Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Financial Economics 69 (2003) 355–373
JOURNAL OF Financial ECONOMICS
www.elsevier.com/locate/econbase

Arbitrage risk and the book-to-market anomaly[☆]

Ashiq Ali[a,*], Lee-Seok Hwang[b], Mark A. Trombley[a]

[a] *Eller College of Business, Public Administration, University of Arizona, Tucson, AZ 85721, USA*
[b] *College of Business Administration, Seoul National University, Seoul, South Korea*

Received 24 August 2001; received in revised form 4 April 2002; accepted 21 March 2003

Abstract

This paper shows that the book-to-market (*B*/*M*) effect is greater for stocks with higher idiosyncratic return volatility, higher transaction costs, and lower investor sophistication, consistent with the market-mispricing explanation for the anomaly. The *B*/*M* effect for high volatility stocks exceeds that for the low volatility stocks in 20 of the 22 sample years. Also, volatility exhibits significant incremental power beyond transaction costs and investor sophistication measures in explaining cross-sectional variation in the *B*/*M* effect. These findings are consistent with the Shleifer and Vishny (1997) thesis that risk associated with the volatility of arbitrage returns deters arbitrage activity and is an important reason why the *B*/*M* effect exists.

JEL classification: G11; G14

Keywords: Arbitrage risk; Book-to-market; Mispricing; Transaction costs; Investor sophistication

1. Introduction

Numerous studies show predictable returns over three to five years for portfolios long in high book-to-market (*B*/*M*) stocks and short in low *B*/*M* stocks

[☆] The comments of Bill Horace, George Jiang, Chris Lamoureux, the workshop participants at Duke University, the University of Oregon, and an anonymous referee are gratefully acknowledged. The analyst estimates data used in this study were provided by I/B/E/S under its program to encourage research on earnings expectations. Professors Ali and Trombley acknowledge financial support provided by the University of Arizona Eller College of Business and Public Administration Faculty Development Fund and by Ernst and Young.

*Corresponding author. Tel.: +1-520-621-3765.
E-mail address: ashiq@u.arizona.edu (A. Ali).

0304-405X/03/$ - see front matter
doi:10.1016/S0304-405X(03)00116-8

(e.g., Rosenberg et al., 1984; Fama and French, 1992; Lakonishok et al., 1994). Two competing explanations for this exist. First, the return to *B/M*-based portfolio strategies represents compensation for risk, as suggested by Fama and French (1992, 1993, 1997). Second, the return to *B/M*-based portfolio strategies results from systematic mispricing of extreme *B/M* securities. Studies supporting the mispricing explanation show that market participants underestimate future earnings for high *B/M* stocks and overestimate future earnings for low *B/M* stocks (La Porta et al., 1997; Skinner and Sloan, 2002).[1] If the *B/M* effect represents mispricing due to systematic bias in expectations, then why don't professional arbitrageurs exploit this opportunity and quickly eliminate the mispricing? Shleifer and Vishny (1997) argue that arbitrage is costly and any systematic mispricing would not be quickly and completely traded away in situations where arbitrage costs exceed arbitrage benefits. They further argue that risk due to the volatility of arbitrage returns (hereafter arbitrage risk) deters arbitrage activity and is likely to be an important reason why the *B/M* effect exists. Our study empirically examines their prediction and in doing so provides additional evidence to discriminate between the risk and mispricing explanations for the *B/M* effect.

Arbitrage resources are concentrated in the hands of a relatively few specialized and poorly diversified traders. These arbitrageurs are risk averse and are concerned about the idiosyncratic risk of their portfolios. Thus, Shleifer and Vishny (1997) predict that volatility will deter arbitrage activities. If the *B/M* effect is due to mispricing, then according to Shleifer and Vishny this effect should be greater for stocks with higher expected volatility. We use the historical volatility of the market model residuals (hereafter referred to as just volatility) as a proxy for idiosyncratic risk and find that the *B/M* effect is greater for stocks with greater volatility, consistent with the mispricing explanation.

