Disentangling Liquidity and Size Effects in Stock Returns: Evidence from China*

Rong Cui† University of Vienna

Youchang Wu‡ University of Vienna

This revision: February 2007

Abstract

Both size effect and liquidity effect in stock returns have been documented extensively in the literature. Since firm size and stock liquidity are closely correlated, it is generally difficult to disentangle these two effects. The Chinese stock market offers a unique opportunity to address this issue due to the widespread coexistence of tradable and nontradable shares in listed companies. We find that the size effect of nontradable shares on expected returns persists after controlling for several common liquidity measures as well as the size of tradable shares. However, it is substantially weaker than the size effect of tradable shares. Our results suggest that the negative relation between stock returns and firm size, as normally measured by the size of tradable shares, is a mixture of a genuine size effect and a liquidity effect. We also find that despite many special features of the market environment, Chinese stock returns exhibit cross-sectional patterns strikingly similar to those documented for developed markets.

JEL codes: G12, G14, G15

Keywords: Size effect, liquidity effect, stock returns

---

*We thank Lorenzo Garlappi, Gabriel Lee, Vasant Naik, Andreas Reschreiter, Alex Stomper, Suresh Sundaresan, Josef Zechner, and seminar participants at University of Vienna, Vienna Institute for Advanced Studies, Vienna Institute for Strategic Capital Market Research, Beijing University, Hong Kong University, and European Summer Symposium in Financial Markets 2005 in Gerzensee for comments and suggestions, and thank Sinofin Information Services Company for the data. Youchang Wu gratefully acknowledges the financial support by the Gutmann Center for Portfolio Management at the University of Vienna.

†Department of Finance, University of Vienna, Bruennerstrasse 72, 1210 Vienna, Austria. Email: a0348291@unet.univie.ac.at.

‡Corresponding author. Department of Finance, University of Vienna, Bruennerstrasse 72, 1210 Vienna, Austria. Phone: 0043-1-4277-38211, Fax: 0043-1-4277-38074, Email: youchang.wu@univie.ac.at.

1 Introduction

One of the central themes in finance is what determines the expected returns of different securities. The mainstream asset pricing theory contends that the expected return of a security is mainly determined by its systematic risks. While supportive evidence has been found in earlier empirical tests (Fama and MacBeth (1973), Roll and Ross (1980)), more recent research has documented some intriguing regularities that cannot be easily reconciled with this paradigm. One particular example of such "anomalies" is the size effect: stock returns, after controlling for standard measures of systematic risks, tend to be negatively related to the firm size. This effect has been documented for both developed countries and emerging markets.[1]

Various theories have been put forth to explain the size effect. One of the most frequently mentioned explanations holds that small stocks contain some systematic risks that are not adequately measured by empirical researchers. Small firms are small because the market uses a high discount rate to capitalize its future cash flows, or because they have lost market values due to poor past performance. They are more likely to have cash flow problems and less likely to survive adverse economic conditions. Since these risks cannot be easily captured by empirical models, small stocks tend to exhibit a higher risk-adjusted return.[2] Another popular explanation for the size effect, dating back to Stoll and Whaley (1983), is based on liquidity. Larger stocks are generally more liquid, and investors are

[1]See Banz (1981) and Fama and French (1992) for results on the US market and Rouwenhorst (1999) for results on 20 emerging markets.

[2]See, for example, Chan and Chen (1991), Fama and French (1996), Berk (1995), Vassalou and Xing (2004), Berk, Green, and Naik (1999), Gomes, Kogan, and Zhang (2003).

willing to compromise returns for higher liquidity. Therefore equilibrium returns of larger stocks are lower.

The importance of liquidity on security returns has been confirmed by numerous empirical studies. A significantly positive relation has been found between expected stock returns and various measures of illiquidity such as the bid-ask spread, price impact of trading, turnover rate or trading volume.[3] In particular, Amihud and Mendelson (1986), Amihud and Mendelson (1989) find that the size effect disappears after controlling for bid-ask spread. Brennan, Chordia and Subrahmanyam (1998, 2005) and Chordia, Subrahmanyam, and Anshuman (2001) find that the size-return relation becomes either insignificant or positive after controlling for trading volume. These results suggest that the size effect may be fully explained by liquidity.

A difficulty in testing the liquidity-based explanation against the risk-based explanation lies in the fact that stock liquidity is very hard to measure, and is usually inextricably correlated with firm size. In fact, the market capitalization of equity, a standard measure of firm size, is also frequently used as a measure of stock liquidity (see Kluger and Stephan (1997)). As a result, it is very difficult to disentangle the genuine size effect and the liquidity effect.

The Chinese stock market offers a unique opportunity to examine this issue due to the special ownership structure of the listed companies. Typically only a fraction of a Chinese listed company's outstanding stocks are freely tradable on the stock exchange. This creates a wedge between firm size and stock liquidity. While the liquidity of a stock is closely related to the size of tradable shares, it is not directly related to the size of nontradable

[3]See for example, Amihud and Mendelson (1986), Brennan and Subrahmanyam (1996), Datar, Naik, and Radcliffe (1998), Chordia, Subrahmanyam, and Anshuman (2001), Pastor and Stambaugh (2003). A notable exception is Spiegel and Wang (2005), who find that the liquidity effect is weak after controlling for idiosyncratic risk.

shares. Therefore, if the size effect solely comes from difference in liquidity, then we may not find any relation between returns and the size of nontradable shares. By contrast, if the size effect arises from some omitted risk factors associated with the firm's operations or cash flows, as is typically argued by most risk-based explanations, then what really matters is the size of the whole firm. As a result, the sizes of both tradable shares and nontradable shares should be equally related to expected returns, and the relative weights of these two types of shares should be irrelevant. In between these two extreme cases, if both risk and liquidity contribute to the size effect, then both size measures should matter, but the size of tradable shares should have a stronger effect due to its close relation with stock liquidity. This is the key observation that guides our empirical tests on the sources of size effect.

From the statistical point of view, the various liquidity measures used in the literature are often highly correlated with the size of tradable shares, this limits the power of empirical tests in separating the size and liquidity effects. However, they are less correlated with the size of nontradable shares. Therefore, if there is a genuine size effect independent of liquidity, it will be easier to detect it by examining the relation between the size of nontradable shares and expected returns.

Besides the size effect, many other stock return "anomalies" have been reported in the literature. For example, a number of studies have found that value stocks (stocks with high ratios of book equity to market equity, earnings to price, and cash flow to price) tend to outperform growth stocks.[4] Furthermore, some striking time series patterns of stock returns have been uncovered, namely, the relative performance of stocks tends to reverse in the short-term (contrarian effect), continue in the intermediate-term (momentum effect),

[4] See Fama and French (1992, 1998).

and reverse again in the long-term (long-term reversal).[5] These previous findings provide important guidance on the design of our empirical tests.

Using monthly individual stock data from 1993 to 2004 and adopting the procedure of Brennan, Chordia, and Subrahmanyam (1998), we first examine the basic cross-sectional patterns of Chinese stock returns. Given the distinct features of the Chinese stock market, analyzing this issue is not only an interesting exercise by itself, but also an important step towards identifying proper model specifications for our tests of the size effect. Surprisingly, we find some striking similarities between the cross-sectional patterns of Chinese stock returns and those documented for developed markets. More specifically, we find a negative size effect and turnover rate effect, as well as a positive book-to-market effect. We also find a strong short-term contrarian effect and a significant intermediate-term momentum effect. These results differ substantially from results of several recent studies based on portfolio returns.

More importantly, our paper provides valuable insights into the sources of the size effect by examining the correlations between expected returns and the sizes of tradable and nontradable shares. We find the negative correlation between expected returns and the size of nontradable shares persists after controlling for common liquidity measures such as turnover rate, trading volume, liquidity ratio, as well as the size of tradable shares. Therefor, in contrast to the previous findings of no independent size effect, our results suggest that there exists a genuine size effect that cannot be attributed to liquidity.

While our results do not support the view that the size effect is simply a liquidity effect, we also find evidence suggesting an important role played by liquidity in the size effect.

[5] See Jegadeesh (1990), Jegadeesh and Titman (1993), and DeBondt and Thaler (1985).

Consistent with the conjecture that the size-return relation is strengthened when size is linked to liquidity, we find that the size effect of tradable shares is substantially stronger than that of nontradable shares. Furthermore, given the size of the whole firm, firms with a larger fraction of tradable equity offer a significantly lower stock return. Using the size effect of nontradable shares as a measure of the genuine size effect, our point estimates indicate that more than half of the negative relation between stock returns and firm size measured by tradable shares can be attributed to liquidity.

