Reference price distributions and stock returns: an analysis based on the disposition effect

Yuegang Zhou, Assistant Professor
Chinese Academy of Finance and Development
Central University of Finance and Economics
39 Xueyuan Nanlu, Haidian District
Beijing, China 100081
Tel: 0086-10-62289266
Email: yuegangz@cufe.edu.cn

Xiaoyan Lei, Assistant Professor
China Center for Economic Research
National School of Development
Peking University
Beijing, China 100871
Tel: 0086-10-62758946
Email: xylei@ccer.pku.edu.cn

Reference price distributions and stock returns: an analysis based on the disposition effect

Abstract

This paper provides evidence that reference price distributions can predict stocks' expected returns. We develop a model based on the disposition effect by considering shareholders' trading activities with different relative capital gains. The model suggests that both investors' incentives to sell shares and their paper capital gains impact stock performance over a subsequent period. From the theoretical model and examples, we conjecture that various moment variables for relative capital gains, which measure the shape of a reference price distribution, are associated with future stock returns. By applying four proxy variables this paper finds that the mean and skewness of relative capital gains play important roles in predicting cross-section of stock returns in the Chinese stock markets. It also shows that the mean is the key explanatory variable for variations in returns for stocks with positive average capital gains. Skewness works as the key variable for stocks with negative average capital gains. These findings hold true when the factors identified in the related literature are taken into consideration.

JEL classification: G1; G12

Keywords: Prospect theory; Disposition effect; Reference price distribution; Cross section in stock return

1. Introduction

Shefrin and Statman (1986) found that investors have a tendency to sell winners too early and hold on to losers for too long. They refer to this phenomenon as the "disposition effect". It is widely held that Kahneman and Tversky's (1979) prospect theory and Thaler's (1980) "mental accounting" constitute the basic explanation for this effect[1]. This paper contributes to the literature

[1] Recent literature in this area contributes to integrating psychological evidence on risk preferences into models of equilibrium prices, e.g., Barberis, Huang, and Santos (2002), Barberis, Huang, and Thaler (2003), and Grinblatt and Han (2005).

on the disposition effect with theoretical modeling and empirical investigations from a market-wide angle. That is, we examine whether or not reference price distributions can predict stock returns based on the disposition effect.

Prospect theory describes how one evaluates choices under uncertainty, and provides possible explanations for phenomena that are not in accordance with classical expected utility theory. It is based on the risk attitudes people have when confronted with different situations. An individual is more likely to be risk seeking when facing a loss, and risk averse when facing a gain. If an investor is subject to prospect theory and mental accounting (a "PT-MA investor" henceforth), her utility function is concave on the positive domain and convex on the negative domain of investment gains as shown in Figure 1. The current positions of four PT-MA investors— G1, G2, G3 and G4—are indicated in the figure. G1 is the most risk-seeking investor, G4 the most risk-averse one, and G2 and G3 in between. Their demand for the stock can be ranked from the highest to the lowest as G1, G2, G3, and G4.

(Insert Figure 1 here)

Critical to the utility value of a PT-MA investor is her reference price, by which she is determined to be in a position of loss or profit. As shown in Figure 1, the reference prices of G1 and G2 are higher than the market price, and those of G3 and G4 are lower than the market price. When the stock experiences an appreciation, all investors' positions on the utility curve will move to the right if their reference prices remain unchanged. Reference prices, however, are usually updated from time to time due to trading activities. For instance, G4 may sell her shares to a new investor $G_4^{'}$, and of $G_4^{'}$'s position will be at a point near the inflection point. An update may also be due to the investor's expectations of future returns. If an appreciation of the stock causes G4 to have a higher expectation for future return, then the reference price also increases. Thus, she will move a shorter distance to the right than under a fixed reference price. As do Odean (1998) and Grinblatt and Han (2005), this paper assumes that reference prices are purchase prices.

For a preliminary understanding of the effect that PT-MA investors have on stock price, visualize following situation. A market consists of two types of investors: One is the PT-MA investor whose risk preference is defined by prospect theory, and the other is the rational investor.

Decision making should be markedly different for these two types of investors. We assume, however, that an investor, regardless of her type, will always make decisions like the rational investor before holding any shares of a given stock. When the stock price converges to its fundamental value from below, PT-MA investors are prone to selling their shares to new stockholders if they face higher paper gains, while rational investors hold onto them if they expect a further price rise in the future. If it turns out that the stock price rises, more and more transactions are initiated by PT-MA investors during the "good" period. These trading behaviors delay the market price's quick approach to the stock's fundamental value, resulting in a longer rise in price over the following period. On the other hand, if a stock price falls to its fundamental, rational investors will sell their shares, while PT-MA ones, if they are suffering losses, will hold onto them. As the market price drops, more and more investors will be in a losing position, and sell-initiated transactions will happen with less frequency during the "bad" period. Again, PT-MA investors are deferring the stock price's approach to its fundamental value. Therefore, a higher average capital loss more likely forecasts a further decrease in the stock price. Likewise, a stock with more winning shares may have higher return in the next period.

This paper models the demand function of a PT-MA investor by adding an extra term to the rational linear demand function as does by Grinblatt and Han (2005), to account for the disposition effect. However, we differentiate reference price for each share: if a share is sold, its reference price changes to the new market price. We derive that the selling probability of each share partially determines expected return. The aggregate effect of PT-MA investors' trading behaviors is associated with how their paper gains are distributed, and therefore the model leads us to take into account the shape of a relative capital gains distribution when studying the variations in stock return caused by PT-MA investors.

The theoretical model and empirical examination by Grinblatt and Han (2005) documents the role of capital gains overhang in the continuation of stock returns. The paper goes on to reassess the issue in two steps. The first step is to investigate how stock price is affected by an individual investor with a certain capital gains. Next is to aggregate the effects of all shareholders' behaviors. This process explains why the shape of a reference price distribution is important. We therefore expect that statistical measures, such as variance, skewness, and kurtosis, should have implications on stock's future performance, because they characterize a distribution.

Besides the average capital gains as shown in Grinblatt and Han (2005), the role of higher moment statistics in predicting future returns. For example, two stocks have the same average capital gains but different variances. It is more likely that the higher-variance stock has a similarly higher proportion of shares with extremely high profits and high losses. If those shareholders trade as the disposition effect suggests, then the high-variance stock is more likely to underreact to information, and its future return will be higher than that of the low-variance stock. Now suppose that both stocks' relative capital gains have identical averages and variances but differ in skewness. The high-skewness stock has a high return over the next period because of its larger proportion of highly winning shares or less high-loss shares. In this paper, we hope to reveal the relationships between statistical variables of reference price distributions and cross-section in stock returns.

To obtain moment variables for relative capital gains in a stock, we need transaction records documenting all trades. However, this type of data is not available in most stock markets[2]. Account-based data do bring an advantage to investigating investor' decisions, since the data provide precise purchase and sell prices, holding periods, and other details. For example, Shefrin and Statman (1985) make use of trading records from individual investors and aggregate data on mutual fund trades, while Odean (1998) randomly selects 10,000 accounts at a brokerage house. These empirical studies based on partial trading records can provide evidence on the existence of the disposition effect in corresponding financial markets. However, the main concern about these data is their inability to comprehensively reflect common investor behavior as they do not consist of all market participants. If we are interested in the responses of stock return to all shareholders' trading behaviors, this data from a segment of the markets may not be sufficient.

The theoretical model in this paper provides us with a channel to employ market data to investigate the effects of investors' behaviors. In our empirical investigation, we use daily average prices and trading volumes from stock-based transaction records to construct four proxy variables, which correspond to the four moment variables mentioned above[3]. Kaustia (2004) also uses

---

[2] The data applied by Grinblatt and Keloharju (2000, 2001) consist of all account-based transaction records in Finnish stocks.

[3] Webber and Camerer (1998) mention that control variables such as investors' expectations and individual decisions are not observable in market data.

market data in a recent empirical investigation of the market-wide disposition effect, focusing on variations in trading volumes when stock prices cross different thresholds following IPOs, and avoiding worry over reference prices.

Our results appear to verify our theoretical model. By applying the Fama-MacBeth methodology, we find evidence that some of the proxy variables are able to predict cross-sectional variations in stock returns. For example, the mean (ARC) and skewness (SRC) of relative capital gains are positively correlated with stock returns the following month. We further examine the associations by dividing the stocks into two groups according to the sign of a stock's ARC in each month. Fama-MacBeth regressions reveal that different proxy variables stand out for stocks in different stock groups. For example, for stocks in G-group (with positive ARCs, or a winning group), ARC plays the key role: a stock with a higher ARC has a higher return the next month. For stocks in L-group (with negative ARCs, or a losing group), the key explanatory variable is SRC: a higher SRC implies a higher future return. We also include factors in our regressions pertaining to returns as identified in the literature and re-examine the effects of the four proxy variables. We find that the explanatory abilities of these variables remain.

The paper is structured as follows. Section 2 sets up the model and specifies regression equation for empirical examinations. Section 3 details our methodology for constructing proxy variables with common daily market data. Section 4 provides descriptive statistics of the proxy variables and explores our preliminary regression results. Possible explanations for the connection between our empirical results and the theoretical model are also offered in this section. Section 5 further examines the results obtained in Section 4 taking into account other factors that potentially affect our findings. Section 6 concludes the paper.

2. The model

2.1 Model setup and stock return

Suppose that the number of shares outstanding for a stock is K, which by mental accounting, are held by K investors. Any investor before holding any shares is assumed to be rational[4]. The fundamental of this stock evolves as

$$V_{t+1} = V_t + \varepsilon_{t+1} \tag{1}$$

where V_t is the fundamental value at time t, and ε_t follows a normal distribution with a mean of zero. Equation (1) assumes that the fundamental of the stock is normally distributed with the mean of the stock's historical values.

The trading behaviors of rational investors drive the market price to revert to the fundamental, while the behaviors of PT-MA investors may have opposite effect. The total effect depends on the aggregate trading strength of PT-MA investors with different relative capital gains.