We also find that the *B/M* effect is greater for stocks with higher transaction costs and stock with less ownership by sophisticated investors, providing further evidence that the *B/M* effect is due to market mispricing. However, relative to these factors, the relationship of return volatility with the *B/M* effect is both stronger and incremental, suggesting that arbitrage risk is an important reason for the existence of the *B/M*-related mispricing.

Our study makes the following contributions to the literature. Lakonishok et al. (1994) predict that returns to *B/M* strategies are smaller for stocks of larger companies because arbitrage costs and investor sophistication affect the existence of mispricing and firm size proxies for these factors. However, La Porta et al. (1997) report that the *B/M* effect is not much different across small and large firms, raising concern whether mispricing is a valid explanation. Using more direct measures of arbitrage costs and investor sophistication, we find that when arbitrage is costly and

[1] Two other explanations are proposed in the literature. First, Kothari et al. (1995) argue that part of the returns to high *B/M* portfolios results from survivorship bias in the Standard and Poor's Compustat database. However, Chan et al. (1995) provide evidence that refutes the claim. Second, Lo and MacKinlay (1988) suggest the data-snooping explanation. However, the existence of the *B/M* effect in other countries (e.g., Chan et al., 1991) makes this explanation unlikely.

investor sophistication is low, the *B/M* effect is high, providing support to the mispricing explanation for the *B/M* anomaly.

Prior empirical studies attempting to explain the existence of systematic mispricing on the basis of arbitrage costs focus mainly on transaction costs, and not much attention is given to arbitrage risk.[2] Some studies on short-term trading strategies find that transaction costs are greater than or equal to the magnitude of the mispricing (e.g., Knez and Ready, 1996; Barber et al., 2001), and conclude that transaction costs provide a sufficient explanation for the existence of mispricing. However, for a long-term strategy like *B/M*, the cost associated with arbitrage risk, which increases with the length of the period the position is held, is likely to be much more than transaction costs. Our results are consistent with this argument.

The remainder of the paper is organized as follows. Section 2 describes the risk and mispricing explanations for the book-to-market effect, explains how arbitrage risk affects mispricing, and reviews various measures of transaction costs and investor sophistication identified in the literature. Section 3 describes the data and presents empirical results. Section 4 concludes the paper.

2. Alternative explanations of the book-to-market effect

Fama and French (1993, 1997) interpret the return to *B/M* as compensation for state-dependent risk related to relative financial distress. They present evidence that industry-specific variation in returns to *B/M* corresponds with periods of industry strength or distress. However, by showing that bankruptcy risk is not related to future returns, Dichev (1998) refutes the financial distress explanation for the *B/M* effect. La Porta et al. (1997) and Skinner and Sloan (2002) argue that the *B/M* effect is due to mispricing. They show that for high (low) *B/M* securities, market participants underestimate (overestimate) future earnings, and that stock price reactions to future earnings announcements of extreme *B/M* stocks are consistent with the correction of the systematically biased expectations. Below, we discuss factors that are likely to be cross-sectionally correlated to the *B/M* effect, if the effect is due to mispricing.

2.1. Arbitrage risk

Shleifer and Vishny (1997) argue the importance of arbitrage risk in the existence of mispricing related to *B/M*. They note that arbitrage resources are concentrated in

[2] Transaction costs have been used to explain several asset-pricing anomalies. They include the small-firm effect (Stoll and Whaley, 1983), the January effect (Reinganum, 1983; Bhardwaj and Brooks, 1992), the post-earnings announcement drift (Bhushan, 1994), closed-end fund discounts (Pontiff, 1996), expected price improvement from a switching strategy (Knez and Ready, 1996), analyst recommendation underreaction (Copeland and Mayers, 1982; Barber et al., 2001), and momentum profits (Lesmond et al., 2001). Only Pontiff (1996) and Wurgler and Zhuravskaya (2002) explicitly consider arbitrage risk to explain asset-pricing anomalies. They do so in the context of the closed-end fund discounts and the inclusion of a stock in the S&P 500 index.

the hands of a relatively few specialized and poorly diversified traders. Arbitrageurs carefully analyze each stock they invest in, a costly process, and hence include only a limited number of stocks in their arbitrage portfolios. The risk-averse arbitrageurs are concerned about the idiosyncratic risk of their portfolios. Thus, Shleifer and Vishny predict that volatility will deter arbitrage activities.