The rest of this paper is structured as follows. Section 2 briefly reviews the institutional background of the Chinese stock market and describes the data. Section 3 describes the methodology. Section 4 presents the basic cross-sectional patterns of expected stock returns. Section 5 tests whether there is a genuine size effect independent of liquidity and evaluate the contribution of liquidity to the size effect. Section 6 concludes.

2 Institutional background and data description

2.1 Institutions

The Chinese stock market has a relatively short history but has enjoyed rapid growth. When the two stock exchanges, one in Shanghai and the other in Shenzhen, commenced operations in December 1990 and July 1991 respectively, there were all together only 14 traded stocks. By the end of 2005, 1381 stocks have been listed on these two exchanges. The market capitalization of all listed companies stands at RMB yuan 3.2 trillion (including the imputed value of nontradable stocks), which is around US dollars 0.4 trillion. The number of registered investor accounts exceeds 7.3 million.[6]

One special feature of the Chinese stock markets is the variety of types of stocks issued by the listed companies. Shares are classified as domestic (A-shares) and foreign (B-shares) by shareholders' country of residence. Most listed companies issue only A-shares. 90 of them issue both A shares and B-shares.[7] Historically, A-shares were only accessible to domestic investors, while B-shares were only available to foreign investors. As a step towards a more integrated and more open capital market, the B-share market was opened to domestic investors in February 2001, which resulted in a sharp increase of B-share prices. Furthermore, limited access to the A-share market has been given to a number of "qualified foreign institutional investors" (so-called QFII) starting from 2003. We focus on the A-share market because it is by far the most important stock market in China.

The A-shares issued by a typical listed company are further classified into three subcategories: (1) state shares, which are held by the central government, local governments, or solely state-owned enterprises; (2) legal person shares, which are owned by other domestic institutions; and (3) tradable A-shares. Among these three subcategories only the last type of stocks are publicly tradable on the exchanges. State shares and legal person shares can only be transferred to institutions through private negotiations or occasionally organized auctions. They are usually referred to as non-tradable shares. On average they constitute about two thirds of the total shares outstanding.

While the coexistence of tradable and non-tradable shares provides us with some unique research opportunities, it has long been regarded as one of the most fundamental structural

[6]Cited from the website of China Securities Regulatory Commission: www.csrc.gov.cn.

[7]31 listed companies issue also H-shares, which are traded on Hongkong Stock Exchange.

problems that hinders the development of Chinese stock market. Various proposals to end this situation have been hotly debated, and a series of attempts have been made by the regulatory agencies. In particular, a plan to allow the state shares to be sold gradually to the public at the market price of tradable shares was announced in 2001. The market responded with a dramatic decline in fear of a looming supply overhang, which eventually forced the government to withdraw the plan after a few months. A new reform agenda was launched in 2005, which allows each listed company to come up with its own solution through negotiations between shareholders. This program proves to be quite successful. By the end of 2006, shareholders in most listed companies have worked out an agreement on how to convert nontradable into tradable shares. As a result, nontradable shares will soon become a historical name in the Chinese stock market.

Apart from the distinct features mentioned above, the Chinese stock market also shares some comment characteristics of other emerging markets. For example, intervention by government is extensive; regulatory environment keeps changing over time; shareholder protection is inadequate; information is not transparent and not readily available; individual investors with limited financial knowledge abound, while institutional investors are still in early stage of development.

2.2 Data and summary statistics

We obtain our data from Sinofin Information Services Company. The data set includes daily trading data for all A-shares from January 1993 to December 2004. It also includes the annual financial statements of all listed companies since 1994.

Using this database, we first construct a value-weighted monthly return index for the A-share market. We do so because the existing stock indexes are not directly useable for our purpose. For example, the two most frequently used stock indexes, i.e., Shanghai Composite Stock Index and Shenzhen Composite Stock Index, have the following drawbacks. First, each of them covers only stocks listed on one of the two exchanges. Second, they weight each stock's returns by the issuing company's total equity instead of the market value of tradable A-shares, therefore they do not accurately reflect the movements of a tradable market portfolio. Third, they do not incorporate dividends. Our return index is free from these problems. It covers all the A-shares traded on both exchanges. Each stock's monthly returns are weighted by the market value of tradable A-shares at the beginning of the month. Dividends are assumed to be reinvested on the ex-dividend date.

We then exclude the following stocks/observations from the sample: (1) stocks of financial firms; (2) stocks whose IPOs take place after 2000; (3) the IPO month of each stock; (4) stocks that are designated as "PT" ("Particular Transfer") stocks, which are stocks of companies that have reported losses for three consecutive years. Financial stocks are usually excluded in empirical asset pricing tests due to their special characteristics. Stocks going public after 2000 are excluded because they do not have enough monthly return observations to estimate the betas. The IPO months are excluded because returns and trading activities in the IPO month are substantially different from those in normal months. Due to strict regulations on IPO pricing, China has seen the highest initial returns of IPOs in the world.[8] PT stocks are excluded because they are subject to special trading restrictions and are traded only on Friday. We take the above sampling procedures to make our tests as clean as possible. However, detailed robust checks show that our results do not depend on

[8] Liu (2003) documents an average market-adjusted initial return of 132.49% for 354 IPOs from 1999 to 2002. Datar and Mao (1998) documents an average initial return of 388.0% for 226 IPOs from 1990 to 1996.

these procedures. The final sample consists of 1053 stocks. Each stock on average has 90 monthly return observations.

Table 1 presents descriptive statistics for our final sample. The number of stock increases rapidly from 155 in 1993 to 1053 in 2000. It decreases slightly after 2001 since a few stocks have become "PT" stocks or have been delisted. The average returns are positive in the first eight years of our sample except in 1995. Ironically, the market goes down for four consecutive years since 2001, despite a remarkable performance of Chinese economy during the same time period. The average market value of tradable shares peaks at RMB yuan 1439 million in 2000, then goes down as the stock price declines, while the average book equity per firm increases steadily from RMB yuan 541 million in 1994 to RMB yuan 1193 million in 2004. The average floating ratio, defined as the ratio of tradable A-shares to total shares outstanding, although gradually increasing over time, remains below 40 percent even at the end of 2004, indicating that the controlling stakes of most Chinese listed companies are not freely tradable.[9]

Table 1 also indicates that the turnover rate of tradable shares is rather high. In the first eight years of our sample, the average monthly turnover rate never falls below 1/3, implying that the average holding period is less then three months. This may have to do with the presence of a large number of individual investors who are interested in short-term speculative gains instead of long-term investment. Although it declines substantially after 2000, the annual turnover rates in the following four years are still above 2. At the first glance, the high turnover rate seems to indicate that the level of stock liquidity is high in China and illiquidity may be of a less concern. However, as the model of Amihud and

[9] The actual ratio of floating shares, after considering tradeable B-shares and H-shares, will be higher, but only slightly, since only less than ten percent of firms issue such shares.

Mendelson (1986) suggests, liquidity has a much stronger impact on short-term investors than on long-term investors. Therefore the demand for liquidity may be even stronger in a market dominated by short-term investors, leading to a larger liquidity effect. Furthermore, most Chinese listed companies, especially when measured by the size of tradable shares, are relatively small compared to the listed companies in developed market. This can make the price impact of stock trading a serious concern for large investors.

Another notable feature showing up in Table 1 is that the book-to-market ratio is rather low compared to the ratios normally observed in developed markets. The average ratio ranges from a bottom of 0.177 in 2000 to a peak of 0.538 in 1995, while the average ratio for the US stocks is 1.25, as reported by Brennan, Chordia, and Subrahmanyam (1998) for their sample. This indicates that a large fraction of the market capitalization of Chinese listed companies comes from anticipated growth potentials.

For our cross-sectional analysis, we calculate for each stock and each month t the following potential explanatory variables of expected returns:

- ME_A: the market value of tradable A-shares at the end of month $t-1$.
- BE: the total book value of the firm's equity. Since the market value nontradable equity is unobservable, we have to use the book value of total equity to measure the size of the whole firm.
- FR: the floating ratio, defined as the ratio of tradable A-shares to total shares outstanding, at the end of month $t-1$.
- BE_A: the book value of tradable A-shares, calculated by multiplying the total book equity by the floating ratio.

- BE_N: the book value of non-tradable shares, calculated by multiplying the total book equity by 1 minus the floating ratio.

- BM: the book-to-market equity ratio, defined as the ratio of book equity per share to the price of A-shares.