As proposed by Grinblatt and Han (2005), the demand functions of the two types of investors are

$$D_{k,t}^{Rational} = 1 + b_t(V_t - P_t), \textit{ if the holder of the } k^{th} \textit{ share is rational};. \tag{2}$$

$$D_{k,t}^{PT-MA} = 1 + b_t\left\{(V_t - P_t) + \lambda\left(C_{k,t} - P_t\right)\right\}, \textit{ if the holder of the } k^{th} \textit{ share is } \text{PT-MA}. \tag{3}$$

$D^{Rational}$ and $D^{PT\text{-}MA}$ denote the demands of rational investors and PT-MA investors respectively; P_t is the market price at date t; $C_{k,t}$ is the reference price (mental cost) of the k^{th} investor at time t if she is subject to the disposition effect; b_t represents the slope of the rational investors' demand function. The positive parameter λ measures the relative strength of the PT-MA investors to induce extra demand. Equation (3) implies that the relative strength is the same for PT-MA investors with the same reference prices. We further assume that a shareholder is a PT-MA investor with probability π. Thus the total market demand D_t is the sum of all individuals' demands[5],

$$D_t = \sum_{k=1}^{K}\left[\left(1-\pi\right)D_t^{Rational} + \pi D_{k,t}^{PT-MA}\right].$$

By substituting (2) and (3) into the market demand function, we rearrange the right-hand side and obtain

---

4 This assumption is to ensure the demands of those investors having no effect on the total demand.

5 The market demand does not consist of the investors who currently do not hold the stock. However, since we assume that these investors without any share are rational, their effect on the return in the equilibrium disappears as we will see later.

$$D_t = K + Kb_t(V_t - P_t) + \lambda b_t \sum_{k=1}^{K} \pi\left(C_{k,t} - P_t\right). \tag{4}$$

The market clearing condition implies that the equilibrium price is

$$P_t = \frac{1}{1+\lambda\pi} V_t + \frac{\lambda\pi}{K\left(1+\lambda\pi\right)} \sum_{k=1}^{K} C_{k,t} \ . \tag{5}$$

Equation (5) indicates that the equilibrium price is the weighted average of the fundamental value and reference prices of all PT-MA investors.

For a given share, its holder's reference price evolves as follows,

$$C_{k,t+1} = P_t 1_{\{sold\}k,t} + \left(1 - 1_{\{sold\}k,t}\right) C_{k,t}, \tag{6}$$

where indicator function $1_{\{sold\}k,t}$ is one if share k is sold at time t, and zero otherwise. This specification is reasonable, because whenever a share is traded, its holder's reference price should be the current market price.

By virtue of the results above, the change in price from date t to date $t+1$ can be represented as

$$P_{t+1} - P_t = \frac{1}{1+\lambda\pi}\left(V_{t+1} - V_t\right) + \frac{\lambda\pi}{K\left(1+\lambda\pi\right)} \sum_{k=1}^{K}\left(C_{k,t+1} - C_{k,t}\right).$$

By substituting $C_{k,t+1}$ in Equation (6), and taking expectation on both sides, we obtain the expected change in price, which is

$$E_t\left[P_{t+1} - P_t\right] = \frac{\lambda\pi}{K\left(1+\lambda\pi\right)} \sum_{k=1}^{K}\left[\Pr\left(1_{\{sold\}k,t} = 1\right)\left(P_t - C_{k,t}\right)\right],$$

since the expected changes in the fundamental are zero by assumption. $\Pr\left(1_{\{sold\}k,t} = 1\right)$ represents the probability that share k is sold at time t. Therefore, the expected stock return is

$$E_t\left[\frac{P_{t+1} - P_t}{P_t}\right] = \frac{\lambda\pi}{K\left(1+\lambda\pi\right)} \sum_{k=1}^{K}\left[\Pr\left(1_{\{sold\}k,t} = 1\right)\left(\frac{P_t - C_{k,t}}{P_t}\right)\right]. \tag{7}$$

In contrast to the equilibrium price presented in Grinblatt and Han (2005), this formula takes into account different reference prices, and the possibility of an investor being PT-MA. Our model is arguably more realistic, since the trading activities of investors at different positions of the utility curve implied by prospect theory should be explicitly distinguished. Equation (7) demonstrates that the expected change in stock price depends on two critical factors. The first is the relative capital gains, $\frac{P_t - C_{k,t}}{P_t}$. The second is the probability that a share is sold, $\Pr\left(1_{\{sold\}k,t} = 1\right)$. Both are positively correlated with expected return. More importantly, Equation (7) implies that the distribution of relative capital gains (RC) determines the expected stock return as the combined result of the two factors.

To discuss the effect of a relative capital gains distribution on expected return, we must make an assumption about the selling probability, $\Pr\left(1_{\{sold\}k,t} = 1\right)$. Since PT-MA investors tend to sell their winning shares and hold onto losing ones, we assume that $\Pr\left(1_{\{sold\}k,t} = 1\right)$ is an non-decreasing function of $\left(\frac{P_t - C_{k,t}}{P_t}\right)$. Although it may hold for rational shareholders, Equation (7) mainly considers the trading behaviors of PT-MA investors.

The fact that expected return is determined by the distribution of relative capital gains can be identified in the summation notation on the right-hand side of Equation (7). It reflects the aggregate effect of trading behaviors of all investors at different positions along the profit axis. Based on the expression, we should estimate the relative capital gains and the selling probability for each shareholder in order to forecast a stock return.

Our expected return formula also suggests the result of Grinblatt and Han (2005), which documents that the expected return is an increasing function of unrealized capital gains. However this may be for two reasons: the probability of selling a share, and the relative capital gains. Our model apparently shows that higher moment variables are also meaningful.

The main interest in this paper is in investigating how reference price distribution predicts expected return. It is well known that a distribution can be represented by its moment variables. Thus we may instead use moment variables to characterize a distribution as discussed in the introduction, such as the mean, variance, skewness, and kurtosis of relative capital gains. By exploring the associations between these variables and returns, we are able to learn how the shape of a reference price distribution is related to future return.

2.2 Return approximation and explanations

Equation (7) provides us with a way to analyze how different shapes of capital gains distributions, or reference price distributions, are associated with future stock returns. First, suppose that two stocks, Stock 1 and Stock 2, are identical except for the paper gains of their shareholders. Assume each stock is equally held by four investors. The relative capital gains of Stock 1's holders are -0.1, -0.05, 0.05, and 0.05; and those of Stock 2 are -0.1, -0.05, 0.05, and 0.1. If the selling probabilities of shares with the same relative gains are equal, then by plugging in the relative capital gains and selling probabilities into equation (7) for each stock, we can obtain that the difference in returns, $r_2 - r_1$, is $a\left(0.1\,\mathrm{Pr}_2 - 0.05\,\mathrm{Pr}_1\right)$. The parameter a is composed of all constant terms in equation (7); and Pr_2 is the selling probability of Stock 2 with 0.1 relative capital gains, and Pr_1 is that of Stock 1 with 0.05 relative capital gains. Since the selling probability is non-decreasing in relative capital gains, it is easy to obtain that $P_2 \geq P_1$. Therefore, $a\left(0.1\,\mathrm{Pr}_2 - 0.05\,\mathrm{Pr}_1\right) \geq 0$. Denote $MEAN_1$ and $MEAN_2$ to be the means of the two stocks' relative capitals. From the simple example, we can conjecture that Stock 2 has a higher return if $MEAN_2 > MEAN_1$. That is, stock return is a non-decreasing function of average relative capital gains. This result obviously confirms what Grinblatt and Han (2005) obtained.

Second, assume that the means of relative capital gains are equal for Stock 1 and Stock 2, but the variances, VAR_1 and VAR_2, are not equal. As an example, we assume again that each stock is equally held by four shareholders. The relative capital gains are -0.1, -0.05, 0.05, and 0.1 for Stock 1, and -0.2, -0.05, 0.05, and 0.2 for Stock 2. Plugging in these relative capital gains and corresponding selling probabilities, we can observe that the difference in returns, $r_2 - r_1$, is now $a\left(-0.2\,\mathrm{Pr}_{21} + 0.2\,\mathrm{Pr}_{24} + 0.1\,\mathrm{Pr}_{11} - 0.1\,\mathrm{Pr}_{14}\right)$, where Pr_{ij}, $i = 1, 2,\ j = 1, 2, 3, 4$, denotes the probability of shareholder *j* selling her holdings of Stock *i*. Under the assumption that selling probability is non-decreasing in relative capital gains, we get $\mathrm{Pr}_{22} - \mathrm{Pr}_{21} \geq \mathrm{Pr}_{12} - \mathrm{Pr}_{11} \geq 0$, thus $r_2 \geq r_1$. As $VAR_2 > VAR_1$, we conjecture that the stock return is non-decreasing in variance of relative capital gains. This result is different from Grinblatt and Han (2005) that predicts equal expected returns for stocks with equal means of relative capital gains.

Third, assume that relative capital gains of Stock 1 and Stock 2 have equal means as well as equal variances, but they have different skewnesses, for instance, $SKEW_1 < SKEW_2$. To illustrate this situation, we instead assume that each stock are equally held by five investors. The relative capital gains for Stock 1 are -0.2, -0.1, 0.05, 0.1 and 0.15; and those for Stock 2 are -0.15, -0.1, -0.05, 0.1, and 0.2. We can verify that both stocks have the same mean and variance, but $SKEW_1 = -0.41$, and $SKEW_2 = 0.41$. Define $\Pr_{ij}$, $i = 1, 2$, $j = 1, ..., 5$, to be the probability of shareholder *j* selling her holdings of Stock *i*. Similar analysis leads to the difference in returns, $r_2 - r_1$, being equal to $a\left[0.15\left(\Pr_{25} - \Pr_{21}\right) + 0.5\left(\Pr_{25} - \Pr_{22}\right) - 0.15\left(\Pr_{15} - \Pr_{11}\right) - 0.5\left(\Pr_{13} - \Pr_{11}\right)\right]$. If we can further assume that $\Pr_{25} - \Pr_{22} \geq \Pr_{13} - \Pr_{11}$ and $\Pr_{25} - \Pr_{21} \geq \Pr_{15} - \Pr_{11}$ [6], then $r_2 - r_1 \geq 0$. As $SKEW_2 > SKEW_1$, we conjecture that expected return is a non-decreasing function of skewness of relative capital gains.

Finally, assume that the two stocks have equal means, variance and skewnesses of relative capital gains, but different kurtoses. It is not easy to give a numerical example of this situation. However, we may explore the implications of kurtosis of relative capital gains by intuitive argument. Kurtosis measures the degree of concentration of a distribution on the mode. Thus the situation is different for positive and negative mode. When the mode is positive (negative), a higher kurtosis implies that more shares concentrate on a level with positive (negative) gains. Thus according to the disposition effect, a higher kurtosis for a positive mode suggests that more shares will potentially simultaneously be sold, while a higher kurtosis for a negative mode suggests that more shares will continue to be held. Denote skewness to be $KURT$. If we assume that the mode is equal to the mean, we conjecture that stock return is a non-decreasing function of $KURT$ when $MEAN > 0$; and that stock return is a non-increasing function of $KURT$ when $MEAN < 0$.

The higher moment variables also convey some information on the shape of a reference price

[6] For stockholders subject to disposition effect, the assumption should be reasonable, since the difference in selling probabilities of different capital gains is smaller for shares with paper losses, and larger for shares with paper profits

distribution. As the literature usually does, we focus only on the four variables in this paper. In our investigation, we also try including higher moment variables, but do not find meaningful results.