Shleifer and Vishny (1997) further argue that after arbitrageurs take positions in mispriced securities, noise traders can push prices further away from fundamentals before disconfirming evidence becomes available. They note that over a one-year horizon, a long position in a diversified portfolio of high *B/M* stocks outperforms the S&P 500 only about 60% of the time, even though superior performance over a five-year horizon is much more likely. But arbitrageurs care more about the short-run performance, because they use capital provided by investors, who tend to withdraw funds if the short-run performance is poor (Ippolito, 1992; Warther, 1995). Arbitrageurs desire to keep the ratio of reward-to-risk over shorter horizons high. This behavior further deters arbitrage activity in high volatility stocks, and as a result these stocks exhibit greater mispricing.

To specialized arbitrageurs, idiosyncratic volatility of the stocks in their portfolio is of greater concern than systematic volatility. For systematic volatility, arbitrageurs get compensated, or alternatively, they can eliminate it by hedging. On the other hand, idiosyncratic volatility cannot be hedged. Also, since arbitrageurs are not well diversified, idiosyncratic volatility adds to total portfolio volatility, without a corresponding increase in expected returns.

In deciding which mispriced stocks to take positions in, arbitrageurs would consider the expected idiosyncratic stock volatility during the holding period. To obtain a measure of expected idiosyncratic volatility, we regress daily returns on a value-weighted market index over a one-year period immediately preceding the holding period and compute the variance of the residual term.[3]

2.2. Transaction costs

When securities are mispriced, transaction costs limit the extent to which investors aware of the mispricing can take advantage of it and eliminate it. Securities with higher transaction costs are therefore likely to exhibit greater residual mispricing. Prior studies discuss measures for three types of transaction costs: direct transaction costs, indirect transaction costs, and costs associated with short selling.

Direct transaction costs include bid-ask spreads and brokerage commissions. Bhardwaj and Brooks (1992) and Blume and Goldstein (1992) suggest that quoted bid-ask spreads and commission per share as a percentage of share price are inversely related to share price. Thus, we use share price as a measure of direct transaction costs. We also use historical bid-ask spread as an additional measure of direct

[3] We repeat the analysis in the paper using total return volatility instead of residual return volatility. The conclusions remain unaffected. For our sample, the rank correlation between the two volatility measures is 0.994.

transaction costs, though the data for this measure are available to us for only a limited number of years.

Indirect transaction costs are the adverse price effects of the trade and the delay in processing the transaction. Theoretical arguments suggest that dollar-trading-volume is an important determinant of these indirect costs (e.g., Kyle, 1985; Admati and Pfleiderer, 1988; Foster and Viswanathan, 1990; Bhushan, 1992). If stocks are thinly traded, transactions are less likely to be completed quickly and are more likely to cause adverse price effects.[4]

Additionally, short selling contributes to the costs of arbitrage. Short sellers must borrow the shorted securities and must repay (return) the securities on demand. This requirement exposes the short seller to the risk of a "short squeeze," in which the borrowed and shorted securities must be repurchased (often at a loss) unless an alternative lender for the securities can be found. The risk of a short squeeze is likely to be lower for stocks with substantial institutional ownership, because it is easier to find alternative lenders of such stocks if the original lender should demand return of the borrowed shares (Dechow et al., 2001). We use percentage of institutional ownership as a proxy for the costs of short selling. A more direct measure of the costs of short selling would be borrowing prices in the equity lending market. However, the data are not easily available. Since short-selling cost applies only to low B/M stocks, we examine the relation between institutional owernship and the B/M effect separately for the low and high B/M stocks. If the B/M effect is due to mispricing and short-selling costs prevent it from being arbitraged away, then the relation would be stronger for the low B/M stocks than for the high B/M stocks.