- TO: the average monthly turnover, defined as the amount of A-shares traded as a fraction of total tradable A-shares outstanding, in months $t-1$ through $t-12$.

- TV: the average monthly trading volume, measured in unit of currency, in months $t-1$ through $t-12$.

- LR: the average monthly liquidity ratio in months $t-1$ through $t-12$. The liquidity ratio in each month is calculated by dividing the total monthly trading volume (in unit of currency) by the sum of the absolute values of all daily returns in that month. This ratio has been used by many authors to measure the price impact of order flow, an important dimension of stock liquidity.[10]

- $R1$: the stock return in month $t-1$.

- $R2$: the stock return in month $t-2$.

- $R12$: the average monthly stock return in months $t-3$ through $t-12$.

Note that there is no any explicit measure of risks on our list of explanatory variables. This is because the dependent variable in our cross-sectional regressions will be the risk-adjusted return, as will be explained in the next section. Following Fama and French (1992)

[10]See for example, Kluger and Stephan (1997), Amihud, Mendelson, and Lauterbach (1997).

and Brennan, Chordia, and Subrahmanyam (1998), for all variables related to accounting information, i.e., BE, BE_A, BE_N, and BM, we adopt the following convention: from July of year i to June of year $i+1$, these variables are calculated using book values (and market price) at the end of year $i-1$. This allows for the delay in accounting disclosure and ensures the availability of such information to investors. We take natural logarithm of all explanatory variables except the lagged returns to reduce skewness of those variables. Furthermore, we transform all explanatory variables into deviations from the cross-sectional means in each month, therefore the average stock has a value of zero for all explanatory variables.

Table 2 presents the time-series averages of the monthly cross-sectional correlations between various transformed stock characteristics. The return is negatively correlated with the market value of tradable equity (MV_A), the trading volume (TV), and liquidity ratio (LR), and positively correlated with the book-to-market ratio (BM). But none of these correlations is higher than 5% in absolute value. As expected, there is a high correlation between the four size measures: MV_A, BE_A, BE_N, and BE. In particular, the correlation between BE_N and BE is as high as 0.974. Using the total book equity, BE, as a measure of the firm size, the correlation matrix suggests that larger firms tend to have a lower floating ratio (with a correlation of -0.290), a higher book-to-market ratio (correlation o.553), a lower turnover rate (correlation -0.238), a higher trading volume (correlation 0.426) and a higher liquidity ratio (correlation 0.485). The correlations between the three liquidity measures — the turnover rate (TO), the trading volume (TV), and the liquidity ratio (LR) — are highly correlated with each other. The latter two liquidity measures are also highly correlated with the market value of tradable equity (correlation 0.779 between TV and ME_A, 0.855 between LR and ME_A), which makes it difficult to disentangle size effect and liquidity effect in stock returns. However their correlations with the book value of nontradable equity (BE_N) are substantially lower (0.336 between TV and BE_N, 0.390 between LR and BE_N). This suggests that if there is a size effect independent of liquidity, it will be easier to detect it when the firm size is measured by the size of nontradable equity.

3 Methodology

A traditional method of empirical research on asset pricing is to form portfolios by sorting stocks on the characteristics of interest and test hypotheses using the portfolio returns. The advantage of this approach is that it allows more accurate measurement of factor loadings, thus mitigating the errors-in-variables bias. However, it may lead to other biases. For example, Roll (1977) points out that it may conceal possibly return relevant security characteristics within portfolio averages and make it difficult to reject the null hypothesis of no effect on security returns. On the contrary, Lo and MacKinlay (1990) argue that this approach may lead to reject the null hypothesis too often due to a "data-snooping" bias. Brennan, Chordia, and Subrahmanyam (2005) highlight the potential biases of this approach by showing that the statistical significance of security characteristics in explaining expected returns is dependent on the way in which the portfolios are formed. As an alternative, they develop a method that allows the use of individual stock data while keeping the errors-in-variables bias under control. We adopt their procedure in this study.

The key innovation in the Brennan-Chordia-Subrahmanyam approach is to use the risk-adjusted return instead of the return itself as the dependent variable in the cross-sectional regression. The starting point (null hypothesis) is a L-factor APT model. Assume that returns are generated by a L-factor model:

$$R_{jt} = E(R_{jt}) + \sum_{k=1}^{L} \beta_{jk} f_{kt} + \varepsilon_{jt} \tag{1}$$

where R_{jt} is the return of stock j, β_{jk} is stock j's loading on factor f_k. Then the APT model implies that the expected excess returns can be written as

$$E(R_{jt}) - R_{ft} = \sum_{k=1}^{L} \beta_{jk} \lambda_{kt}, \tag{2}$$

where R_{ft} is the return on the riskless asset and λ_k is the risk premium of factor f_k. Substituting Equation (2) into (1), we have

$$R_{jt} - R_{ft} = \sum_{k=1}^{L} \beta_{jk} F_{kt} + \varepsilon_{jt}, \tag{3}$$

where $F_{kt} \equiv \lambda_{kt} + f_{kt}$ is the sum of the realization and risk premium of factor f_k.

A standard application of Fama and MacBeth (1973) consists of regressing excess returns jointly on factor loadings β and various stock characteristics. This procedure suffers from the errors-in-variables bias since the independent variables β can not be accurately measured. Brennan, Chordia, and Subrahmanyam (1998) bypass this problem by impounding the errors in factor loadings in the dependent variable. They first calculate the risk-adjusted return of each stock in each month t:

$$R^*_{jt} \equiv R_{jt} - R_{ft} - \sum_{k=1}^{L} \hat{\beta}_{jk} F_{kt}, \tag{4}$$

where $\hat{\beta}$ is the factor loadings estimated from historical data. They then run the month-by-month cross-sectional regressions of the risk-adjusted return R_t^* on various stock characteristics that are observable by the end of month $t-1$, i.e., they estimate the following equation:

$$R_{jt}^* = c_0 + \sum_{m=1}^{M} c_m Z_{mj,t-1} + e_{jt}, \tag{5}$$

where $Z_{mj,t-1}$ is the value of characteristic m for stock j observable in month $t-1$.

The advantage of this alternative approach is that the errors in factor loadings are impounded in the dependent variable. As a result, the estimated coefficients will be unbiased as long as errors in estimated loadings are uncorrelated with stock characteristics. Although the loadings are generally estimated with errors, there is no *a priori* reason to believe that those errors will be correlated with stock characteristics. Therefore there is no *a priori* reason to believe that estimated coefficients will be biased.

Brennan, Chordia, and Subrahmanyam (1998) use two alternative estimators to obtain the overall estimate of the coefficient of each characteristic $Z_m (m = 0, 1, ..., M)$. The first is the Fama-MacBeth estimator, i.e., the time series average of the coefficients obtained in the month-by-month estimation of Equation (5). This estimator is unbiased as long as the errors are not correlated with stock characteristics. The statistical inference can then be made based on the standard Fama-MacBeth t-statistic, i.e., the time series average of the coefficient estimates divided by the time series standard error. The second is the purged estimator first introduced by Black and Scholes (1974). For each characteristic, it is obtained as the intercept from the time series regression of the coefficients estimated in each month on the factor realizations, i.e., c_m^p in the following equation:

$$\hat{c}_{mt} = c_m^p + \sum_{k=1}^{L} d_{mk} F_{kt} + u_{mt}, \tag{6}$$

where $\hat{c}_{mt}$ is the coefficient of characteristic $Z_m (m = 0, 1, ..., M)$ estimated in month t. It can be shown that as long as the factor realizations are serially uncorrelated, the intercept estimated in this time series regression is an unbiased estimator of c_m in Equation (5) even if the errors in estimated loadings are correlated with stock characteristics.[11] The t-statistics obtained in this time series regression can be used to make statistical inference.

While Brennan, Chordia, and Subrahmanyam (1998) use the Fama-French three-factor model and Connor and Korajczyk (1988) five-factor model to calculate the risk-adjusted return, we use the simplest factor model, i.e., the market model, for most of our analysis. Our null hypothesis is that the expected excess return of stock j is determined solely by its loading on the market factor. This is a natural choice given the fact that little is known about the factor structure of Chinese stock returns.

