	17.8
6	0.910	0.895	0.922	22.2	20.1	23.8	11.5	6.87	15.3
7	0.802	0.788	0.813	19.6	17.5	21.2	10.2	6.00	13.5
8	0.725	0.719	0.730	17.7	15.8	19.2	9.16	5.43	12.1
9	0.661	0.657	0.664	16.1	14.2	17.6	8.32	4.91	11.0
10	0.610	0.609	0.610	14.8	12.9	16.3	7.64	4.50	10.2
11	0.565	0.564	0.565	13.7	11.8	15.2	7.08	4.15	9.42
12	0.529	0.531	0.527	12.8	10.9	14.2	6.61	3.87	8.79
13	0.501	0.507	0.496	12.1	10.3	13.5	6.23	3.66	8.28
14	0.474	0.484	0.466	11.4	9.71	12.7	5.87	3.47	7.79
15	0.453	0.467	0.442	10.9	9.28	12.2	5.58	3.32	7.39
16	0.433	0.448	0.421	10.4	8.77	11.7	5.33	3.17	7.06
17	0.416	0.435	0.401	9.96	8.45	11.2	5.10	3.06	6.73
18	0.400	0.420	0.384	9.57	8.08	10.8	4.89	2.93	6.45
19	0.387	0.410	0.369	9.24	7.81	10.4	4.70	2.84	6.19
20	0.373	0.396	0.356	8.90	7.46	10.1	4.53	2.72	5.98
21	0.362	0.384	0.344	8.61	7.15	9.77	4.38	2.63	5.79
22	0.350	0.373	0.332	8.33	6.86	9.50	4.24	2.53	5.60
23	0.341	0.364	0.322	8.08	6.64	9.23	4.11	2.47	5.42
24	0.332	0.356	0.314	7.89	6.44	9.05	4.01	2.40	5.29
25	0.325	0.349	0.306	7.71	6.25	8.88	3.91	2.34	5.17
26	0.318	0.342	0.298	7.53	6.09	8.68	3.81	2.28	5.04
27	0.311	0.335	0.292	7.35	5.89	8.52	3.73	2.22	4.93
28	0.304	0.328	0.284	7.17	5.71	8.35	3.64	2.17	4.82
29	0.296	0.320	0.277	6.98	5.50	8.17	3.55	2.11	4.70
30	0.290	0.314	0.271	6.80	5.33	7.99	3.46	2.06	4.59

Table 4: **Portfolio Size and Turnover**

This table reports the needed portfolio size in order to observe certain level of average annual returns over a four-year period with different turnover under different confidence level. These numbers are computed with and without transactions costs. "$E.\ 2$" stands for holding a portfolios for two years. The four year cumulative returns are computed using the consecutive two-year returns. "At Least α% Return" stand for at least observing an average cumulative annual return of α%.

Confidence Level	At Least 2% Return			At Least 5% Return			At Least 8% Return			At Least 10% Return		
	E. 1	E. 2	E. 4	E. 1	E. 2	E. 4	E. 1	E. 2	E. 4	E. 1	E. 2	E. 4
	Without transactions costs											
99%	7	5	4	12	9	8	30+	30	25	30+	30+	30+
98%	5	4	3	8	6	5	27	18	14	30+	30+	30+
95%	3	3	2	5	4	3	10	7	6	30+	22	17
90%	2	2	-	3	2	2	5	4	3	8	6	5
85%	1	1	1	2	-	-	3	3	2	5	4	3
	With 1% transactions costs											
99%	8	6	5	15	10	9	30+	30+	27	30+	30+	30+
98%	6	4	3	11	7	6	30+	25	16	30+	30+	30+
95%	3	3	2	6	4	3	16	9	7	30+	30+	21
90%	2	2	-	3	3	2	7	4	3	14	8	6
85%	-	-	1	2	-	-	4	3	2	6	4	3