Finally, we consider a comprehensive measure of transaction costs. Lesmond et al. (1999) argue that an investor will trade on information not reflected in the price of a security only if the profit net of all the transaction costs is expected to be positive. As a result, a security with high transaction costs will exhibit more frequent daily returns that are zero than a security with low transaction costs. Thus, the frequency of zero daily returns represents a comprehensive measure of transaction costs.[5]

[4] BARRA, an investment consulting company, argues that market price impact is a function of not only trading volume but also of response of order flows to price signals, return volatility, how often the stock trades, and the distribution of trade sizes. Unfortunately, we do not have the data to estimate BARRA's measure of price impact. We use dollar-trading volume as a measure of indirect transaction costs, noting that by itself it may not do as good a job of measuring the price impact of trade. Also note that since return volatility contributes to the price impact of trade, the explanatory power of return volatility for the B/M effect could be partly due to transaction costs. We attempt to isolate the explanatory power of return volatility due to arbitrage risk in a multiple regression context, discussed later.

[5] Lesmond et al. (1999) present a model for estimating the effective transaction cost of a stock, based on the time series of daily security returns and the frequency of zero returns. We use their model to estimate the effective transaction costs and repeat our analysis with this measure. The results lead to the same conclusions as based on the zero return frequency measure. We chose to report in the paper results based on the zero return frequency measure because we lose observations in estimating the Lesmond et al. model due to the lack of convergence of the estimation process.

2.3. Investor sophistication

Even when costs of arbitrage are high, stocks may not be mispriced if sophisticated investors are actively involved in trading for reasons other than arbitrage, such as liquidity or risk management. These investors are less likely than naïve investors to have systematically biased expectations about future firm earnings, a possible source of *B*/*M*-related mispricing. We use analyst following as a measure of investor sophistication. A large number of analysts following a stock would be associated with a greater number of market participants having access to sophisticated analysis about the stock (Brennan et al., 1993; Hong et al., 2000).

However, Dechow and Sloan (1997) show that analysts themselves overestimate future prospects of low *B*/*M* firms. Thus, we use number of institutional owners as an alternative measure of the level of investor sophistication (Chen et al., 2002; Bartov et al., 2000; Bhushan, 1994). Compared to the percentage of institutional ownership, the number of institutional owners is probably a better measure for investor sophistication. Percentage of ownership can be high when there are just one or two large but potentially unsophisticated investors.

2.4. Firm size

Lakonishok et al. (1994) use firm size to proxy for arbitrage costs and investor sophistication. Our analysis considers more direct measures of arbitrage costs and investor sophistication proposed in the literature. However, to facilitate comparability with prior studies, we also examine the effect of firm size on the ability of *B*/*M* to predict returns.

If we observe that the *B*/*M* effect is cross-sectionally correlated with arbitrage costs and investor sophistication measures, we will conclude that the *B*/*M* effect is likely due to market mispricing. For this conclusion, it does not matter that a measure presumed to represent one type of arbitrage costs or investor sophistication also captures the effect of another type of arbitrage cost and/or investor sophistication. And this is likely to happen if the measures are correlated. However, in this study we are also interested in examining whether arbitrage risk contributes incrementally to the existence of the *B*/*M* effect, and for that we rely on multivariate analysis.

3. Data and empirical results

3.1. Returns to *B*/*M* based portfolio strategies

Our sample consists of all firms on NYSE and AMEX from the period 1976 to 1997, having book value data available from the Standard and Poor's Compustat database and having price and return data for the subsequent three-year period available from the Center for Research in Security Prices (CRSP) database. *B*/*M* is calculated as book value in year *t*-1 divided by market value of equity at the end of

June of year t. Following Fama and French (1992), we drop firms with negative book value and observations with the highest and lowest 0.5% of values for B/M. The remaining data includes 64,486 firm-year observations.

For each year, we form quintile portfolios with the lowest B/M observations placed in portfolio Q1 and the highest B/M observations placed in portfolio Q5. Buy-and-hold returns are measured over one-, two-, and three-year holding periods beginning in July of year t. We compute difference in returns of the extreme B/M portfolios, Q5–Q1 returns, for each sample year and use the time-series variation over the sample period to compute the significance level. Our t-statistic is computed as mean divided by standard error of the annual estimates and our Z-statistic is computed in a similar manner except that rank measures of annual estimates are used (Wilcoxon rank-sum test). To correct for serial correlation in returns induced by overlapping holding periods for return horizons greater than one year, we use the Newey and West (1987) procedure.