To obtain true reference price distributions, we have to collect all transaction records of each investor for all points of time. However, this type of data is usually not available. Alternatively, based on the examples and analysis above, we can approximate a stock's expected return with a liner function

$$E_t\left[\frac{P_{t+1}-P_t}{P_t}\right]=h_0+h_1 MEAN_{1t}+h_2 VAR_t+h_3 SKEW_t+h_4 KURT_t, \tag{8}$$

where $MEAN_t$, VAR_t, $SKEW_t$, and $KURT_t$ represent mean, variance, skewness, and kurtosis of relative capital gains for a stock at time *t*; h_0 captures the proportion of the expected return related to higher moment terms and other factors. Coefficients h_i, $i=1,...,4$, denote the loadings of expected return on the corresponding moment variables. The four variables in Equation (8) measure different aspects of a relative capital gains distribution.

The theoretical model and analysis above promote us to develop the following hypothesis .

Hypothesis: If the disposition effect has a market-wide effect, then empirical examination based on equation (8) should verify that $h_1>0$, or $h_2>0$, or $h_3>0$; and $h_4>0$ if $MEAN>0$, or $h_4<0$ if $MEAN<0$, whenever they are significant.

Among the four moment variables, *VAR* is also related to financial theory such as dispersion of opinion. When the difference of investors' evaluation of a stock is larger, the stock's reference price distribution has a higher variance. However, the theories on dispersion of opinion have not got a unanimous prediction to the sign of h_2. This paper may give a certain attempt to investigation in this area.

3. Variables and data

3.1 Variable construction

As mentioned in the introduction, we apply the transaction records at the individual stock level to construct the moment variables. The procedure is as follows:

For a given stock, we start with the last trading day T.

Step 1: Accumulate daily turnover ratios backwards until the cumulative turnover reaches 100%. The latest day satisfying the criterion of 100% cumulative turnover is marked as S_T [7].

Step 2: At the time interval $[S_T, T]$, apply the formulas, introduced later, to calculate the variables' values at day T.

Step 3: For day $T-1$, repeat Steps 1 and 2 to obtain the variables' values.

Step 4: Continue the process up to the day the backward cumulative turnover cannot reach 100%.

It seems natural to select the time interval at which the turnovers accumulate to 100%, because it is one of the reasonable approximate periods during which all shares of the stock change hands. The main concerns for our method may consist of the following aspects. 1) The shares never traded in the target interval are not taken into account. We argue that these shares are usually not held by PT-MA investors, thus omitting them may not significantly affect our results. 2) Some shares having been traded for multiple times in the interval are repeatedly counted. Since PT-MA investors tend to sell their holdings when they have paper gains, the repeated count may not improve the results. 3) Our method does not adjust the turnovers as does in Grinblatt and Han (2005) in which higher weights are assigned to transactions happened more recently. However, due to lack of true reference prices, it is not easy to verify which method is more reasonable. In addition, we cannot see that our method will result in change in our findings in one direction.[8] In summary, this paper assumes that the daily prices and daily volumes in a target time interval form the distribution of reference prices at a target date.

Given the length of the time interval and the target date N [9], we construct proxy variables using trading volumes and average prices over the period. Before that, we need the relative capital gains, *RC*[10], for the shares of a stock purchased at date *n*, $1 \le n \le N$, which are

[7] It is worth noting that if another end date is selected, the start date changes and the length of the time interval also changes, so the letter S_T represents different numbers for different stocks and different end dates.

[8] We do not preclude other alternatives. We also take advantage of the adjusted turnovers of Grinblatt and Han (2005) and our flexible time intervals to construct proxy variables. The empirical results from those variables appear better than, but are consistent with what are reported in this paper.

[9] Thus, the start date of the time interval is 1 for target date *N*.

[10] The so-called "relative capital gains" are not the "return rates" of an investment, but the absolute capital gains normalized by the current price. Moreover, our methodology does not deny the legitimacy of other plausible options.

$$RC_n = \frac{AC_N - AC_n}{AC_N}. \tag{9}$$

where AC_n is the average price of the stock at date n. If a stockholder purchases a certain number of shares at date n at an average price lower than the market price, i.e., $AC_n < AC_N$, then $RC_n > 0$, and the holder possesses paper capital gains.

Now we can compute proxy variables for the reference price distribution. The formulas are listed as follows:

$$ARC_N = \frac{\sum_{n=1}^{N} VOL_n RC_n}{\sum_{n=1}^{N} VOL_n}, \qquad VRC_N = \frac{N \sum_{n=1}^{N} VOL_n \left(RC_n - ARC_N\right)^2}{(N-1)\sum_{n=1}^{N} VOL_n},$$
$$SRC_N = \frac{N \sum_{n=1}^{N} VOL_n (RC_n - ARC_N)^3}{\left(N-1\right) VRC_N^{3/2} \sum_{n=1}^{N} VOL_n}, \quad KRC_N = \frac{N \sum_{n=1}^{N} VOL_n (RC_n - ARC_N)^4}{\left(N-1\right) VRC_N^{2} \sum_{n=1}^{N} VOL_n}. \tag{10}$$

ARC_N, VRC_N, SRC_N, and KRC_N represent the mean, variance, skewness, and kurtosis of a reference price distribution at date N; VOL_n is the number of shares traded at date n. ARC_N is a volume-weighted average of relative capital gains, RC_n, over the target time interval. These four measures proxy for the variables in Equation (8).

We denote ARC_{it}, VRC_{it}, SRC_{it}, and KRC_{it} to be the values of the corresponding variables for stock i at date t. In our empirical investigation, the regression model becomes

$$r_{i,t+1} = h_0 + h_1 ARC_{it} + h_2 VRC_{it} + h_3 SRC_{it} + h_4 KRC_{it} + \varepsilon_i. \tag{11}$$

This regression model can capture the association between reference price distributions and stock returns in cross sections.

3.2 The data

The data used in this paper are from Shanghai Wind Information Co., Ltd (Wind). As a financial services company, Wind collects and sorts various financial data from the Chinese financial markets. Their products have been used by large securities companies, fund management companies, and investment institutions. We select daily trading records of all A-share stocks publicly traded on either the Shanghai Stock Exchange or the Shenzhen Stock Exchange. The data consist not only of trading information such as opening price, closing price, average price, volume (in number of shares or RMB), and turnover ratio, but also of several financial statement variables such as the earnings-price ratio and market-to-book ratio. The raw sample period, from April 1991 to March 2010, consists of 228 months.

We first construct the four proxy variables defined in the previous section using daily data, and then translate them to monthly data, according to rules appropriate to each specific variable. For example, the monthly return r_m is obtained with $\frac{P_{m,L} - P_{m-1,L}}{P_{m-1,L}} \times 100\%$, where $P_{m,L}$ is the closing price on the last trading day in month m [11]. The monthly turnover ratio is the sum of daily turnover ratios in that month, and the book-to-market ratio is the reciprocal of the market-to-book ratio on the last trading day of that month.

When transferring from daily to monthly data, we delete observations that fall within a month following a stock's IPO due to the dramatic jumps in stock prices that are usually seen[12]. We also exclude stocks in months where book values were negative, and those whose trading days were less than 15 days in the month. We also exclude data before January 1996 from the final monthly data, because there were too few stocks traded. Thus the final data used in our empirical analysis consist of 171 months, with 309 individual stocks in January 1996, and 1,426 by March 2010.

4. Empirical results

4.1 Summary statistics

Figure 2 plots the time series of the 10th, 50th, and 90th percentiles for the cross-section of ARCs and SRCs from January 1996 to March 2010, for Chinese firms listed on either the Shanghai Stock Exchange or the Shenzhen Stock Exchange. The bottommost dotted line denotes the 10th percentile, the solid line in the middle denotes the 50th percentile, and the dashed line on

[11] The prices and turnovers are adjusted according to dividends, splits and other issues.
[12] For a discussion on this issue, please refer to Wang and Xu (2004).

the top denotes the 90^{th} percentile. The left graph indicates that the ARCs display wide monthly variations over that period. And in most of those months, more than 50 percent of stocks had a negative ARC. This roughly shows that the investors indeed tend to sell their winning shares and hold on to losing ones. Most ARCs dropped dramatically in 2008 and 2009 when the Chinese stock markets experienced a bear market due to the world-wide financial crisis. The right graph shows that the SRCs across stocks also exhibit wide dispersions during the period. Most of the stocks have positive SRCs in most months, in contrast to their ARCs.

(Insert Figure 2 here.)

Panel A of Table 1 reports the summary statistics on the returns and the four proxy variables. Since Kahneman and Tversky's (1979) prospect theory suggests different utility functions for investors with negative and positive RCs, we divide the stocks into two groups for each month: G-group (the winning group, ARC > 0) and L-group (the losing group, ARC < 0). Since this is done for every month, a stock does not necessarily remain in the same group in different months. Within each group, we further divide the stocks into four ARC quartiles and calculate the equally-weighted average monthly returns and the simple averages of the ARCs, VRCs, SRCs and KRCs in each quartile. This process provides us with time-series averages of these five variables for each subgroup.

The average returns of the stocks in the two groups exhibit different patterns. The time series mean for stock returns in G-group is much greater than in L-group (3.11% vs. 0.85%). In L-group, the average return declines from the lowest ARC quartile at 1.43% through the highest ARC quartile at 0.56%. In G-group, it declines from the first quartile at 3.09% through the third quartile at 2.86%, and it reaches its highest value in the fourth quartile, at 3.46%. This pattern seems to contradict to prediction of Grinblatt and Han's (2005) model. However, this may be suggested by our model for two reasons. This first is that shareholders with negative capital gains are reluctant to sell their holdings regardless of their loss ratios. It is most likely that the selling probabilities for these investors are all zeros by the disposition effect. Therefore, the difference in ARCs for stocks in L-group may not predict variations in stocks' return. The second is that our model requires us to consider more moment variables besides ARC.

Different from the average monthly return, the means of VRCs and SRCs of stocks in G-group are significantly smaller than those in L-group. The patterns of VRCs and SRCs are also different between the two groups. The means of VRCs and SRCs of stocks in L-group decline with the ARC quartiles, whereas those in G-group exhibit a nonlinear pattern.

(Insert Table 1 here)

In Panel B of Table 1, stocks are first divided into two groups in the same way as in Panel A, i.e., a G-group and an L-group based on the signs of the stocks' ARCs. However, the quartiles are formed in terms of the individual stocks' SRCs, from low to high in each group. Again, the time series averages of equally-weighted monthly stock returns and the four proxy variables are reported. The average monthly return increases with the SRC quartiles in L-group, while there is no obvious pattern in G-group. This pattern is also suggested by our model. If the probabilities for investors with negative capital gains to sell their shares are zeros, the behaviors of investors with positive capital gains will play main role in predicting future stock returns. SRC just indicates the allocations of shares at the tails and around the median of a relative capital gains distribution. The association between SRCs and returns for stocks in L-group is consistent with our model. In addition for stocks in G-group, difference in SRCs may not account for sufficient difference in capital gains to predict variation in future stock returns.

The time series averages of the quartile ARCs decrease with SRC in both groups. This has also been observed in Panel A. The average VRC of L-group is greater than that of G-group, and there is no obvious pattern for KRC in either group.