To test the ability of the widely-used Fama-French three-factor model in characterizing Chinese stock returns, for the basic model specification presented in the next section, we also report results using the Fama-French model as the benchmark. If equilibrium stock returns are fully described by the Fama-French factors, returns adjusted by these factors

[11] See Brennan, Chordia, and Subrahmanyam (1998). The intuition behind the purged estimator is straightforward. If, for example, the kth loading of firms with a high Z_m is systematically under-estimated, then the risk-adjusted return of firms with a high Z_m will be over-estimated in the periods with a positive realization of F_k, and vice versa. Accordingly the estimated $\hat{c}_{mt}$ will be positively correlated with F_{kt}. On average, c_m will be over-estimated by an amount proportional to the risk premium of factor F_k. Running a regression of $\hat{c}_{mt}$ on factor realizations eliminates the F_k-dependent component in $\hat{c}_m$ and yields an unbiased of c_m as the intercept.

should be uncorrelated to any stock characteristics. To implement such a test, we construct two risk factors in addition to the market risk factor following Fama and French (1993), one as a proxy for the size-related systematic risk and the other for the systematic risk related to the book-to-market equity ratio. At the beginning of July of each year, all stocks in our sample are ranked based on the market value of the tradable A-shares at the end of June and grouped into two equally weighted portfolios (Small and Big) with the same number of stocks. Each group again is divided into three sub-portfolios (High, Medium, Low) based on the book-to-market ratio at the end of last December,[12] with the High and Low groups each containing about 30% of stocks. The size-related risk factor, SMB, is then calculated as the difference between the returns on the portfolio Small and the portfolio Big. The book-to-market-ratio-related risk factor, HML, is the difference between the simple averages of returns on the two high book-to-market sub-portfolios and the two low book-to-market sub-portfolios. Returns on each individual portfolio are value-weighted.

The factor loadings for month t are estimated using data of month $t-1$ though $t-48$.[13] To allow for thin trading, we follow the procedure of Dimson (1979) to estimate the betas, i.e., we regress excess returns of each stock on both current and lagged factor realizations and use the sum of the two slopes as the stock's betas. If a stock has missing return data for more than 12 of the 48 past months, then its betas for the current month are set as missing. We use the three-month bank deposit rate to compute the monthly risk-free rate. After estimating all the betas, we calculate the risk-adjusted return according to Equation (2) and run the month-by-month cross-sectional regressions.

[12] For years 1993 and 1994, we use the book-to-market ratio in 1994 to classify the stocks.

[13] A previous version of this paper use 36 months of returns to estimated the loadings and yields similar results.

Following Brennan, Chordia, and Subrahmanyam (1998), for the cross-section of the risk-adjusted return, we report both results based on the Fama-MacBeth estimator and on the Black-Scholes purged estimator. For comparison, we also report results from the traditional Fama and MacBeth (1973) regression using the unadjusted excess return as the dependent variable. To account for potential autocorrelations and heteroscedasticities, all the t statistics are computed using Newey and West (1987) corrected standard errors.

4 The cross-section of Chinese stock returns

4.1 The main determinants of expected stock returns

We start our empirical analysis with an examination of the cross-section of Chinese stock returns. In the light of the stylized facts found on other markets as reviewed in the introduction, we construct a basic regression model that contains the following stock characteristics as explanatory variables for expected returns: the market value of tradable shares (MV_A), the book-to-market equity (BM), the turnover rate (TO), and three lagged returns ($R1$, $R2$, and $R12$).

Table 3 presents the results of the basic model. The first column reports the results when returns are not adjusted for risks, the next two columns report the results when returns are adjusted by the market factor, while the last two columns report results when returns are adjusted by the Fama-French factors. The coefficients in columns labeled "FM" are estimated using the Fama-MacBeth estimator while those in columns labeled "BS" are estimated using the Black-Scholes purged estimator. The results are based on 96 month-by-month cross-sectional regressions, covering the period from January 1997 to December

2002. The cross-sectional regressions start in January 1997 since the first four years are used to estimate factor exposures.

The results are rather striking. The Chinese expected stock returns show patterns very similar to those documented for developed markets, in particular, the patterns reported by Brennan, Chordia, and Subrahmanyam (1998) for the US. Except for the last column, results reported in different columns are very consistent with each other. First, there is a negative size effect and a positive book-to-market effect, as indicated by the coefficients of ME_A and BM. Furthermore, our results also show a negative turnover rate effect, as indicated by the negative coefficients of TO; a short-term contrarian effect, as indicated by the negative coefficients of $R1$ and $R2$; and a positive intermediate-term momentum effect, as indicated by the positive coefficients of $R12$. These results differ significantly from results of several recent studies on the Chinese stock market based on portfolio returns. For example, Wang and Xu (2005) and Bailey, Cai, Cheung, and Zheng (2003)) find no turnover rate effect and no serial correlations. Wang and Xu (2005) also report an insignificant book-to-market effect.

The Fama-MacBeth estimator and the Black-Scholes estimator yield almost identical results for returns adjusted by the one-factor model, suggesting that the measurement errors in the loadings to the market factor is uncorrelated with stock characteristics. However, after the risk-adjustment using the Fama-French factors, results based on these two estimators become quite different. Based on the Fama-MacBeth estimator, it seems that the Fama-French model has very little advantage over the market model, since none of the patterns mentioned above disappears or becomes weaker. However, results based on the Black-Scholes estimator, reported in the last column, suggest that the Fama-French model does explain the book-to-market effect and a substantial part of the size effect. The book-to-market effect becomes insignificant, the size effect, while still significant at the level of 10%, become substantially weaker in magnitude: the coefficient of ME_A changes from -0.007 in columns two and three to -0.003 in column five. The differences between the results from the Fama-MacBeth estimator and those from the Black-Scholes estimator suggest that the measurement errors in the exposures to the size and book-to-market factors may be correlated with stock characteristics.

The significant relations between expected returns and various stock characteristics documented above not only reject the null hypothesis that expected returns are solely determined by exposures to the market factor, but also question the validity of the Fama-French three factor model in describing the equilibrium returns of Chinese stocks. In particular, the Fama-French model fails to fully explain the size effect, suggesting that non-risk factors such as liquidity may play a role in this effect.

In summary, stock returns in China seem to follow very similar cross-sectional patterns as stock returns in the US. This result not only adds further evidence for the generality of those previously-uncovered patterns, but also suggests that any new finding we may get from China may be relevant for other markets as well.

4.2 Economic importance of the cross-sectional patterns

To assess the economic importance of the cross-sectional patterns we document above, we examine the profitability of a hypothetical portfolio strategy based-on our basic model, following the procedure of Haugen and Baker (1996). In each month t, we forecast the expected risk-adjusted return of each stock in month $t+1$ based on current stock characteristics and the cross-sectional relations between stock characteristics and risk-adjusted returns estimated from historical data. We then rank the stocks by their expected risk-adjusted returns and form them into ten equally-weighted deciles, with decile 1 containing the stocks with the lowest expected risk-adjusted returns.

When forecasting the expected risk-adjusted return, we use the 12-month trailing averages of the coefficients estimated from the cross-sectional regressions. For example, the expected risk-adjusted return of stock j in month $t+1$ is given by:

$$E(r^*_{j,t+1}) = \overline{c}_{0t} + \sum_{m=1}^{m=M} \overline{c}_{mt} Z_{mjt}, \tag{7}$$

where $\overline{c}_{mt}(m = 0, 1, ..., M)$ is the average of the coefficients estimated from Equation (5) in months t through $t-11$, and Z_{mjt} is the characteristic Z_m of stock j observed at the end of month t.

Table 4 reports the average realized returns of each decile from 1998 to 2004 when the forecast is based on the cross-sectional relations between stock characteristics and returns adjusted by the market factor. Panel A presents the unadjusted realized returns. Panel B presents the realized returns adjusted by the market factor. It can be seen that the out-of-sample predictability of our basic model is rather high. Both the realized unadjusted return and the realized risk-adjusted return increase steadily from decile 1 to decile 10, both in the bullish market of 1998 to 2000 and in the bearish market of 2001 and 2004. A hypothetical strategy that shorts the stocks in decile 1 and longs the stocks in deciles 10 would yield an unadjusted return of 2.17% per month over our sample period. In terms of market-factor-adjusted returns, the performance of such a strategy would be 2.33% per month.[14]

Table 5 reports the same set of results when the forecast is based on the cross-sectional relations between stock characteristics and returns adjusted by the Fama-French factors. It can be seen that the year-by-year forecasts are less reliable. However, a hypothetical strategy that shorts the stocks in decile 1 and longs the stocks in deciles 10 would still yield an unadjusted return of 1.62% per month, and a risk-adjusted return of 1.84% per month, over our sample period. While these numbers are smaller than in the case of the market factor model, the profit of the characteristics-based is still sizable, indicating that the explanatory power of Fama-French model is very limited.