Table 1 shows that B/M-based portfolio strategies produce positive returns for our data set. The Q5−Q1 returns are 8.9%, 21.6%, and 30.7%, over one-, two-, and three-year holding periods, respectively (*Ret12*, *Ret24*, and *Ret36*), all significant at

Table 1

Characteristics of the book-to-market (B/M) quintile portfolios over the 1976–1997 period

For each year from 1976 to 1997, stocks are assigned to five quintile portfolios based on the value of B/M, calculated as book value in year $t-1$ divided by market value of equity at the end of June of year t. For each B/M quintile, means of the following variables are calculated: *ME* is market value of equity in millions at the end of June of year t. *Beta* is systematic risk estimated using monthly returns over a maximum of 36 months beginning July of year t. *Ret12*, *Ret24*, and *Ret36* are the one-year, two-year, and three-year buy-and-hold return, respectively, beginning July of year t. *SRet12*, *SRet24*, and *SRet36* are the size-adjusted one-year, two-year, and three-year buy-and-hold return, respectively, beginning in July of year t, defined as raw buy-and-hold return less size-decile return, where size deciles are based on NYSE and AMEX firms. Firms with negative book value and observations with the highest and lowest 0.5% of the values of B/M are dropped. Statistical significance is reported for difference in the values of Q5 and Q1 portfolios, *Q5−Q1 Diff.* The t-statistic is computed as mean divided by standard error of the annual estimates. To correct for serial correlation in returns induced by overlapping holding periods for return horizons greater than one year, we use the Newey and West (1987) procedure. *** and ** indicate significance at better than the 1% and 5% levels (two-tailed), respectively.

Variable	*Q1* (Low B/M)	*Q2*	*Q3*	*Q4*	*Q5* (High B/M)	*All* firms	*Q5–Q1 Diff.*
B/M	0.198	0.431	0.657	0.942	1.877	0.821	1.679***
ME	1,083	1,001	726	479	199	698	−884***
Beta	1.213	1.071	0.925	0.847	0.814	0.974	−0.399***
Ret12	0.130	0.169	0.177	0.196	0.219	0.178	0.089***
Ret24	0.224	0.330	0.349	0.389	0.440	0.346	0.216***
Ret36	0.374	0.519	0.542	0.618	0.681	0.547	0.307***
SRet12	−0.043	−0.007	0.001	0.014	0.037	0.000	0.081**
SRet24	−0.111	−0.012	0.006	0.037	0.095	0.003	0.206***
SRet36	−0.135	−0.006	0.018	0.083	0.161	0.024	0.297***
Number of obs.						64,486	

less than the 1% level. These results are similar in magnitude to results in prior studies (e.g., Lakonishok et al., 1994) and suggest that the *B/M* ratio has the ability to predict returns. Table 1 also indicates that market capitalization (*ME*) is significantly greater for the low *B/M* quintile than for the high *B/M* quintile ($1,083 million versus $199 million). Higher returns to the high *B/M* quintile could be due to the higher risk premium associated with smaller firms. To control for the effect of size differences, we calculate size-adjusted returns, equal to raw returns minus the corresponding NYSE/AMEX CRSP size-decile index returns (as in Lakonishok et al., 1994). Q5−Q1 size-adjusted returns are 8.1%, 20.6%, and 29.7%, over one-, two-, and three-year holding periods, respectively (*SRet*12, *SRet24*, and *SRet36*), all significant at less than the 5% level. Thus, the raw return results in Table 1 do not seem driven by size differences between the high-*B/M* and low-*B/M* portfolios. In addition, the systematic risk measure *Beta* decreases monotonically from *B/M* quintile Q1–Q5, indicating that the higher returns of the high *B/M* portfolio do not appear to represent compensation for risk related to *Beta*. These results are also consistent with the results in previous studies, including Lakonishok et al. (1994).