Table 2 reports the summary statistics of several variables identified by the extant literature, including market capitalization, book-to-market ratio, average turnover in the past 12 months, and three types of past returns. The two panels in Table 2 apply the same grouping procedure as the corresponding panels in Table 1.

Market capitalization in Table 2 is measured by the logarithm of the market value of tradable A-shares, in RMB, denoted by ln(ME). Panel A shows that the average market capitalization of stocks in G-group is greater than in L-group. The average size (market capitalization) increases slightly with the ARC quartiles in G-group. These patterns are due to the stocks in G-group experiencing more appreciation on average over the past periods than the stocks in L-group.

The book-to-market ratio in Table 2 is the logarithm of book value over market value per share, or ln(BE/ME). As exhibited in Panel A, the simple average ln(BE/ME) displays consistent patterns with ARC, within and between L-group and G-group. The quartile average ln(BE/ME) descends monotonically from the low ARC in L-group, at -1.24, to the high ARC in G-group, at -1.78. These findings suggest that when stocks experience appreciations, ARC increases and BE/ME decreases.

In Table 2, the turnover ratio corresponding to month t is the average monthly turnover ratio over the previous 12 months, namely during the period from month t-1 to month t-12. Turnover is the trading share volume divided by the number of tradable A-shares in each month. Panel A presents interesting results for average turnover over the previous 12 months. First, the average turnover ratio of stocks in L-group, 43.42%, is much less than that of stocks in G-group, at 49.32%. Second, quartile average turnover goes up with ARC in L-group, but decreases in G-group. These results imply that trading activities of investors truly are related to their unrealized capital gains or losses. In particular, the trading volume is higher for stocks with ARC closer to zero.

We also consider the three types of past returns, which are short-term, medium-term, and long-term historical returns. At month t, the short-term past return r_{-1} is the stock's return in the past month, i.e., in month t-1. The medium-term past return, $r_{-4,-6}$, is the average monthly return during the three-month period from month t-4 to month t-6, and the long-term past return, $r_{-13,-24}$, during the twelve-month period from month t-13 to month t-24. The reasoning for choosing these three types of returns is based on relevant papers in this area. Jegadeesh (1990) and Lehmann (1990) show that contrarian investment strategies based on short-term return reversals of one week or one month can bring abnormal returns. Jegadeesh and Titman (1993) document significant profits for relative strength trading strategies over 3- to 12- month horizons. Long-term return reversals have also been identified by the related literature such as by De Bondt and Thaler (1985). The three types of prior returns in the paper aim to capture the three types of stock return patterns over different horizons. Our empirical investigation shows that long-term return reversals occur over relatively shorter periods in Chinese stock markets than in developed markets. This seems to be in accordance with the viewpoint that a higher proportion of investors are speculative in Chinese stock markets, and is supported by the average annual turnover ratio which can exceed 500%. The high-frequency trading activities in Chinese stock markets lead to relatively faster return reversals.

Panel A suggests that the return in the previous month, r_{-1}, and average return over the period from month t-4 to month t-6, $r_{-4,-6}$, are related to ARC. The time series average r_{-1} of stocks in L-group is -4.38% — much less than that of stocks in G-group, which have an average of 13.72%. Moreover, quartile averages of r_{-1} and $r_{-4,-6}$ both exhibit monotonically increasing patterns with ARC in both groups. This means that those stocks experiencing more appreciation during the previous half year have higher ARCs. On the other hand, $r_{-13,-24}$ does not display obvious patterns with ARC quartiles. However, the average $r_{-13,-24}$, 1.35%, for stocks in L-group, is significantly higher than the 0.72% for stocks in G-group. It indicates that long-term return reversals exist to some extent in the Chinese stock markets.

(Insert table 2 here.)

Panel B of Table 2 shows that both ln(ME) and ln(BE/ME) are related to SRC, but in opposite directions. The quartile with a higher SRC has a smaller ln(ME), but a higher ln(BE/ME) in both L-group and G-group. When compared to the results in Panel A, this suggests that ARC and SRC indeed capture different aspects of the impact of reference price distribution on cross-sectional returns. We observe an increasing pattern of average turnover with SRC quartiles from the low SRC column in L-group to the high SRC column in G-group.

As for the associations between SRC and the three types of past returns, we find that r_{-1} decreases from -3.18% in the low SRC quartile to -5.95% in the high SRC quartile in L-group, and from 16.50% to 12.00% in G-group. The average $r_{-4,-6}$ exhibits a different pattern with SRC in the groups. $r_{-4,-6}$ increases with the SRC quartile in L-group, while it displays a nonlinear association with SRC in G-group. As shown in Panel A of Table 2, SRC seems to be independent of $r_{-13,-24}$.

These summary statistics reveal that variations of cross-sectional returns are related to the proxy variables we construct in this paper, specifically the ARC and SRC of stockholders' relative capital gains. Meanwhile, our variables seem not to be disentangled from some factors identified in extant literature, indicating a need for further investigation.

4.2 Preliminary results

We use the Fama-MacBeth[13] (1973) method to analyze the relationship between stock returns and our proxy variables. Since the combination of four proxy variables rather than single variable can more precisely characterize a distribution, we include all proxy variables simultaneously as in Equation (11). We use r_{it}, the monthly return of stock i in month t, as a dependent variable, and the four proxy variables of mean, variance, skewness, and kurtosis of relative capital gains of stock *i*, on the last trading day of month *t*-1 as explanatory variables. ε_{t-1} is the error term containing the cross-sectional variation across stocks in month *t*, which is not captured by the four variables.

The regression preliminarily aims to explore the implications of our model. If prospect theory and mental accounting determine the utility valuations of investors in the Chinese stock markets, the distribution of investors' reference prices should play a role in stock returns. Thus, cross-sectional variation across stocks' returns should be explainable with proxy variables we construct. As described in the previous section, we divide the stocks into a G-group and an L-group based on the signs of their ARCs, given the asymmetry property of investors' risk attitudes.

We run cross-sectional OLS regressions of individual stock returns on the ARC, VRC, SRC, and KRC of relative capital gains, for each of the 171 months from January 1996 to March 2010, and obtain estimates for the 171 time series coefficients for each variable. The time series averages of the estimates and their t-statistics are reported in Table 3. The three panels in Table 3 present the results of the estimations with the full sample, the L-group sample and the G-group sample, respectively.

(Insert Table 3 here)

[13] Special thanks to Dayong Huang for suggesting the improved Fama-MacBeth regression. We following Mitchell A. Peterson's correction, and employ the Stata ado file xtfmb.ado provided by Daniel Hoechle for the Fama-MacBeth regressions throughout the empirical investigation in this paper.

Panel A of Table 3 reports estimation results with the full data. It suggests that ARC and SRC play significant roles in predicting the cross-section of stock returns, but that VRC and KRC do not. The estimated coefficient for ARC is positive and significant at the 5% level. A stock with a higher ARC has likely recently experienced a higher proportion of sell-initiated transactions than has a stock with a lower ARC. From our model, both aggregate selling probability and relative capital gains are higher for a stock with a higher ARC, leading to a higher future stock return. This result is consistent with the prediction of our theoretical framework. It suggests that a stock with more winning shares will perform better in the following period than a stock with fewer winning shares. As such, the results of ARC verify what Grinblatt and Han (2005) find. It confirms the hypothesis on coefficient h_1 in equation (8). The disposition effect underlies explanation to the finding.

The coefficient estimate of SRC is also positive and even more significant (at the 1% level), demonstrating that a stock with a long tail distribution to the right (or short left right tail) has a higher return in the next month when other variables are fixed. This empirical examination illustrates that in addition to ARC, SRC is also important in predicting stock returns. Given the ARC, VRC, and KRC, a higher positive SRC means that more shares hold very high relative capital gains (or less shares with high losses). An explanation similar to the one for ARC also applies here: with a higher SRC, a stock has a lower proportion of shares with capital gains close to the mean, but more shares hold large capital gains. Thus the aggregate selling probability is higher for the stock, forecasting a higher return in the near future, even if ARC holds unchanged. This result verifies our hypothesis on the coefficient h_3 of skewness in equation (8). This also provides evidence on our conjecture that reference price distribution rather than only average capital gains is critical to expected returns. However, we do not find that the proxy variables VRC and KRC hold explanatory abilities in the regressions[14].

In order to differentiate the effects of reference price distribution on winning and losing stocks, we repeat the process on stocks in G-group and L-group separately. The results are reported in Panels B and C of Table 3. ARC plays a key role in determining the returns of stocks in

[14]Our examination does not exclude the possibility that VRC and KRC are able to predict a stock's return over spans of other than one month. As a matter of fact, examinations not reported in the paper show that VRC and KRC can predict a stock's return over a one week period. These regression results are available upon request.

G-group, while SRC does so for stocks in L-group. For stocks in L-group, the estimated coefficient of ARC is insignificant, and likewise for stocks in G-group, SRC does not have explanatory ability.

For stocks in L-group, their average relative capital gains are negative, and difference in these stocks' ARCs just implies difference in holders' loss ratios. According to the disposition effect, if the selling probabilities for losing shares are zeros, the difference in ARCs may not lead to difference in selling probabilities. Our model predicts that in this situation, stock return may not necessarily vary with ARC. Therefore, even though these shareholders with negative capital gains have unsatisfied demands, stocks returns do not reflect the various demands due to constraint in supply. On the other hand, if we fix a negative ARC, an increase in SRC results in more large-profit shares (fewer large-loss shares) as well as more small-loss shares (fewer small-profit shares). If we assume the selling probabilities of losing shares are zeros, then only change in shares with positive capital gain determines the supply of the stock. Therefore, an increase in SRC leads to higher selling probabilities for more shares with positive capital gains[15], and finally higher stock returns. This finding can be explained by our model but not by Grinblatt and Han's (2005).

As shown in panel C of Table 3, for stocks with positive ARCs, of the four proxy variables ARC is the only significant explanatory variable. It is reasonable to observe that ARC plays its role in predicting future stock returns for stocks in G-group. The explanation is similar to what we obtain from the full data. The empirical result confirms the prediction of our model again, and provides us with a coefficient for ARC with a larger magnitude and higher significance level than obtained in the full data.

The insignificance of SRC in G-group shows that the effect of ARC on the returns of stocks with a positive ARC dominates that of SRC. Fixed a positive ARC, we consider variation in SRC for stocks with positive and negative SRC separately. If SRC is negative, a small increase in SRC leads to fewer shares with large losses as well as fewer shares with small profits. The changes in capital gains may not result in remarkable change in stock returns according to our model if the

[15] Notice that stocks in L-group most have positive SRCs. Therefore, a higher SRC implies increase in supplies of the stock by the disposition effect.

selling probabilities of shares with negative capital gains are all zeros[16]. If SRC is positive, a small increase in SRC leads to fewer shares with small profits as well as more shares with large profits. In this case, the decreasing number of shares with small profits is more than the increasing number of shares with large profits. Thus, the aggregate effect of increasing high-profit shares and decreasing low-profit shares may not significantly causes change in stock return.