Figure 1. Skewness and Kurtosis of Individual Stock Returns

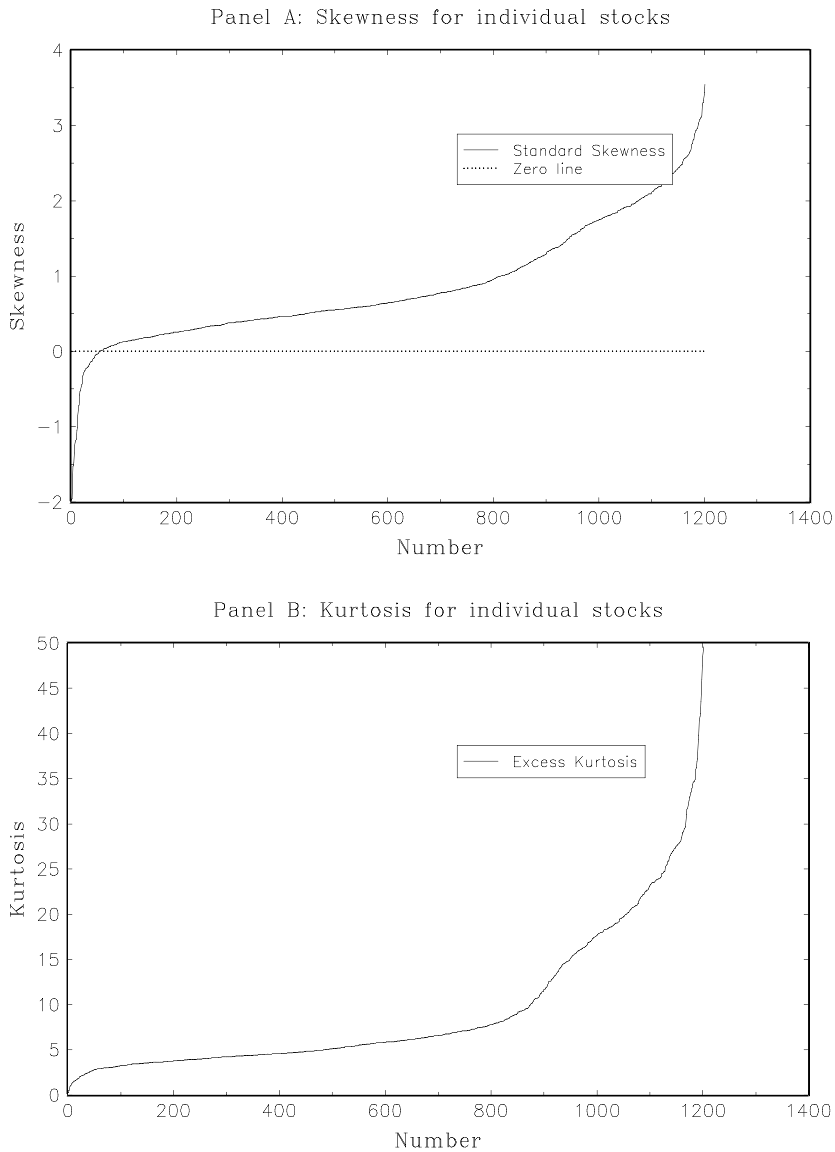


Figure 2. Market