Although such a hypothetical portfolio strategy is not directly implementable since short sales are not allowed in the Chinese stock market, and transaction costs will reduce its desirability even if it is implementable, the substantial abnormal performance of our characteristics-based strategy indicates that the cross-sectional patterns we find in this paper is economically very important.

5 Liquidity and the size effect

In the last section we show a strong and negative relation between the size of tradable shares and expected returns. However, since the size of tradable shares measures both the firm size and stock liquidity, it is not clear whether this negative relation is due to liquidity or other reasons. We now test the liquidity-based explanation for the size effect against the risk-

[14]The risk-adjusted return is higher than the unadjusted return since our strategy maximizes the risk-adjusted return.

based explanation. Since we wish to start with the "raw" size effect rather than impose a strong structure such as the Fama-French model from the beginning, in the following tests we focus only on returns adjusted by the market factor model.

5.1 Size effect independent of liquidity

The liquidity of an asset is defined as the degree of ease with which this asset can be traded. While various costs of trading a stock, such as the cost of searching for a counterparty, and the price concession in executing a trade, may be negatively related the size of tradable shares, they are not directly related to the size of nontradable shares. Therefore, if the size-return relation is solely due to the difference in liquidity between large and small stocks, then the size of nontradable shares may not have any incremental explanatory power for expected returns. By contrast, according to the risk-based explanation of the size effect, the size effect arises from some hidden risk factors which are associated with the firm's operations or cash flows. Since these risk factors are correlated with the size of the whole firm, the sizes of both nontradable shares and tradable shares should matter for stock returns. We now test these alternative predictions directly by examining whether the size of non-tradable equity is significantly related to expected returns after controlling for liquidity as well as other relevant stock characteristics.

Since the market value of non-tradable equity is not observable, we use the book value to measure its size.[15] We consider three alternative model specifications, each containing a different measure of stock liquidity. The market value of tradable shares is included in all

[15] The prices of the non-tradable equities occasionally observed in negotiated block transfers are substantially below the market price of the tradable A-shares. See for example, Chen and Xiong (2001).

the models. This is important since otherwise the size of nontradable shares may capture the liquidity effect of tradable shares due to the correlation between these two size measures.

The first model is a straightforward extension of the basic model reported in Table 3, i.e., we simply add the book equity of nontradable shares, BE_N, into the regression. In this model liquidity is measured by the turnover rate and the market value of tradable shares. The results are reported in the first three columns of Table 6. As one can see, adding BE_N to the regression changes very little the point estimates and the significance of the coefficients of the original explanatory variables in the basic model. Again, the Fama-MacBeth estimator and Black-Scholes estimator yield almost identical results. The coefficients of BE_N are negative and significant at the 5% level in all the three columns. Since the size of nontradable shares itself is not directly related to stock liquidity, and since we have controlled for stock liquidity, this result suggests that the size effect is not simply a liquidity effect.

Brennan, Chordia and Subrahmanyam (1998, 2005) and Chordia, Subrahmanyam, and Anshuman (2001) find that the size-return relation becomes either insignificant or positive when liquidity is measured by trading volume instead of turnover rate. As a further test, we examine whether the size effect of nontradable shares remains significant when the turnover rate is replaced by trading volume in the regression. We keep all the other explanatory variables unchanged. The results are reported in the last three columns of Table 6. Again the three sets of regressions yield very similar results. Consistent with the results of the previous studies mentioned above, the size of tradable shares, ME_A, loses its significance. However, the coefficient of the size of nontradable shares, BE_N, remains negative at the 10% significance level. The coefficients of the trading volume are negative and highly significant. The coefficients of the other explanatory variables are largely unchanged. These results further suggest that there is a genuine size effect that cannot be explained by liquidity.

While the turnover rate and trading volume reflect how actively a stock is traded on the market, they do not directly measure the price impact of order flow, i.e., the discount that a seller concedes or the premium that a buyer pays when executing an order. A commonly-used liquidity measure that explicitly takes this into account is so-called liquidity ratio (also called Amivest ratio). It measures the trading volume associated with a 1% change in the stock price. If the price impact of order flow is high, due to for example severe information asymmetry between potential sellers and buyers, then this ratio will be low. By contrast, if order flow does not move price very much because information is relatively transparent, then this ratio will be high. Our last model tests whether the size effect of nontradable equity persists when liquidity is measured by the liquidity ratio. The results, presented in Table 7, are very similar to what we have when liquidity is measured by the trading volume. The coefficients of ME_A are insignificant and the coefficients of BE_N are significantly negative.

In summary, the size effect of nontradable shares persists after controlling for all the three common liquidity measures we consider. This is in sharp contrast with previous findings of no size effect independent of liquidity. Clearly this finding results from the increased statistical power due to the relatively low correlation between the size of nontradable shares and the liquidity measures. If we just look at the size effect of tradable shares, we will find it insignificant in two of the three models we consider. Since there is no reason to believe that the genuine size effect of the tradable shares is weaker than that of nontradable shares, the insignificance of the former is better interpreted as a result of low statistical power due to high multi-collinearity rather than evidence of inexistence of such an effect.

5.2 Contribution of liquidity to the size effect

In the last subsection we have presented evidence showing that there is a genuine size effect that cannot be attributed to liquidity, we now examine the role of liquidity in the size effect by comparing the size effects of tradable and nontradable shares. If the size effect arises solely from some hidden risk factors and has nothing to do with liquidity, then these two size effects will be equally strong, and the relative weights of tradable and nontradable shares will be irrelevant. Otherwise we would expect that the size effect of tradable is stronger because it contains both a genuine size effect and a liquidity effect.

First note that the results reported in Table 6 do show that the size effect of tradable shares is substantially stronger than that of nontradable shares: compared to the coefficients of BE_N, the coefficients of ME_A are larger in absolute value (-0.006 vs. -0.004 when returns are not adjusted for risks, -0.007 vs. -0.003 after risk-adjustment).[16] This suggests that the size-return relation can be reinforced when size is linked to liquidity.

To make sure that the difference we find above does not come from the fact that the size of nontradable shares is measured by book value while the size of tradable shares is measured by market value, we replace ME_A by BE_A and rerun the regression, i.e., we make a fairer comparison by measuring both types of shares by their book values.[17] The results are reported in the first three column of Table 8. Statistically the size effect of nontradable

[16]The results of the second and third tests are not suitable for making such a comparison since ME_A is more correlated with TV and LR than is BE_N.

[17]Note that the market value is measured with a lag of one month, while the book value is measured with a lag of 7-19 months, therefore BE_A is not simply a linear combination of ME_A and BM.

shares becomes weaker under the new specification due to the high correlation between the two book values: the coefficients of BE_N are only significant at the level of 10%. The size effect of tradable shares (measured by the coefficients of BE_A) remains to be significant at the level of 5%.

The relative magnitude of the two size effects remain almost the same: the coefficients of BE_N are less than half of the coefficients of BE_A in absolute value. If we assume that the coefficient of BE_N measures the genuine size effect, and that the coefficient of BE_A measures the joint effect of the firm size and liquidity, then more than half of the size effect of tradable shares on returns can be attributed to liquidity.

Compared to the results in Table 6, the book-to-market effect has become stronger. This is to be expected. Due to the negative correlation between the market equity and the book-to-market ratio, when size is measured by the book equity, part of the negative size effect will be captured as a stronger book-to-market effect. The momentum effect becomes weaker but remains significant at the level of 10%.

As a further test of the potential role of liquidity in the size effect, we now control for the total book equity, BE, of the whole firm and examine whether the fraction of tradable shares, i.e., the floating ratio, FR, matters for expected stock returns. If tradable shares and nontradable shares are equally important in the size-return relation, then the coefficient of FR will be insignificantly different from zero. However, if the size of tradable shares, due to its positive correlation with liquidity, has a stronger negative relation with stock returns, the coefficient of FR will be significantly negative. The regression results, which are reported in the last three columns of Table 8, show a significantly negative coefficient of FR in all the three columns. This suggests that not only the total firm size but also the relative weights of tradable and nontradable shares matter for the expected returns. Given the firm size, stocks of firms with a higher floating ratio are more liquid, therefore, in equilibrium they deliver a lower return. This result provides further evidence for the importance of liquidity on the size effect.

Interestingly, both Bailey, Cai, Cheung, and Zheng (2003) and Wang and Xu (2005) report a positive relation between the floating ratio and stock returns. They interpret it as evidence of better corporate governance in firms with a higher floating ratio.[18] A crucial difference between their model specifications and ours is that they only control for the size of tradable shares while we control for the size of the whole firm. While the floating ratio can be naturally interpreted as a measure of liquidity if the total firm size is given, its relation with liquidity is not clear once the size of tradable shares is given. As a result, their models do not capture the liquidity implication of the floating ratio.[19] Our results indicate that the liquidity effect is large enough to dominate the potential corporate governance effect.