The significance of SRC as opposed to ARC for stocks in L-group, and the opposite for those in G-group, provides evidence on the asymmetry property of prospect theory. It suggests that we should not only pay attention to the average capital gains of stockholders' investments, but also to the shape of the distributions formed by shareholders' relative capital gains.

5. Controlling for other factors

The literature on investigating cross-section of stock returns has identified a number of explanatory variables. Fama and French (1992) show that size and book-to-market ratio capture the cross-section of expected stock returns. Jegadeesh and Titman (1993) document that strategies of buying past well-performing stocks and selling poorly-performing stocks can bring investors significantly positive abnormal returns. In addition, some studies have postulated that cross-sectional stock returns decrease with stock turnover and/or volume (Datar et al., 1998; Hu, 1997; Rouwenhorst, 1998; and Chordia et al., 2001). In this section, we reassess the relationships explored in the previous section by controlling for factors that potentially have an impact on the returns.

5.1 Controlling for size and book-to-market ratio

Since size and book-to-market ratio are two well-known factors in explaining cross-section of stock returns, we add them into the regression model (11) to inspect the obtained results. The same as in the preliminary investigation, we run Fama-MacBeth regressions for the full data, then the two groups separately. For each sample, we run three regressions: The first includes only size (ln(ME)), the second includes only book-to-market ratio (ln(BE/ME)), and the third includes both

[16] Moreover, there is a possible decline in stock return due to fewer shares with small positive capital gains. This seems to receive some evidence from the negative coefficient of SRC for stocks in G-group.

as control variables. In order to be consistent with the four proxy variables[17], we use tradable market capitalization and book-to-market ratio from the beginning of each month in each monthly regression. The results of the regressions for the three groups are reported in the three panels of Table 4.

(Insert Table 4 here)

Panel A of Table 4 demonstrates that our proxy variables hold their ability to predict the cross-sectional variations in stock returns when we control for size effect. For the full sample, similar to the results without any control in the previous section, both ARC and SRC are positively associated with stock returns over the next period when market equity, ln(ME), is controlled. Both coefficient estimates are significant at the 10% level. For L-group and G-group, the results in Panel A show that market capitalization cannot eliminate the effects found in preliminary regressions. Moreover, ln(ME) loses its explanatory ability in G-group. As for L-group, the coefficient estimate for SRC is 0.49 and is significant at the 1% level; while the estimate for ARC is 8.20 and significant at the 5% level in G-group. As with the results reported in Table 3, ARC is insignificant for L-group, and SRC is insignificant for G-group. The insignificance of the size variable for stocks in G-group suggests that the relative strength of trading activities of PT-MA investors, who hold shares of large firms, are stronger than of those who hold shares of small firms. When a high enough proportion of shares are owned by winning stockholders, the strength of the behaviors of investors' with high risk aversion can offset the effect of the firm's market capitalization. To confirm this, we examine the association between firm size and ARC. Panel A of Table 2 indicates that the larger firms in G-group indeed have higher ARCs on average. But the association between SRC and firm size for stocks in L-group is negative as shown in Panel B of Table 2. The results at higher significance levels for both SRC and ln(ME) in L-group, are comparable to those for full data.

When controlling for only the book-to-market ratio as documented in panel B, the variables, which are able to explain the differences in stock returns still appear to be ARC and SRC. In the full sample, the coefficient estimate of ARC is 4.62 and is significant at the 5% level; that of SRC

[17] We also test our model using ln(ME) for June of each year, as well as ln(BE/ME) for December of the previous year, which are used in Fama and MacBeth (1973). The results are similar to those reported in this paper.

is 0.35 and also significant at the 5% level. The regression results show that SRC is the key variable in explaining the cross-sectional returns of stocks in L-group, and the effect of ARC is still prominent for stocks in G-group. These results justify our findings in the preliminary regressions. In addition, unlike size, the book-to-market ratio plays a significant role in determining the returns of stocks in all three groups, indicating that its explanatory power is independent of the proxy variables constructed in this paper.

Panel C of Table 4 reports the regression results when both size and book-to-market ratio are controlled for. It provides further evidence on the findings in the last section. Both ARC and SRC are significant at the 5% level in the full sample, and the coefficient estimate of SRC is 0.40 with a 5% significance level for L-group while ARC's coefficient estimate is 9.77 with a 1% significance level for G-group. An interesting finding in the full sample is that the coefficient estimate of VRC is positive and significant at the 5% level. If VRC contains information on the trading activities of PT-MA investors, it may be able to forecast stock returns. Given an ARC, a stock with a higher VRC means that more shares are distributed with deeper losses and wins. As discussed above, stockholders in deep losses are more likely to continue holding shares, and those in deep profits are more likely to sell their holdings. According to our model, a stock with a larger VRC may likely have higher returns over the following period.

5.2 Controlling for past returns

Grinblatt and Han (2005) apply prospect theory to explain the momentum effect of past returns, which implies that the disposition effect has a possible relationship with stocks' historical returns. In addition, when constructing our proxy variables, we use past stock prices from those periods where cumulative turnover ratios reach 100%. Previous papers, such as those by De Bondt and Thaler (1985, 1987), Jegadeesh (1990), Lehmann (1990), and Jegadeesh and Titman (1993), provide evidence that past returns can predict future returns. Three patterns of past returns have been commonly identified in the literature: long-term return reversals, medium-term return continuations, and short-term return reversals. The three variables, r_{-1}, $r_{-4,-6}$, and $r_{-13,-24}$, are used to capture these three patterns.

Table 5 presents the results of the Fama-MacBeth regressions controlled for the past returns of the three sample groups (the full sample, L-group and G-group). M1 denotes the regression model controlling for r_{-1}, with M2 controlling for $r_{-4,-6}$, and M3 controlling for all r_{-1}, $r_{-4,-6}$ and $r_{-13,-24}$.

Panel A of Table 5 presents the results for the full sample. The first and third rows show that returns for the past month, r_{-1}, negatively affect the expected stock return in both regressions M1 and M3. Including r_{-1} in the regressions does not remove the predictive abilities of ARC and SRC, although their coefficients and significance levels somewhat vary across the two settings. Controlling only for r_{-1} raises the coefficients' values and the significance levels for both ARC and SRC, while controlling for $r_{-4,-6}$ reduces them. Nevertheless, even in the latter case, ARC and SRC still remain significant at the 1% and 5% levels, respectively. As Grinblatt and Han (2005) document, the momentum effect appears to be closely related to prospect theory. However, in our framework, the momentum effect and the disposition effect are not completely equal. When all three types of past returns are controlled for, results do not change much, though there is a slight increase in the t-statistic for ARC's coefficient estimate.

Similar results are obtained for L-group. As shown in panel B of Table 5, the short-term return, r_{-1}, is always able to predict the cross-section of stock returns. A higher r_{-1} predicts a lower future return. Although the medium-term past return, $r_{-4,-6}$, reduces the explanatory power of SRC, the coefficient estimate of SRC remains significant at the 10% level with a value of 0.33. When all past return variables are controlled, SRC's coefficient becomes 0.58, and the t-statistic increases to 3.58. Panel B shows that for stocks in L-group, including past return variables does not change the effect of SRC. SRC still plays the key explanatory role in predicting the cross-section of returns for stocks in L-group. Another interesting finding is that the average return variable $r_{-13,-24}$ becomes significant when the three types of past returns are included, with a t-statistic of -2.29. This suggests that r_{-1} indeed captures short-term reversals, $r_{-3,-6}$ the medium-term continuations, and $r_{-13,-24}$ the long-term reversals in Chinese markets' stock returns.

Panel C of Table 5 shows that ARC exhibits a stronger effect on stocks in G-group. Regardless of which past return is controlled for, both coefficient estimates and significance levels are improved. For all the three regression models, the coefficient estimate is greater than 12, and the significance level is greater than 5%. However, in the regression with all three past returns included, coefficient estimates of all past return variables are not significant. This suggests that for stocks with a positive ARC, trading activities of investors subject to prospect theory and mental accounting determine the future performance of these stocks and the roles of past returns in affecting the expected return are eliminated. This further verifies our framework and provides evidence for the findings of Grinblatt and Han (2005). If we compare the results of the two M3 regressions for stocks in L-group and those in G-group, we may obtain further evidence on the existence of the disposition effect. Most holders of stocks in L-group possess losing shares and are reluctant to realize their losses, resulting in relatively lower trading frequency insufficient to eliminating the influence of past returns. For G-group stocks, however, high trading frequencies due to PT-MA investors result in the reference price distribution suggested by our model overshadowing the effects of past returns.

It is worth noting, nevertheless, that the impacts of the three past returns are still observed in some of our regressions. In most situations, the coefficient estimate $r_{-4,-6}$ is positive, and those of r_{-1} and $r_{-13,-24}$ are negative. This suggests that past returns and our proxy variables contain different information on the stock markets. Although not reported, the coefficients of $r_{-13,-24}$ are insignificant, and the paper's results are not affected in regressions for all sample groups where only $r_{-13,-24}$ is controlled for.

5.3 Controlling for turnover ratio

The four proxy variables constructed in this paper are based on turnover ratio and trading volume of a stock over previous trading days. Thus, skeptics may reasonably claim that these variables capture only the explanatory abilities of turnover ratio and/or volumes identified in the related literature. For example, Hu (1997) finds that in the Tokyo Stock Exchange, a stock with higher turnover has a lower expected return in a cross-sectional examination. We find this pattern in our data as well. Using the Fama-MacBeth method, we run regressions of stock returns on average turnover ratios over a 12-month period. The t-statistic for the coefficient estimate of the average turnover is -1.94, which is significant at the 10% level. In this subsection, we include turnover ratios in our estimation equation and examine whether it leads to any changes in our findings.

(Insert Table 6 here)

Table 6 presents regression coefficients with their t-statistics for the three sample groups. Using full data, we find that the coefficient estimate for ARC is insignificant when turnover ratio is included, while SRC remains significant with a t-statistic of 2.62. Although ARC loses its significance in the full sample, it is still a key variable in explaining the variations of stock returns in G-group. This shows that turnover cannot eliminate the effect of ARC on stocks in G-group. Panel B of Table 6 suggests that, for stocks in L-group, SRC is the primary factor indicating the cross-sectional differences in stock returns. Moreover, the coefficient estimate for the turnover ratio is insignificant. This shows that SRC is a better indicator than a stock's historical turnover to predicting its future returns, even though the coefficient estimate for the turnover is significant at the 10% level in cross-sectional regressions where turnover is the only regressor.

One way to account for the strength of ARC's effect is to investigate a stock's turnover in the subsequent period. If PT-MA investors are indeed prone to selling their winning shares, then the turnover ratio should be much higher for stocks with positive ARCs on average than for those with negative ARCs. A t-test, not shown in the paper, indicates that the average turnover for stocks in G-group is significantly greater than for those in L-group the following month. The average monthly turnover of stocks in G-group is 49.32, while in L-group it is 43.42. The t-statistic of the difference is 33.84.