Volatility and Aggregate Idiosyncratic Volatility

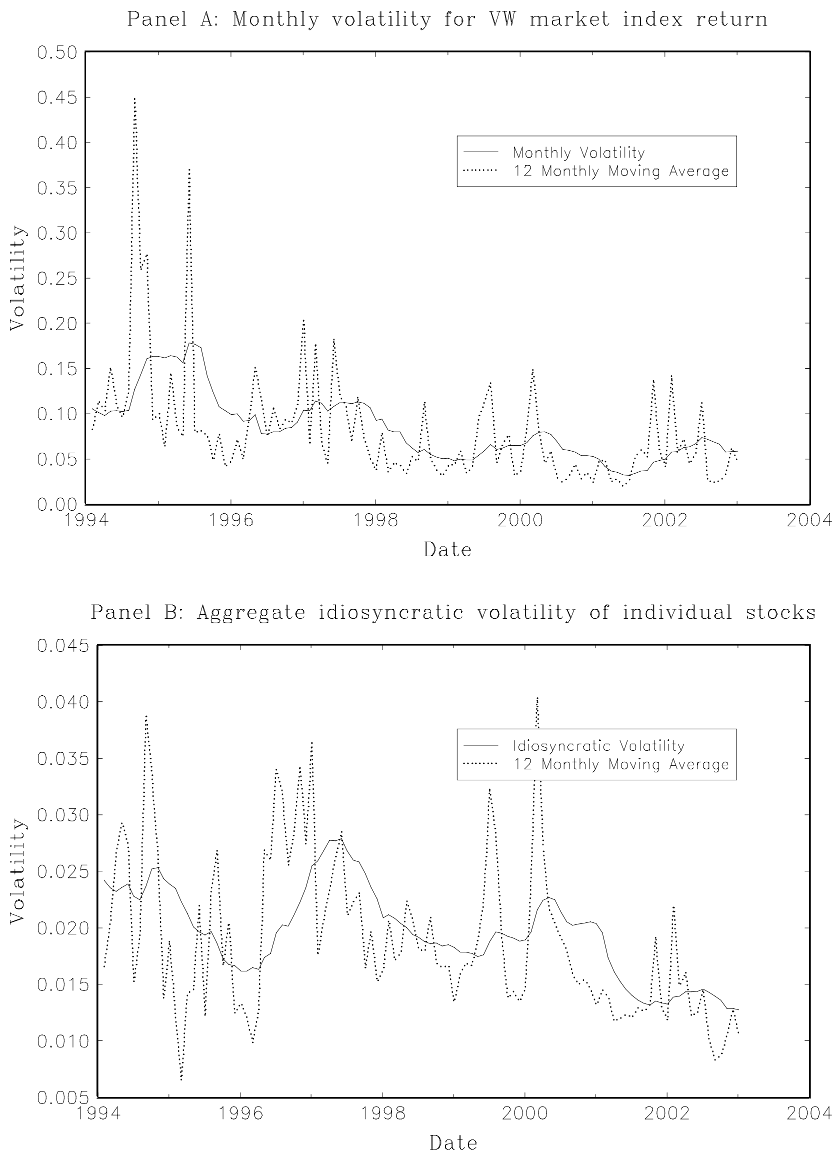


Figure 3. Correlations Among Individual