6 Conclusion

The size effect in stock returns has been documented extensively in the literature. However, the sources of this effect is still subject to debate. In particular, due to the close relation between firm size and stock liquidity, it remains unclear to what extent the size effect can be attributed to liquidity.

Using a special feature of the Chinese stock market, namely, the coexistence of a trad-

[18]Implicitly, they assume that the positive effect of good corporate governance is not fully anticipated by the market.

[19]Given the size of tradable shares, the floating ratio is an inverse measure of the size of nontradable shares. Therefore, it may also be the case that the positive coefficient in their models simply captures the negative size effect of nontradable shares.

able and nontradable shares in the listed companies, we perform a test of the liquidity-based explanation for the size-effect against the risk-based explanation. We find that the size of nontradable shares is negatively related to expected stock returns after controlling for several common liquidity measures such as turnover rate, trading volume and liquidity ratio, as well as the size of tradable shares. This is in sharp contrast to the results of several previous studies and indicates the existence of a genuine size effect independent of liquidity.

While our results do not support the view that the size effect can be fully explained by liquidity, we also find some evidence in support of the importance of liquidity in the size effect. Since the size of tradable shares is directly related to stock liquidity, the size effect of tradabe shares is substantially stronger than that of nontradable shares. This indicates that a significant component of the size effect, when size is measured by tradable shares, comes from the difference in liquidity between large and small stocks.

Our analysis based on individual stock data also adds further evidence for the generality of the cross-sectional patterns in stock returns documented in the previous literature. In contrast to studies based on portfolio returns, we find that despite many distinct features in the market environment, Chinese stock returns exhibit patterns strikingly similar to those documented for developed markets. More specifically, the expected returns are negatively related to the firm size and the turnover rate, and positively related to the book-to-market equity ratio. Furthermore, short-term performance tends to reverse itself, while the intermediate-term performance tends to continue. These cross-sectional patterns are not only statistically significant but also economically important.

References

Amihud, Y., and H. Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223–249.

Amihud, Y., and H. Mendelson, 1989, The Effects of Beta, Bid-Ask Spread, Residual Risk, and Size on Stock Returns, *Journal of Finance* 44, 479–486.

Amihud, Y., H. Mendelson, and B. Lauterbach, 1997, Market microstructure and securities values: Evidence from the Tel Aviv Stock Exchange, *Journal of Financial Economics* 45, 365–390.

Bailey, W., J. Cai, Y. Cheung, and Z. Zheng, 2003, The cross-sectional determinants of stock returns in China: Further evidence on characteristics, factors, and momentum, Working paper.

Banz, R. W., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3–18.

Berk, J., 1995, A critique of size related anomalies, *Review of Financial Studies* 8, 275–286.

Berk, J., R. C. Green, and V. Naik, 1999, Optimal investment, growth options and security returns, *Journal of Finance* 54, 1553–1608.

Black, F., and M. Scholes, 1974, The effects of dividend yield and dividend policy on common stock prices and returns, *Journal of Financial Economics* 1, 1–22.

Brennan, M. J., T. Chordia, and A. Subrahmanyam, 1998, Alternative factor specications, security characteristics, and the cross-section of expected stock returns, *Journal of Financial Economics* 49, 345–373.

Brennan, M. J., T. Chordia, and A. Subrahmanyam, 2005, Cross-sectional determinants of expected returns, in Bruce Lehmann (ed.), *The Legacy of Fischer Black*, Oxford University Press.

Brennan, M. J., and A. Subrahmanyam, 1996, Market microstructure and asset pricing: On the compensation for illiquidity in stock returns, *Journal of Financial Economics* 41, 341–364.

Chan, K. C., and N.-F. Chen, 1991, Structural and return characteristics of small and large firms, *Journal of Finance* 46, 529–54.

Chen, Z., and P. Xiong, 2001, Discounts on Illiquid Stocks: Evidence from China, Working Paper, Yale University.

Chordia, T., A. Subrahmanyam, and V. R. Anshuman, 2001, Trading activity and expected stock returns, *Journal of Financial Economics* 59, 3–32.

Connor, G., and R. Korajczyk, 1988, Risk and return in an equilibrium APT: Application of a new test methodology, *Journal of Financial Economics* 21, 255–290.

Datar, V., and Mao, 1998, Initial public offerings in China: Why is underpricing so severe, Working paper, Seattle University.

Datar, V., N. Naik, and R. Radcliffe, 1998, Liquidity and asset returns: An alternative test, *Journal of Financial Markets* 1, 203–219.

DeBondt, W., and R. Thaler, 1985, Does the stock market overreact?, *Journal of Finance* 40, 793–805.

Dimson, E., 1979, Risk measurement when shares are subject to infrequent trading, *Journal of Financial Economics* 7, 197–226.

Fama, E. F., and K. R. French, 1992, The cross section of expected stock returns, *Journal of Finance* 47, 427–466.

Fama, E. F., and K. R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, E. F., and K. R. French, 1996, Multifactor explanations for asset pricing anomalies, *Journal of Finance* 51, 55–84.

Fama, E. F., and J. MacBeth, 1973, Risk and return: Some empirical tests, *Journal of Political Economy* 81, 607–636.

Gomes, J. F., L. Kogan, and L. Zhang, 2003, Equilibrium cross section of returns, *Journal of Political Economy* 111, 693–732.

Haugen, R., and N. Baker, 1996, Commonality in the determinants of expected stock returns, *Journal of Financial Economics* 41, 401–439.

Jegadeesh, N., 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881–898.

Jegadeesh, N., and S. Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–92.

Kluger, B. D., and J. Stephan, 1997, Alternative Liquidity Measures and Stock Returns, *Review of Quantitative Finance and Accounting* 8, 1936.

Liu, T., 2003, Investment without risk: An investigation into IPO underpricing in China, Working paper, The Royal Institute of International Affairs.

Lo, A. W., and A. C. MacKinlay, 1990, Data-snooping biases in tests of financial asset pricing models, *Review of Financial Studies* 3, 431–468.

Newey, W., and K. West, 1987, A simple, positive, semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Pastor, L., and R. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* pp. 642–685.

Roll, R., 1977, A critique of the asset pricing theory tests: On past and potential testability of theory, *Journal of Financial Economics* 4, 129–176.

Roll, R., and S. Ross, 1980, An empirical investigation of the arbitrage pricing theory, *Journal of Finance* 35, 1073–1103.

Rouwenhorst, K., 1999, Local return factors and turnover in emerging stock markets, *Journal of Finance* 54, 1439–1464.

Spiegel, M., and X. Wang, 2005, Cross-sectional variation in stock returns: liquidity and idiosyncratic risk, Working Paper, Yale University.

Stoll, H. R., and R. E. Whaley, 1983, Transactions Costs and the Small Firm Effect, *Journal of Financial Economics* 12, 57–79.

Vassalou, M., and Y. Xing, 2004, Default risk in equity returns, *Journal of Finance* 59, 831–868.

Wang, F., and Y. Xu, 2005, What Determines Chinese Stock Returns?, *Financial Analysts Journal* 60, 6, 65–77.

Table 1: Summary Statistics

This table presents annual statistics for our final sample. **N** is the number of listed companies. **Return** is the average monthly return. **Market Equity(A)** is the average market value of tradable A-shares per firm in millions of RMB. **Book Equity** is the average total book equity per firm, denominated in millions of RMB. **BM Ratio** is the average book-to-market ratio, calculated by dividing the book equity per share by the share price at the end of the year. **Turnover** is the average monthly turnover rate, defined as the number of traded A-shares to the number of tradable A-shares. **Trading Volume** is the average monthly trading volume, measured in millions of RMB. **Floating Ratio** is the ratio of tradable A-shares to total shares outstanding. **Market Equity (A)**, **Book Equity**, **BM Ratio**, and **Floating Ratio** are averaged across firms using end-of-year data, while **Return**, **Turnover**, and **Trading Volume** are simple averages of all monthly observations available in each year.