Table 6 also shows a negative association between volume and expected stock return in the Chinese stock markets, consistent with findings in extant literature. These results suggest that turnover captures a different aspect of the stock market than do our proxy variables, particularly ARC and SRC, which hold their explanatory abilities in almost all multivariate regressions.

5.4 Comprehensive tests

In the previous subsection, we examined the validity of our model by separately controlling for different categories of factors potentially affecting our results. This subsection aims to investigate the joint effects of these factors on the coefficient estimates of our variables. Again, we test our model by running Fama-MacBecth regressions on the full sample first, and then on G-group and L-group separately. In each sample, four regression models, denoted as M1, M2, M3, and M4, include different combinations of control variables. Results are reported in Table 7.

(Insert Table 7 here)

Panel A of Table 7 indicates that SRC always possesses an explanatory ability for differences in cross-sectional stock returns with the full data, regardless of the combinations of control variables used. The estimates of ARC's coefficients are positive and significant in three of the four regression models, M1, M3, and M4. Result from model M2, in which market capitalization, book-to-market ratio, and turnover ratio are controlled for, shows that ARC seems not to be a critical variable to predicting stock returns. The insignificance of ARC in full data is mainly due to the inclusion of average turnover over the previous 12 months, as we have observed previously in Panel A of Table 6. However, the effect of ARC becomes more prominent in models M1 and M4 than is indicated in Table 3, where stock returns are only regressed on our four proxy variables. In addition, the coefficient estimates of SRC are positive and significant at 5% level or above for all regression models applied to the full sample.

The regression results for stocks in L-group are reported in panel B of Table 7. The coefficient estimates for SRC are around 0.4, and the significance levels are 5% or above for all four regression models. This demonstrates that among our four proxy variables, it is still SRC that plays the most important role in predicting returns in L-group. The coefficient estimates and significance levels for market capitalization, book-to-market ratio, past returns, and turnover ratio in Panel B are similar to those in Panel A.

When Fama-MacBeth regressions are employed only on stocks with positive ARCs, as indicated in Panel C of Table 7, the coefficient estimates for ARC are all positive, and more significant than those reported in Panel A for full data (significance levels are above the 1% level). However, in M4, in which all control variables are included, the variable VRC is significant and has a very large negative value. At the same time, all control variables except turnover ratio lose their significance. This shows that ARC plays a more important role in predicting returns of stocks in G-group than most of the control variables.

In summary, the main proxy variables with explanatory powers are ARC and SRC as obtained in previous section, even though various combinations of control variables are included in the regression for different sample groups. Additionally, the results also confirm a number of empirical findings from existing research; for example, the size effect and book-to-market effect present in the Chinese stock markets have similar patterns to those identified in developed financial markets. In addition, Table 7 illustrates that the coefficients of returns in the previous month, r_{-1}, and average monthly turnover over the previous 12 months are important predictors of stock returns. Moreover, their direct effects on stock returns seem independent of our four proxy variables.

Table 7 provides us with more confidence in the theory that the reference price distribution has predictive ability in cross-sectional stock returns. Based on analysis in this section and in previous sections, we believe that our model identifies important factors in forecasting future stock performance which have not been explored by existing literature. All tests provide evidence that our model and the proxy variables, specifically ARC and SRC, indeed have the ability to explain cross-section of stock returns. This also proves that the disposition effect, based on prospect theory and mental accounting, has a wide effect in the Chinese stock markets[18].

6. Conclusions

This paper develops a model based on the disposition effect that provides us with a channel through which we can investigate the relationship between reference price distributions and stock returns. By introducing extra demand from PT-MA investors and selling probability, we derive a formula for the expected return. When PT-MA investors have unrealized capital gains, their declining demand drives them to short their holdings, and when they have unrealized capital losses (negative capital gains), their increasing demand makes them reluctant to sell their shares. The aggregate effect of these investors' trading activities causes the stock price to respond slowly to exogenous news. This confirms observations by Odean (1998) and Shefrin and Statman (1985). The relative strength of winning and losing shares of a stock has implications for its expected return over the following period. For a given stock, the total effect of the PT-MA investors is determined by reference prices relative to the market price. Proxy variables constructed from market data characterize the distribution of reference prices, and enable us to examine the

[18] All regressions in the paper are repeated excluding the 1% outliers for every variable pertinent to each specific individual regression. Since similar results are obtained, we do not report them in the paper.

implications of our model.

By applying the Fama-MacBeth regressions on our data, we find that two of our four proxy variables, ARC and SRC, i.e., the mean and the skewness of the distribution of stockholders' relative capital gains, play important roles in cross-section of stock returns. Stock returns increase monotonically with ARC and SRC. This finding supports the predictions of our model based on the disposition effect. A positive ARC suggests that most stockholders possess winning shares. A higher positive SRC means that there are more investors with positive capital gains in the right tail of the distribution of stockholders' reference prices, and/or more investors have an extremely high paper profits. These investors' trading activities have a pervasive effect in the market, and determine the differences in stock returns.

In order to further explore the asymmetry of investor behavior and the relative strength of PT-MA stockholders' trading activities, we test our model conditional on the sign of stocks' ARC. Our empirical tests provide evidence on the existence of trading asymmetry when investors are in different situations. The subgroup regressions demonstrate that for stocks with a negative ARC, SRC is the main proxy variable explaining the cross-section of stock returns, and ARC is likewise the main proxy variable for stocks with positive ARCs. Our findings based on the subgroup regressions seem to support our model and are consistent with the disposition effect.

The empirical results are robust when factors potentially affecting our proxy variables as well as the cross-section of stock returns are controlled. More specifically, when market capitalization, book-to-market ratio, three types of past returns, and turnover ratio over the previous twelve months are included in our regressions, the significances of ARC and SRC estimates still hold. These analyses illustrate that our model captures variations of cross-sectional stock returns unexplained by other factors. Our model's proxies for stockholders' reference price distributions indeed explain a market-wide disposition effect resulting from investors' risk preferences in accordance with prospect theory and mental accounting, which are not eliminated by arbitrageurs.

References

Barberis, Nicholas and Ming Huang, (2008). Stocks as lotteries: the implications of probability weighting for security prices, *American Economic Review*, 98:5, 2066-2100.

Barberis, Nicholas, Ming Huang, and Tano Santos, (2001). Prospect Theory and Asset Prices, *Quarterly Journal of Economics*, 116(1): 1–53.

Barberis, Nicholas, Ming Huang, and Richard H. Thaler, (2006). Individual Preferences, Monetary Gambles, and Stock Market Participation: A Case for Narrow Framing. *American Economic Review*, 96(4): 1069–90.

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

Datar Vinay T, Narayan Y. Naik, and Robert Radcliffe, (1998). Liquidity and stock returns: An alternative test, *Journal of Financial Markets* 1, 203-219.

De Bondt, Werner F. M. and Richard Thaler, (1985), Does the stock market overreact? *Journal of Finance* 40, 793-805. ,

----(1987). Further evidence of investor overreaction and stock market seasonality, *Journal of Finance* 42, 557-581.

Ding, David K., Charlie Charoenwong, and Raymond Seetoh, (2004). Prospect theory, analyst forecasts, and stock returns, *Journal of Multinational Financial Management*, 14, 425-442.

Fama, Eugene F, and Kenneth R. French, (1992). The cross-section of expected returns, *The Journal of Finance*, Vol. XLVII, No. 2. 422-465.

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

Fiegenbaum, Avi, and Howard Thomas, (1988). Attitudes toward risk and the risk-return paradox: prospect theory explanations, *Academy of Management Journal*, Vol. 21, No. 1, 85-106.

Frazzini, Andrea, (2006). The disposition effect and underreaction to news, *The Journal of Finance*, Vol. 61, No, 4, pp. 2017-2046.

Grinblatt, Mark and Bing Han, (2005). Prospect theory, mental accounting, and momentum, *Journal of Financial Economics* 78, 311–339.

Hu, S.-Y., (1997). Trading turnover and expected stock returns: the trading frequency hypothesis and evidence from the Tokyo Stock Exchange. Working Paper, National Taiwan University.

Jegadeesh, Narasimhan, (1990). Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881-898.

Jegadeesh, Narashimhan and Sheridan and Titman, (1993). Returns to buying winners and selling losers: implication for stock market efficiency, *The Journal of Finance*, Vol. 48, No. 1, pp. 65-91.

Kaustia, Markku, (2004). Market-wide impact of the disposition effect: evidence from IPO trading volume, *Journal of Financial Markets* 7(2), 207-235.

Lee, Charles M. and Bhaskaran Swamination, (2000). Price Momentum and Trading Volume, *The Journal of Finance*, Vol. 55, No. 5 (Oct., 2000), pp. 2017-2069.

Lehmann, Bruce, (1990), Fads, martingales and market efficiency, *Quarterly Journal of Economics* 105, 1-28.

Odean, Terrance, (1998). Are investors reluctant to realize their losses? *The Journal of Finance*, Vol. 53, No. 5, pp. 1775-1798.

Peterson, Mitchell A, (2009), Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches, *Review of Financial Studies*, January, Volume 22, pp. 435-480.

Rouwenhorst, K.G., (1998). Local return factors and turnover in emerging stock markets, *The Journal of Finance*, Vol. LIV, No. 4, pp. 1439-1464.

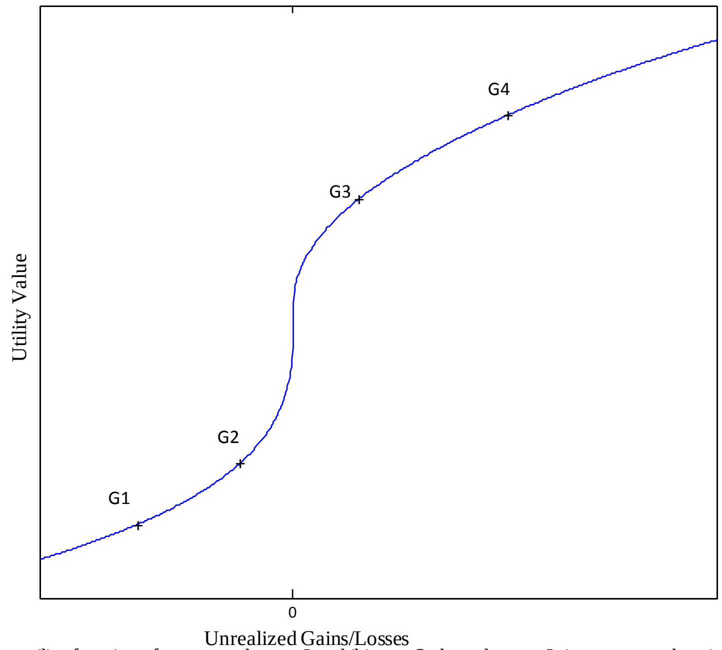
Shefrin, Hersh and Meir Statman, (1985). The disposition to sell winners too early and ride losers too long: theory and evidence, *The Journal of Finance*, Vol. XL, No. 3. 777-790.