Stock Returns and Residual Returns

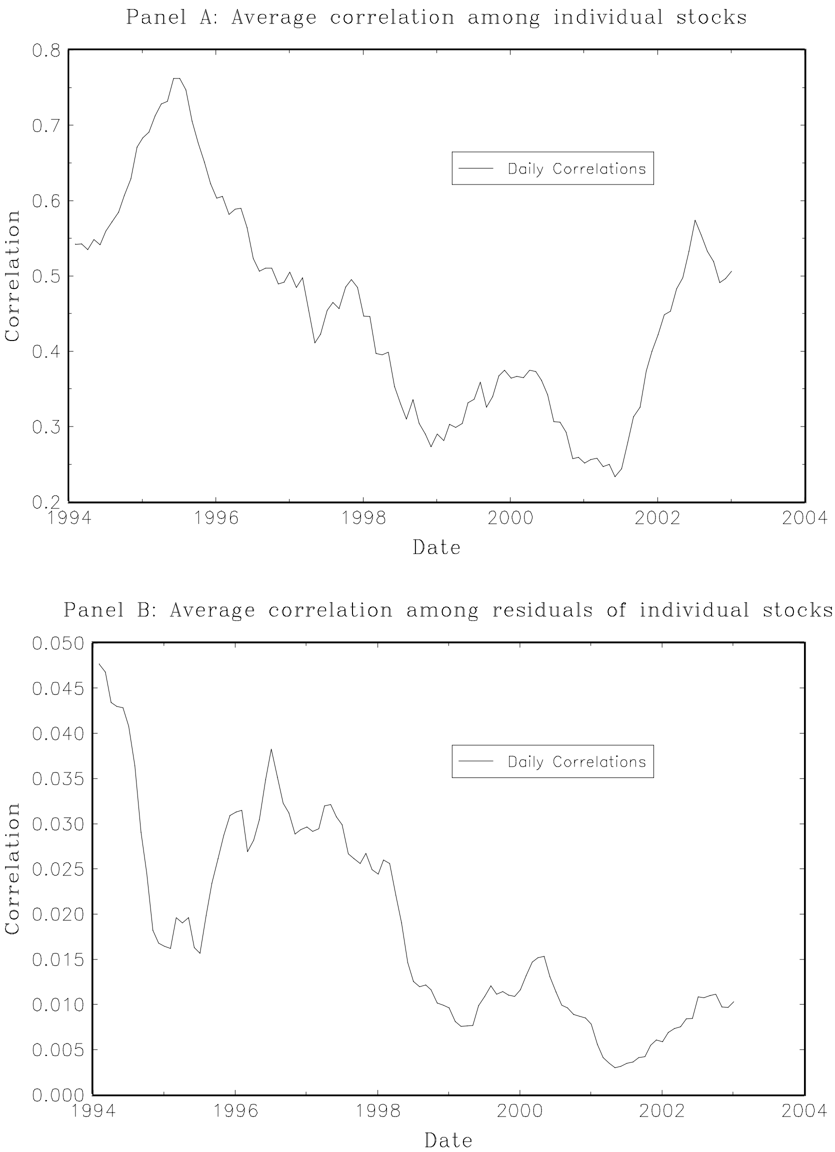
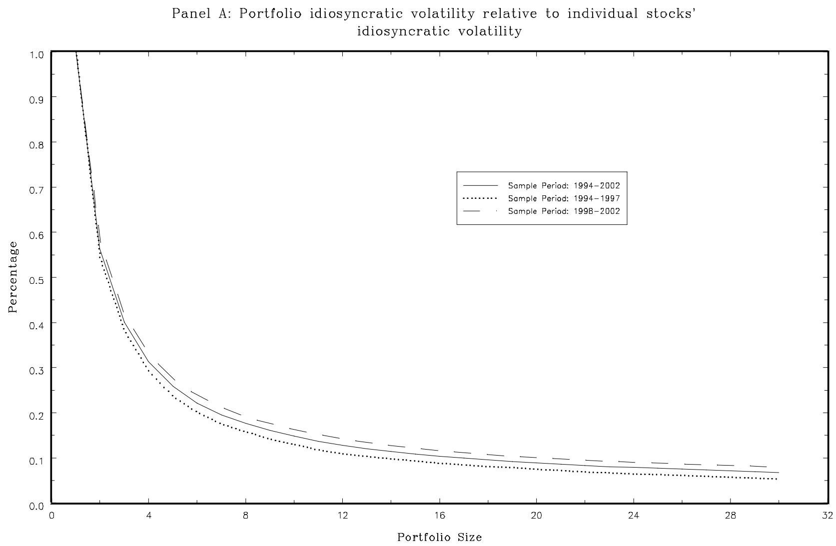