year	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
N	155	282	306	509	715	821	918	1053	1053	1050	1037	1027
Return	0.015	0.036	-0.003	0.062	0.024	0.009	0.020	0.039	-0.015	-0.017	-0.011	-0.012
Market Equity (A)	373	266	249	470	635	649	834	1439	1154	967	940	802
Book Equity		541	607	547	637	728	799	917	956	1000	1097	1193
Floating Ratio	0.258	0.292	0.312	0.301	0.302	0.306	0.326	0.349	0.364	0.373	0.378	0.381
BM Ratio		0.435	0.538	0.285	0.263	0.279	0.256	0.177	0.238	0.318	0.377	0.461
Turnover	0.764	0.879	0.351	0.909	0.606	0.393	0.358	0.428	0.181	0.169	0.179	0.249
Trading Volume	267	241	108	426	360	240	281	491	233	175	182	242

Table 2: Cross-correlations between stock characteristics: 1995.7-2004.12

This table presents the time-series average of the monthly cross-sectional correlations between various transformed stock characteristics, including the return, over the 114 months from July 1995 to December 2004 (note that variables related to accounting information are only observed after July 1995). **Return** is the monthly stock return. ME_A and BE_A are the market value and book value of tradable A-shares, respectively. BE_N and **BE** are the book values of nontradable equity and total equity, respectively. **FR** is the floating ratio. **BM** is the book-to-market equity ratio. **TO** is the turnover rate. **TV** is the trading volume in unit of currency. **LR** is the liquidity ratio. **R1, R2, R12** denote the average monthly returns in months $t-1$, $t-2$, $t-3$ to $t-12$, respectively (t represents the current month). ME_A and **FR** are measured at the end of month $t-1$. **TO**, **TV** and **LR** are measured over the 12 months prior to the current month. From July of year i to June of year $i+1$, **BE** and **BM** are calculated using the data at the end of year $i-1$. All the variables are logarithmized except the return and lagged returns. Furthermore, they are all transformed into deviations from the cross-sectional means in each month. **BM** for firms with a negative book equity is set as missing.

	Return	ME_A	BE_A	BE_N	*BE*	*FR*	*BM*	*TO*	*TV*	*LR*	*R*1	*R*2	*R*12
Return	1												
ME_A	-0.033	1											
BE_A	0.001	0.785	1										
BE_N	-0.007	0.477	0.666	1									
BE	-0.007	0.594	0.801	0.974	1								
FR	0.007	0.318	0.323	-0.467	-0.290	1							
BM	0.046	0.204	0.666	0.470	0.553	0.179	1						
TO	-0.014	-0.171	-0.242	-0.214	-0.238	-0.047	-0.023	1					
TV	-0.047	0.779	0.585	0.336	0.426	0.243	0.134	0.407	1				
LR	-0.039	0.855	0.646	0.390	0.484	0.253	0.123	0.235	0.953	1			
*R*1	-0.017	0.085	0.002	-0.007	-0.006	0.008	0.048	0.033	-0.006	-0.012	1		
*R*2	-0.037	0.082	-0.000	-0.008	-0.008	0.006	0.043	0.061	0.024	0.016	-0.020	1	
*R*12	0.011	0.264	0.004	-0.015	-0.015	0.018	0.052	0.299	0.299	0.280	-0.006	-.005	1

Table 3: The determinants of expected stock returns: basic model

This table presents the cross-sectional regression results of our basic model. The results are based on 96 month-by-month cross-sectional regressions, covering the period from January 1997 to December 2004. The dependent variable in the first column is the unadjusted excess return while in the next two columns it is the return adjusted by the market factor, in the last two columns it is the return adjusted by the Fama-French factors. ME_A is the market value of tradable A-shares. **BM** is the book-to-market equity ratio. **TO** is the turnover rate. **R1, R2, R12** denote the average monthly returns in months $t-1$, $t-2$, $t-3$ to $t-12$, respectively (t represents the current month). The coefficients in columns labelled "FM" are estimated using the Fama-MacBeth estimator, while those columns labelled "BS" are estimated using the purged estimator introduced by Black and Scholes (1974). Newey-West corrected t-statistics are in parentheses. *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

	Excess return	One-factor model		Three-factor model	
	FM	FM	BS	FM	BS
Intercept	0.004	0.003	0.003	0.000	-0.000
	(0.52)	(1.08)	(1.10)	(0.30)	(-0.36)
ME_A	-0.006*	-0.007**	-0.007**	-0.005**	-0.003*
	(-1.88)	(-2.24)	(-2.24)	(-2.61)	(-1.82)
BM	0.006**	0.008**	0.008**	0.007**	0.004
	(2.29)	(2.54)	(2.57)	(2.25)	(1.62)
TO	-0.008***	-0.008***	-0.008***	-0.012***	-0.009**
	(-2.93)	(-2.98)	(-2.92)	(-3.12)	(-2.12)
$R1$	-0.039**	-0.038**	-0.037**	-0.050***	-0.039**
	(-2.62)	(-2.49)	(-2.48)	(-3.49)	(-2.53)
$R2$	-0.031**	-0.035**	-0.035**	-0.035**	-0.024
	(-2.23)	(-2.50)	(-2.49)	(-2.51)	(-1.53)
$R12$	0.151***	0.159***	0.158***	0.080**	0.115***
	(2.81)	(2.69)	(2.77)	(2.26)	(3.42)

Table 4: Return predicability: risk adjustment by the one factor model

In each month t, the expected market-factor-adjusted return of each stock in month $t+1$ is forecast based on stock characteristics observed at the end of month t, and the cross-sectional relations between stock characteristics and the market-factor-adjusted returns estimated in months t through $t-11$. Stocks are then ranked by their relative expected market-factor-adjusted returns and formed into ten equally-weighted deciles, with decile 1 containing the stocks expected to have the worst performance. Panel A and B present the realized unadjusted returns and market-factor-adjusted returns (in percentage) of each decile, respectively.

	1998	1999	2000	2001	2002	2003	2004	1998-2004
Panel A: unadjusted realized returns (%)								
D1	-1.39	1.60	2.54	-2.60	-2.44	-2.70	-1.86	-0.98
D2	-1.07	1.52	2.62	-2.44	-2.36	-2.03	-1.26	-0.72
D3	-0.43	0.94	2.79	-2.15	-1.88	-1.60	-1.64	-0.57
D4	-0.29	1.61	3.92	-1.88	-1.72	-1.41	-1.33	-0.16
D5	0.72	2.63	4.62	-1.71	-1.73	-1.20	-1.23	0.30
D6	1.23	1.91	3.98	-1.35	-1.73	-0.99	-0.94	0.30
D7	1.51	1.51	4.51	-1.25	-1.41	-0.70	-1.08	0.44
D8	1.20	2.12	4.53	-0.93	-1.35	-0.44	-0.91	0.60
D9	1.59	2.33	5.24	-0.71	-1.02	-0.07	-0.79	0.94
D10	2.72	2.53	4.90	-0.56	-0.89	0.29	-0.68	1.19
Panel B: realized returns adjusted by the market factor (%)								
D1	-1.46	0.41	-0.54	-0.36	-0.90	-2.17	-0.89	-0.84
D2	-0.97	-1.29	0.07	-1.07	-0.63	-1.71	0.07	-0.79
D3	0.57	-0.65	-0.12	-0.52	-0.00	-1.36	-0.28	-0.34
D4	0.95	-0.71	1.78	-0.04	0.08	-1.05	0.06	0.15
D5	1.21	0.88	1.70	-0.05	-0.26	-1.09	0.15	0.36
D6	2.12	0.06	0.68	0.48	-0.11	-0.97	0.43	0.39
D7	2.73	0.83	1.84	0.39	0.31	-0.55	0.39	0.85
D8	2.45	0.82	1.68	0.65	0.33	-0.36	0.66	0.89
D9	1.85	2.07	2.87	1.21	0.71	0.00	0.74	1.35
D10	3.07	1.65	2.96	0.70	0.84	0.36	0.88	1.49

Table 5: Return predicability: risk-adjustment using the Fama-French factors

In each month t, the expected return of each stock in month $t+1$, adjusted by the Fama-French factors, is forecast based on stock characteristics observed at the end of month t, and the cross-sectional relations between stock characteristics and returns adjusted by the Fama-French factors estimated in months t through $t-11$. Stocks are then ranked by their relative expected risk-adjusted returns and formed into ten equally-weighted deciles, with decile 1 containing the stocks expected to have the worst performance. Panel A and B present the realized unadjusted returns and returns adjusted by the Fama-French factors (in percentage) of each decile, respectively.