Thaler, Richard , (1980). Toward a Positive Theory of Consumer Choice, *Journal of Economic Behavior and Organization*, 1 (March), 39-60.

Thaler, Richard, (1985). Mental accounting and consumer choice, *Marketing Science*, Vol. 4, No. 3, pp. 199-214.

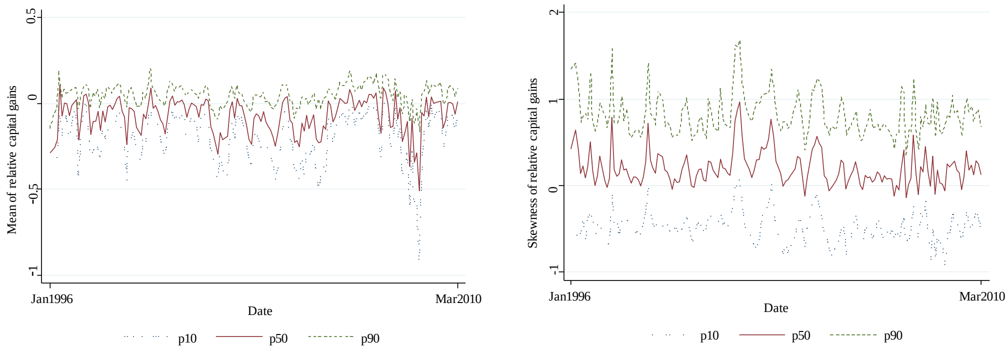
Weber, Martin and Colin F. Carmerer, (1998). The disposition effect in securities trading: an experimental analysis, *Journal of Economic Behavior & Organization*, Vol. 33, 167-184.

Figure 1: Utility function based on prospect theory



This figure plots utility function of prospect theory. It exhibits an S-shaped curve.It is concave when input is positive, and convex when input is negative. Positions G1 and G2 represent investors with capital losses, and positions G3 and G4 represent investors with capital gains.

Figure 2: Time series of cross-sectional percentiles of the means and skewness of relative capital gains



Left graph plots the time series of the 10th, 50th, and 90th percentiles of the cross-sectional mean of relative capital gains. Right graph plots the time series of the 10th, 50th, and 90th percentiles of the cross-sectional skewness of relative capital gains. The time period is from January 1996 to March 2010.

Table 1: Summary statistics of return and statistical variables

Panel A: time series average of statistical variables with stocks sorted on ARC (average of relative capital gains)

	L-group (Negative average capital gains)					G-group (Positive average capital gains)				
	All	Low	2	3	High	All	Low	2	3	High
Return	0.85	1.43	0.87	0.56	0.56	3.11	3.09	3.04	2.86	3.46
	[17.01]	[20.46]	[15.74]	[15.64]	[15.69]	[16.50]	[15.48]	[15.95]	[16.17]	[18.23]
ARC	-0.16	-0.29	-0.17	-0.11	-0.05	0.06	0.01	0.04	0.07	0.14
	[0.19]	[0.31]	[0.11]	[0.08]	[0.05]	[0.06]	[0.01]	[0.02]	[0.02]	[0.07]
VRC	0.03	0.06	0.02	0.01	0.01	0.01	0.004	0.01	0.01	0.01
	[0.32]	[0.59]	[0.23]	[0.03]	[0.03]	[0.17]	[0.01]	[0.34]	[0.01]	[0.01]
SRC	0.20	0.31	0.23	0.15	0.09	0.10	0.19	0.23	0.14	-0.15
	[1.15]	[1.30]	[1.00]	[1.25]	[1.02]	[0.61]	[0.68]	[0.48]	[0.59]	[0.59]
KRC	3.70	4.01	3.17	3.75	3.87	2.73	3.67	2.42	2.43	2.40
	[145.02]	[155.80]	[108.53]	[156.80]	[153.30]	[69.66]	[138.37]	[1.56]	[17.73]	[6.83]

Panel B: time series averages of statistical variables with stocks sorted on SRC (skewness of relative capital gains)

	L-group (Negative average capital gains)					G-group (Positive average capital gains)				
	All	Low	2	3	High	All	Low	2	3	High
Return	0.85	0.31	0.72	1.14	1.25	3.11	2.95	3.17	3.17	3.17
	[17.01]	[15.71]	[15.73]	[20.30]	[15.82]	[16.50]	[16.98]	[16.60]	[16.04]	[16.34]
ARC	-0.16	-0.14	-0.15	-0.16	-0.18	0.06	0.09	0.07	0.05	0.05
	[0.19]	[0.17]	[0.16]	[0.17]	[0.25]	[0.06]	[0.08]	[0.06]	[0.04]	[0.04]
VRC	0.03	0.03	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01
	[0.32]	[0.34]	[0.31]	[0.24]	[0.37]	[0.17]	[0.01]	[0.01]	[0.01]	[0.34]
SRC	0.20	-0.52	0.05	0.38	0.86	0.10	-0.51	-0.03	0.26	0.69
	[1.15]	[2.01]	[0.27]	[0.27]	[0.39]	[0.61]	[0.72]	[0.25]	[0.20]	[0.32]
KRC	3.70	7.23	2.11	2.30	3.17	2.73	3.87	2.12	2.17	2.77
	[145.02]	[290.42]	[0.74]	[0.61]	[1.74]	[69.66]	[139.63]	[3.13]	[0.92]	[1.24]

Summary statistics of return and statistical variables for stocks quartiles sorting on ARC and SRC. In panel A, in each month, stocks are divided into two groups in terms of ARC: stock with negative ARC to L-group and stock with positive ARC to G-group, then in each group, stocks are further sorted on values of their ARCs, and quartiles are formed from low to high ARC in each group. Then calculate the time series mean of the equal-weighted averages of monthly returns and of four statistical variables (ARC, VRC, SRC and KRC) in each quartile.

In panel B, in each month, stocks are divided into two groups in terms of ARC: stock with negative ARC to L-group and stock with positive ARC to G-group, then in each group, stocks are further sorted on values of their SRCs, and quartiles are formed from low to high SRC in each group. Then calculate the time series mean of the equal-weighted averages of monthly returns and of four statistical variables (ARC, VRC, SRC and KRC) in each quartile.

t-satistics are reported in brackets.

Table 2: Summary statistics of several extant factors

Panel A: time series average of variables with stock groups sorted on ARC

	L-group (Negative average capital gains)					G-group (Positive average capital gains)				
	All	Low	2	3	High	All	Low	2	3	High
ln(ME)	20.51	20.52	20.45	20.48	20.58	20.86	20.70	20.76	20.85	21.14
	[1.06]	[1.14]	[1.02]	[1.01]	[1.05]	[1.12]	[1.03]	[1.05]	[1.09]	[1.25]
ln(BE/ME)	-1.32	-1.24	-1.27	-1.33	-1.45	-1.57	-1.47	-1.49	-1.55	-1.78
	[0.76]	[0.78]	[0.74]	[0.74]	[0.75]	[0.79]	[0.76]	[0.76]	[0.77]	[0.83]
Turnover	43.42	40.29	43.78	44.59	45.26	49.32	51.96	51.37	50.00	44.19
	[33.49]	[33.02]	[34.03]	[33.17]	[33.53]	[31.34]	[32.27]	[32.30]	[31.19]	[29.00]
r_{-1}	-4.38	-8.72	-5.68	-3.41	0.31	13.72	8.17	10.67	14.60	21.34
	[12.02]	[12.10]	[11.75]	[11.02]	[11.34]	[18.86]	[22.18]	[12.86]	[15.15]	[20.90]
$r_{-4,-6}$	0.69	0.31	0.56	0.71	1.19	3.17	2.74	2.98	3.14	3.80
	[9.95]	[10.35]	[10.93]	[9.04]	[9.30]	[9.95]	[9.61]	[9.93]	[10.09]	[10.12]
$r_{-13,-24}$	1.35	1.42	1.31	1.30	1.38	0.72	0.70	0.60	0.66	0.89
	[5.38]	[5.67]	[5.40]	[5.23]	[5.17]	[4.87]	[4.85]	[4.81]	[4.90]	[4.93]

Panel B: time series average of variables with stock groups sorted on SRC

	L-group (Negative average capital gains)					G-group (Positive average capital gains)				
	All	Low	2	3	High	All	Low	2	3	High
ln(ME)	20.51	20.57	20.52	20.48	20.45	20.86	20.99	20.88	20.81	20.77
	[1.06]	[1.12]	[1.06]	[1.03]	[1.01]	[1.12]	[1.21]	[1.12]	[1.08]	[1.05]
ln(BE/ME)	-1.32	-1.35	-1.32	-1.31	-1.30	-1.57	-1.67	-1.56	-1.53	-1.53
	[0.76]	[0.80]	[0.76]	[0.73]	[0.74]	[0.79]	[0.82]	[0.79]	[0.77]	[0.78]
Turnover	43.42	41.14	42.81	43.91	45.83	49.32	47.22	49.24	49.84	50.94
	[33.49]	[32.45]	[32.83]	[33.31]	[35.16]	[31.34]	[30.88]	[30.99]	[30.93]	[32.43]
r_{-1}	-4.38	-3.18	-3.73	-4.65	-5.95	13.72	16.50	13.98	12.44	12.00
	[12.02]	[11.67]	[11.82]	[11.57]	[12.81]	[18.86]	[18.98]	[24.17]	[15.45]	[15.06]
$r_{-4,-6}$	0.69	0.43	0.58	0.75	0.99	3.17	3.46	3.15	3.01	3.06
	[9.95]	[9.39]	[10.13]	[9.14]	[11.02]	[9.95]	[10.03]	[9.89]	[9.90]	[9.97]
$r_{-13,-24}$	1.35	1.45	1.38	1.29	1.30	0.72	0.86	0.68	0.62	0.70
	[5.38]	[5.46]	[5.39]	[5.37]	[5.30]	[4.87]	[4.91]	[4.83]	[4.87]	[4.89]

Summary statistical of several variables based on sorting on average relative capital gains and skewness of average relative capital gains. In panel A, in each month, stocks are divided into two groups in terms of average relative gains: negative and positive, then in each group, stocks are further sorted on mean of relative capital gains, and four subgroups are formed equally from low to high mean in each group. Then calculate the time series means of the natural logarithm of market capitalizations, the natural logarithm of book-to-market ratios, average turnover ratio of past twelve months, and three types of past returns in each subgroup.

In panel B, in each month, stocks are divided into two groups in terms of average relative gains: negative and positive, then in each group, stocks are further sorted on skewness of relative capital gains, and four subgroups are formed equally from low to high skewness in each group. Then calculate the time series mean of the natural logarithm of market capitalizations, the natural logarithm of book-to-market ratios, average turnover ratio of past twelve months, and three types of past returns in each subgroup.

t-satistics are reported in brackets.