Figure 4. Diversification Graph



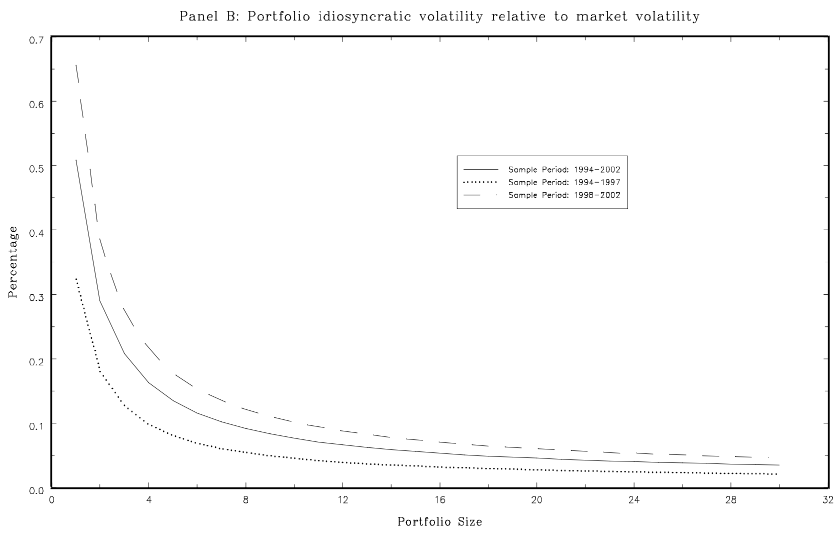


Figure 5. Diversification At Different Percentile

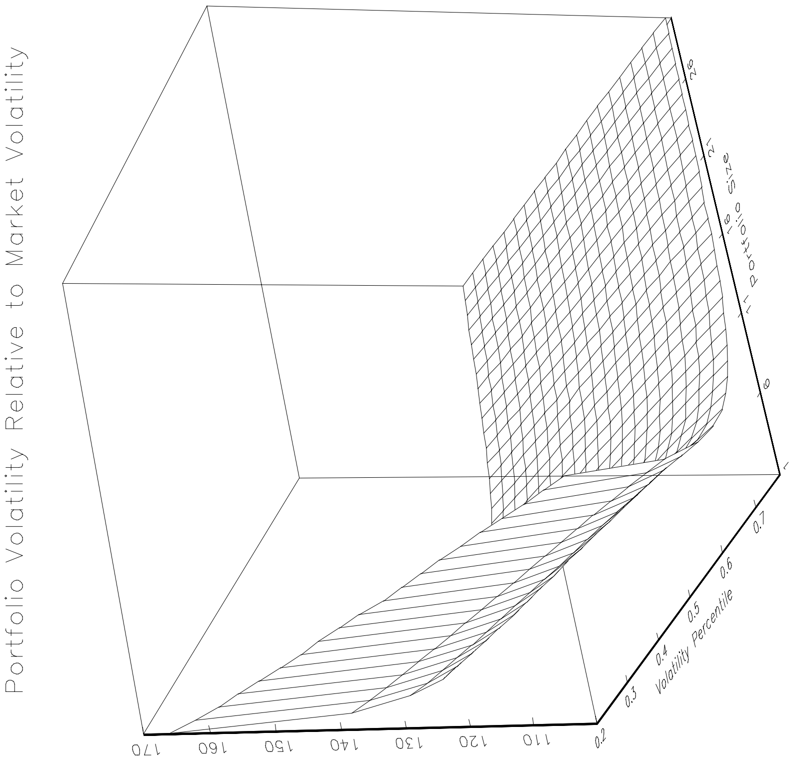


Figure 6. Excess Returns for Different Portfolio Size