	1998	1999	2000	2001	2002	2003	2004	1998-2004
Panel A: unadjusted realized return (%)								
D1	-0.93	1.67	2.34	-2.55	-2.15	-1.65	-2.67	-0.85
D2	-1.02	1.49	2.37	-2.57	-2.14	-0.91	-1.55	-0.62
D3	-0.07	1.54	3.81	-1.98	-2.19	-1.21	-1.48	-0.23
D4	-0.24	1.52	5.36	-1.98	-2.02	-1.01	-1.29	0.05
D5	0.19	1.87	4.00	-1.27	-2.05	-0.87	-1.00	0.13
D6	1.26	2.04	4.16	-1.28	-1.49	-1.13	-0.91	0.38
D7	0.77	2.77	4.75	-1.23	-1.36	-1.17	-0.60	0.56
D8	1.58	1.41	4.20	-1.01	-1.19	-1.09	-0.58	0.48
D9	2.13	2.07	4.49	-1.09	-1.06	-0.69	-0.89	0.71
D10	2.11	2.32	4.22	-0.64	-0.80	-1.16	-0.69	0.77
Panel B: realized returns adjusted by Fama-French factors (%)								
D1	-2.42	-0.10	-1.17	-1.01	-0.21	-1.05	-1.75	-1.10
D2	-2.95	-0.03	-1.09	-1.16	-0.32	-0.33	-0.69	-0.94
D3	-0.84	0.02	0.08	-0.90	-0.18	-0.47	-0.42	-0.39
D4	-0.21	0.38	2.21	-0.73	-0.25	-0.07	0.01	0.19
D5	-0.68	-0.18	0.18	0.03	-0.10	0.03	0.25	-0.07
D6	0.47	1.15	1.22	-0.10	0.41	-0.37	0.39	0.45
D7	1.20	1.52	2.08	0.34	0.41	-0.21	0.69	0.86
D8	1.91	-0.31	0.32	0.24	0.64	0.08	1.05	0.56
D9	1.21	0.45	1.11	0.42	0.58	0.41	0.78	0.71
D10	0.53	0.36	0.50	1.30	1.20	0.12	1.12	0.74

Table 6: Size of nontradable equity and stock returns

The results are based on 96 month-by-month cross-sectional regressions, covering the period from January 1997 to December 2004. The dependent variable in the first and the fourth columns is the unadjusted excess return while in the rest columns it is the return adjusted by the market factor. ME_A is the market value of tradable A-shares. BE_N is the book value of nontradable equity. **BM** is the book-to-market equity ratio. **TO** is the turnover rate. **TV** is the trading volume in unit of currency. **R1, R2, R12** denote the average monthly returns in months $t-1$, $t-2$, $t-3$ to $t-12$, respectively (t represents the current month). The coefficients in the column labelled "FM" are estimated using the Fama-MacBeth estimator, while those in the column labelled "BS" are estimated using the purged estimator introduced by Black and Scholes (1974). Newey-West corrected t-statistics are in parentheses. *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

	Turnover			Trading volume		
	Excess return	Risk-adjusted return		Excess return	Risk-adjusted return	
	FM	FM	BS	FM	FM	BS
$Intercept$	0.004	0.003	0.003	0.004	0.003	0.003
	(0.555)	(1.19)	(1.22)	(0.55)	(1.13)	(1.17)
BE_N	-0.004**	-0.003**	-0.003**	-0.003**	-0.003*	-0.003*
	(-2.05)	(-2.03)	(-2.03)	(-2.00)	(-1.98)	(-1.98)
ME_A	-0.006*	-0.007**	-0.007**	0.006	0.005	0.005
	(-1.88)	(-2.38)	(-2.38)	(1.35)	(1.07)	(1.10)
BM	0.008**	0.011**	0.011**	0.009**	0.011**	0.011**
	(2.39)	(2.39)	(2.49)	(2.44)	(2.42)	(2.53)
TO	-0.011***	-0.011***	-0.011***			
	(-3.75)	(-3.82)	(-3.80)			
TV				-0.013***	-0.012***	-0.012***
				(-4.00)	(-4.20)	(-4.20)
$R1$	-0.038***	-0.037**	-0.037**	-0.047***	-0.046***	-0.045***
	(-2.69)	(-2.48)	(-2.47)	(-3.40)	(-3.11)	(-3.12)
$R2$	-0.033**	-0.037***	-0.037***	-0.042***	-0.046***	-0.046***
	(-2.39)	(-2.72)	(-2.70)	(-3.15)	(-3.43)	(-3.44)
$R12$	0.146***	0.156**	0.155**	0.118**	0.128**	0.127**
	(2.68)	(2.56)	(2.63)	(2.19)	(2.10)	(2.15)

Table 7: Size of nontradable equity and stock returns: further tests

The results are based on 96 month-by-month cross-sectional regressions, covering the period from January 1997 to December 2004. The dependent variable in the first is the unadjusted excess return while in the rest two columns it is the return adjusted by the market factor. BE_N is the book value of nontradable equity. ME_A is the market value of tradable A-shares. **BM** is the book-to-market equity ratio. **LR** is the liquidity ratio. **R1, R2, R12** denote the average monthly returns in months $t-1$, $t-2$, $t-3$ to $t-12$, respectively (t represents the current month). The coefficients in the column labelled "FM" are estimated using the Fama-MacBeth estimator, while those in the column labelled "BS" are estimated using the purged estimator introduced by Black and Scholes (1974). Newey-West corrected t-statistics are in parentheses. *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

	Excess Return	Risk-adjusted return	
	FM	FM	BS
Intercept	0.006	0.003	0.003
	(0.76)	(1.13)	(1.16)
BE_N	-0.003*	-0.003*	-0.003*
	(-1.90)	(-1.77)	(-1.75)
ME_A	0.005	0.003	0.003
	(1.06)	(0.76)	(0.82)
BM	0.008**	0.010**	0.009**
	(1.99)	(2.17)	(2.22)
LR	-0.010***	-0.009***	-0.010***
	(-2.85)	(-3.52)	(-3.46)
$R1$	-0.047***	-0.045***	-0.045***
	(-3.18)	(-3.07)	(-3.08)
$R2$	-0.043***	-0.046***	-0.046***
	(-2.88)	(-3.44)	(-3.45)
$R12$	0.119**	0.126**	0.125**
	(2.26)	(2.09)	(2.13)

Table 8: Contribution of liquidity to the size effect

This table presents the results based on 96 month-by-month cross-sectional regressions from January 1997 to December 2004. The dependent variable in the first and the fourth columns is the unadjusted excess return while in the rest columns it is the return adjusted by the market factor. BE_A and BE_N are the book values of tradable and nontradable shares, respectively. **BE** is the book value of total equity. **FR** is the floating ratio. **BM** is the book-to-market equity ratio. **TO** is the turnover rate. **R1, R2, R12** denote the average monthly returns in months $t-1$, $t-2$, $t-3$ to $t-12$, respectively (t represents the current month). The coefficients in the column labelled "FM" are estimated using the Fama-MacBeth estimator, while those in the column labelled "BS" are estimated using the purged estimator introduced by Black and Scholes (1974). Newey-West corrected t-statistics are in parentheses. *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

	BE_A and BE_N			BE and FR		
	Excess return	Risk-adjusted return		Excess return	Risk-adjusted return	
	FM	FM	BS	FM	FM	BS
$Intercept$	0.004	0.003	0.003	0.005	0.003	0.003
	(0.55)	(1.19)	(1.22)	(0.59)	(1.31)	(1.33)
BE_N	-0.003*	-0.003*	-0.003*			
	(-1.91)	(-1.77)	(-1.76)			
BE_A	-0.006**	-0.007**	-0.007**			
	(-2.05)	(-2.61)	(-2.61)			
BE				-0.009***	-0.010***	-0.010***
				(-2.71)	(-2.99)	(-3.01)
FR				-0.004*	-0.006**	-0.006**
				(-1.72)	(-2.23)	(-2.22)
BM	0.014***	0.018***	0.017***	0.014***	0.016***	0.016***
	(2.78)	(3.04)	(3.19)	(3.11)	(3.29)	(3.48)
TO	-0.011***	-0.011***	-0.011***	-0.011***	-0.011***	-0.011***
	(-3.76)	(-3.93)	(-3.92)	(-3.75)	(-3.91)	(-3.87)
$R1$	-0.042***	-0.042***	-0.042***	-0.045***	-0.045***	-0.045***
	(-2.82)	(-2.71)	(-2.71)	(-2.99)	(-2.89)	(-2.90)
$R2$	-0.038**	-0.043***	-0.043***	-0.040***	-0.044***	-0.043***
	(-2.63)	(-3.02)	(-3.03)	(-2.69)	(-2.97)	(-2.99)
$R12$	0.115*	0.111*	0.111*	0.112*	0.110	0.110*
	(1.72)	(1.64)	(1.67)	(1.69)	(1.63)	(1.67)