Table 3: Cross-sectional regressions of stock returns on four statistical variables

	ARC	VRC	SRC	KRC	intercept
Panel A: Full data					
	3.43	1.33	**0.40**	-0.09	**1.88**
	[1.98]	[0.15]	[3.08]	[-1.26]	[2.23]
Panel B: Negative average capital gains (negative ARC)					
	-0.54	17.76	**0.59**	-0.07	1.58
	[-0.09]	[0.79]	[3.7]	[-0.93]	[1.78]
Panel C: Positive average capital gains (positive ARC)					
	9.51	-20.44	-0.36	0.57	0.48
	[2.67]	[-0.61]	[-0.52]	[-1.17]	[0.36]

Data is from Shanghai Stock Exchange and Shenzhen Stock Exchange. Statistical variables are calcualted from daily market trading data, and then monthly data containing relavent variables are derived from daily data. The sample period is from January 1996 to March 2010. For each sample group, the monthly stock returns are cross-sectionally regressed on four statistical variables by Fama-MacBeth (1973) method. The ARC, VRC, SRC, and KRC are the mean, variance, skewness, and kurtosis of relative capital gains. The full sample panel is results of regressions including all observations. Panel A consists of stocks with negative ARC, and Panel B consists of those with positive ARC.

t-statistics are in brackets. Bold numbers denote siginificant estimate at the 5% level.

Table 4: Cross-sectional regressions with controlling for size and/or book-to-market ratio

	ARC	VRC	SRC	KRC	ln(ME)	ln(BE/ME)
Panel A: control for size						
Full data	2.72	9.67	**0.31**	-0.06	-0.34	
	[1.79]	[1.27]	[2.51]	[-0.77]	[-1.67]	
L-group	-2.20	20.74	**0.49**	-0.01	**-0.55**	
	[-0.37]	[1.01]	[3.19]	[0.11]	[-2.48]	
G-group	**8.20**	-18.05	0.26	0.05	-0.08	
	[2.42]	[-0.58]	[0.88]	[0.19]	[-0.35]	
Panel B: Control for book-to-market ratio						
Full data	**4.62**	5.84	**0.35**	-0.07		**0.74**
	[2.68]	[0.68]	[2.97]	[-1.01]		[2.88]
L-group	-0.58	19.47	**0.53**	-0.04		**0.94**
	[-0.10]	[0.86]	[3.39]	[-0.53]		[3.50]
G-group	**11.75**	-22.30	-0.37	0.61		**0.62**
	[3.48]	[-0.65]	[0.55]	[1.26]		[2.27]
Panel C: Control for both size and book-to-market ratio						
Full data	**3.64**	15.23	**0.25**	-0.03	-0.35	**0.63**
	[2.41]	[2.14]	[2.26]	[-0.50]	[-1.75]	[2.47]
L-group	-2.60	23.09	**0.40**	0.02	**-0.56**	**0.87**
	[-0.45]	[1.12]	[2.69]	[0.32]	[-2.61]	[3.26]
G-group	**9.77**	-17.07	0.22	0.09	-0.07	0.43
	[2.94]	[-0.54]	[0.76]	[0.33]	[-0.28]	[1.60]

Fama-MacBeth regressions are applied to variants of our primary regression model. Panel A exhibits the results when controling for stocks' market capitalization. Market capitalization, ME, is the market value of tradable shares of stocks traded in Chinese stock market. Panel B controls for book-to-market ratio, ln(BE/ME), and panel C controls both size and book-to-market ratio. In each panel, regressions are applied to three groups of observations, full data, L-group and G-group.

t-statistics are in brackets. Bold numbers denote siginificant estimate at the 5% level.

Table 5: Cross-sectional regressions with controlling for past returns

	ARC	VRC	SRC	KRC	r_{-1}	$r_{-4,-6}$	$r_{-13,-24}$
Panel A: Full data							
M1	**6.21**	3.05	**0.43**	-0.11	**-0.04**		
	[3.83]	[0.34]	[3.57]	[-1.49]	[-3.15]		
M2	2.84	2.04	**0.29**	-0.11		**0.04**	
	[1.67]	[0.23]	[2.27]	[-1.31]		[2.56]	
M3	**4.50**	9.23	**0.34**	-0.04	**-0.04**	0.03	-0.06
	[2.75]	[0.75]	[2.95]	[-0.49]	[-3.55]	[1.72]	[-1.59]
Panel B: L-group							
M1	2.65	24.95	**0.52**	-0.10	**-0.04**		
	[0.49]	[1.00]	[3.35]	[-1.13]	[-2.88]		
M2	-1.39	8.53	0.33	-0.09		**0.04**	
	[-0.24]	[0.40]	[1.94]	[-1.03]		[2.25]	
M3	2.14	38.60	**0.58**	0.07	**-0.04**	0.03	**-0.11**
	[0.36]	[1.50]	[3.58]	[0.56]	[-2.27]	[1.14]	[-2.29]
Panel C: G-group							
M1	**12.57**	-3.38	0.37	0.03	**-0.05**		
	[3.56]	[-0.10]	[1.23]	[0.12]	[-2.43]		
M2	**12.02**	-2.86	0.47	0.07		-0.02	
	[3.19]	[-0.08]	[1.23]	[0.29]		[-0.73]	
M3	**16.17**	-38.07	0.96	0.41	-0.02	0.02	0.16
	[4.04]	[-0.84]	[1.33]	[0.42]	[-0.85]	[0.16]	[0.55]

Fama-MacBeth regressions are applied to variants of our primary regression model. Panel A presents the results of regressions for the full data with controlling for three types of the past returns, panel B those for the L-group, and panel C those for the G-group. In each panel, three variant regressions are employed. M1 controls for only the return in last month, r_{-1} M2 controls for the average monthly return during the period from month t-4 to month t-6, $r_{-4,-6}$, and M3 controls for all three types of past returns, r_{-1}, $r_{-4,-6}$, and $r_{-13,-24}$.

t-statistics are in brackets. Bold numbers denote siginificant estimate at the 5% level.

Table 6: Cross-sectional regressions of stock retuns with turnover ratio

	ARC	VRC	SRC	KRC	Turnover
Panel A: Full data					
	2.12	-20.48	**0.34**	-0.12	**-0.14**
	[1.20]	[-1.52]	[2.62]	[-1.33]	[-2.01]
Panel B: L-group					
	-3.01	-1.56	**0.44**	-0.08	-0.01
	[-0.52]	[-0.08]	[2.54]	[-0.87]	[-1.24]
Panel C: G-group					
	11.4	-37.9	0.71	0.18	**-0.02**
	[2.76]	[-1.03]	[1.79]	[0.71]	[-2.10]

Fama-MacBeth regressions are applied to variants of our primary regression model. Panel A is the regression of full data with turnover ratio controlled, panel B is that of stocks in L-group, and panel C is that of stocks in G-group.

Table 7: Cross-sectional regressions with jointly controlling for different combinations of factors

	ARC	VRC	SRC	KRC	ln(ME)	ln(BE/ME)	r_{-1}	$r_{-4,-6}$	$r_{-13,-24}$	Turnover
Panel A: Full sample										
M1	**5.72**	**22.91**	**0.26**	-0.01	-0.36	**0.74**	**-0.05**	**0.04**	-0.01	
	[4.00]	[2.31]	[2.27]	[-0.09]	[-1.80]	[2.45]	[-4.58]	[2.11]	[-0.43]	
M2	2.06	-5.33	**0.24**	-0.07	**-0.40**	0.62				**-0.02**
	[1.44]	[-0.57]	[2.03]	[-0.80]	[-2.00]	[1.91]				[-3.28]
M3	**3.90**	-1.91	**0.32**	-0.05			**-0.04**	0.03	-0.07	-0.01
	[2.49]	[-0.15]	[2.76]	[-0.58]			[-3.62]	[1.66]	[-1.76]	[-1.76]
M4	**4.74**	6.93	**0.26**	-0.04	**-0.43**	**0.66**	**-0.04**	**0.04**	-0.01	**-0.02**
	[3.51]	[0.62]	[2.33]	[-0.45]	[-2.20]	[2.20]	[-4.40]	[2.60]	[-0.46]	[-3.64]
Panel B: L-group										
M1	0.26	45.23	**0.49**	0.16	**-0.62**	**0.95**	**-0.05**	**0.04**	-0.05	
	[0.04]	[1.82]	[2.85]	[1.24]	[-2.67]	[2.95]	[-2.72]	[2.08]	[-1.17]	
M2	-5.16	4.97	**0.33**	0.01	**-0.70**	**1.01**				**-0.02**
	[-0.86]	[0.28]	[2.00]	[0.08]	[-3.00]	[3.11]				[-2.66]
M3	-0.48	14.74	**0.45**	0.01			**-0.04**	0.02	**-0.12**	-0.01
	[-0.09]	[0.63]	[2.87]	[0.07]			[-2.91]	[0.95]	[-2.63]	[-1.58]
M4	-2.10	17.51	**0.39**	0.07	**-0.69**	**0.91**	**-0.05**	**0.05**	-0.05	**-0.02**
	[-0.35]	[0.83]	[2.76]	[0.60]	[-2.99]	[2.90]	[-3.47]	[2.41]	[-1.34]	[-3.23]
Panel C: G-group										
M1	**18.14**	-39.50	1.06	-0.32	-0.20	**0.90**	-0.04	-0.07	-0.04	
	[3.92]	[-0.80]	[1.33]	[-0.56]	[-0.63]	[2.31]	[-1.49]	[-0.95]	[-0.53]	
M2	**12.17**	-37.07	0.49	0.16	-0.17	**0.91**				**-0.03**
	[3.21]	[-1.01]	[1.28]	[0.75]	[-0.69]	[2.26]				[-2.20]
M3	**15.74**	-74.21	0.92	-0.57			-0.06	-0.05	-0.09	**-0.04**
	[4.09]	[-1.66]	[1.30]	[-1.15]			[-1.68]	[-0.08]	[-1.25]	[-2.78]
M4	**18.75**	**-132.64**	1.72	-1.05	-0.37	1.71	-0.05	0.01	-0.03	**-0.05**
	[3.94]	[-2.23]	[1.52]	[-1.33]	[-1.31]	[1.80]	[-1.23]	[0.07]	[-0.29]	[-2.95]

Fama-MacBeth regressions are applied to variants of our primary regression model. Joint effects of control variables are considered. Panel A reports the results of the regressions of full data with four types of combinations of control variables; panel B reports those of stocks group with nagative average capital gains; and panel C reports those of stocks group with positive average capital gains. In each panel, M1 controls for the combination of market capitalization, which is logarithm of the market values of tradable shares of stocks traded in Chinese stock markets, logarithm of the book-to-market ratio and three categories of past returns. M2 controls for market capitalization, book-to-market ratio and turnover ratio, which is the montly average turnover ratio of the past twelve months. M3 controls for the three categories of past returns and turnover ratio. M4 for controls all control variables.
t-statistics are reported in brackets. Bold numbers denote siginificant estimate at the 5% level.