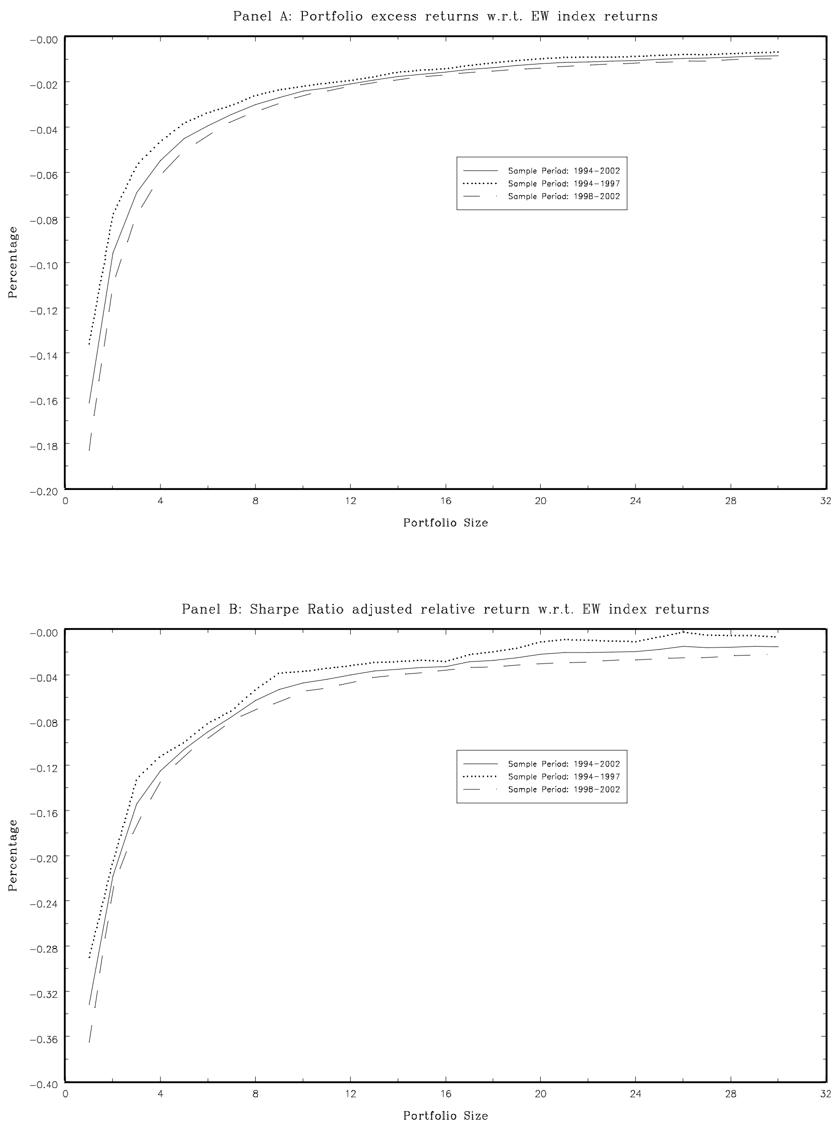


Figure 7. Return Likelihood for Different Portfolio Size

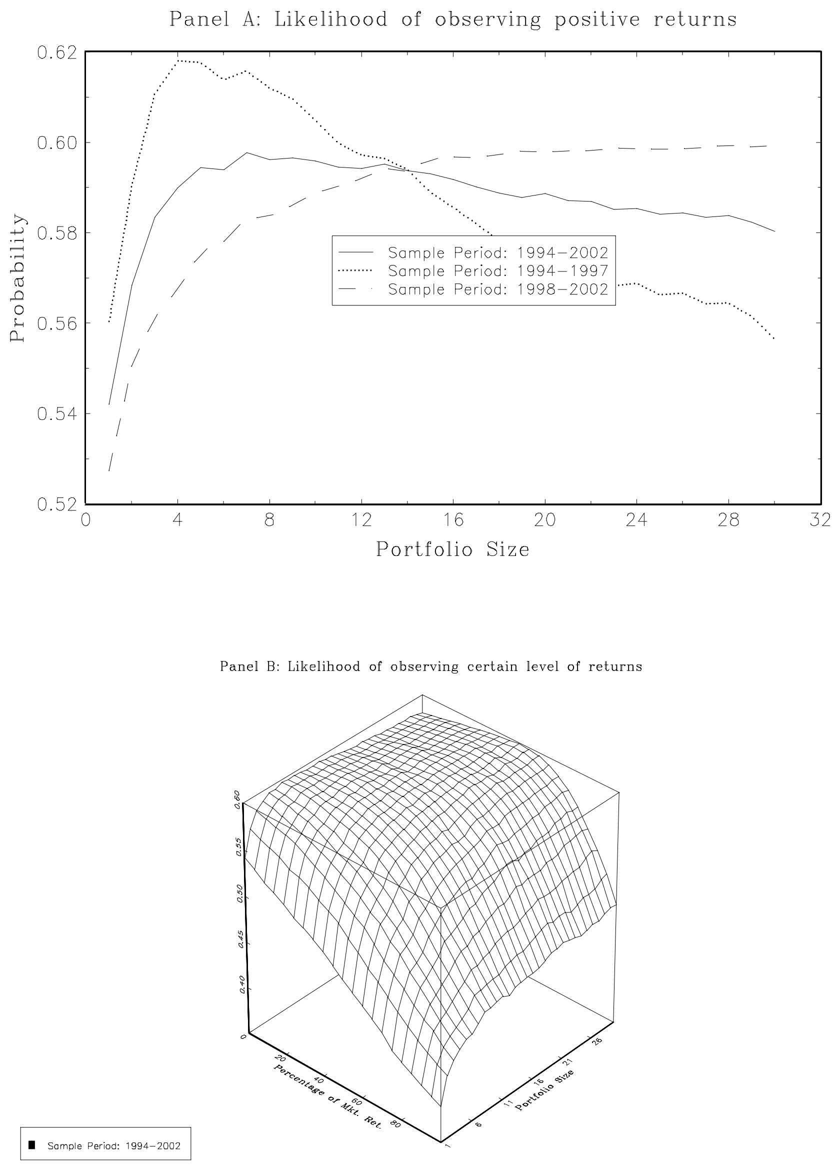


Figure 8. Portfolio Turnover and Return Likelihood