3.2. Definitions, data sources, and descriptive statistics of the measures of arbitrage risk, transaction costs and investor sophistication

Our measure of expected arbitrage risk is *Ivolatility,* which we obtain by regressing daily returns on a value-weighted market index over a maximum of 250 days ending on June 30 of year t and then computing the variance of the residuals.[6] Note that the *B/M* portfolio returns are measured starting from July of year t. *Ivolatility* is estimated using CRSP data. The measures of direct transaction costs are (1) *Bid-Ask*, percentage bid-ask spread, defined as (bid-ask)/(0.5*(bid + ask)) averaged over the last hour of the last trading day of each of the 12 months, beginning from July of year $t-1$ and ending in June of year t, and (2) *Price*, the closing price per share in June of year t. Data for *Bid-Ask* are obtained from the TAQ (NYSE Trade-and-Quote) database. We have TAQ data available to us for the years 1994 to 1997. Data for *Price* are obtained from the CRSP database. Our measure of indirect transaction costs is *Volume*, the annual volume of trade in the firm's shares from July 1 of year $t-1$ to June 30 of year t, in millions of dollars. Data for *Volume* are obtained from the CRSP database. Our proxy for short-selling costs is *Inst%*, the percentage of common stock owned by institutions at the end of year $t-1$, obtained from the Compact Disclosure database, with data availability beginning in 1987. Our comprehensive measure of transaction costs is *Zerofreq*, the frequency of zero daily returns over the period July of year $t-1$ to June 30 of year t. Our measures of investor sophistication are (1) *Analysts*, the number of analysts' estimates included in the I/B/E/S database in May of year t, with data available beginning in 1976, and if a firm is not included in the database, *Analysts* is coded as zero (Bhushan, 1994), and

[6] Our results are not sensitive to the period used for the measurement of *Ivolatility*. We consider a shorter period, 125 daily returns, as well as a longer period, 36 monthly returns, both ending in June of year t, and obtain results that are consistent with those reported in the paper.

Table 2
Correlations among measures of arbitrage risk, transaction costs, investor sophistication and firm-size
Ivolatility is obtained by regressing daily returns on a value-weighted market index over a maximum of 250 days ending on June 30 of year t, where t is the B/M measurement year, and then computing the variance of the residuals. *Bid-Ask*, percentage bid-ask spread, defined as (bid-ask)/(0.5*(bid + ask)) averaged over the last hour of the last trading day of each of the 12 months, beginning from July of year $t-1$ and ending in June of year t. *Price* is the closing price of a share of common stock at end of June of year t. *Volume* is the annual volume of trade in the firm's shares ending in June of year t, in millions of dollars. *Inst%* is the percentage of common stock owned by institutions at end of year $t-1$. *Zerofreq* is the frequency of zero daily returns over one year ending in June of year t; *Analysts* is the number of analysts' estimates included in the IBES database in May of year t and is coded as 0 if the firm is not listed in the IBES database. *Inst#* is the number of institutional owners at end of year $t-1$. *ME* is market value of equity in millions of dollars at the end of June of year t. The number of observations is 64,486 (1976 to 1997), except for *Bid-Ask* (8,690, 1994–1997), *Inst%* (28,949, 1987–1997) and *Inst#* (29,001, 1987–1997). Spearman correlation coefficients are calculated each year from 1976 to 1997 (except for correlations involving *Inst%* and *Inst#* which are calculated each year from 1987 to 1997, and for *Bid-Ask* which are calculated each year from 1994 to 1997), and mean of the annual correlation coefficients are reported. Significance levels are computed using the mean and the standard error of the annual coefficients. The magnitudes of the mean correlation coefficients are all significantly different from both zero and one at the 1% level.

Variable	*Ivolatility*	*Bid-Ask*	*Price*	*Volume*	*Inst%*	*Zerofreq*	*Analysts*	*Inst#*
Bid-Ask	0.697							
Price	−0.708	−0.808						
Volume	−0.422	−0.772	0.725					
Inst%	−0.532	−0.567	0.675	0.696				
Zerofreq	0.311	0.650	−0.709	−0.811	−0.625			
Analysts	−0.328	−0.508	0.496	0.590	0.624	−0.522		
Inst#	−0.649	−0.803	0.780	0.878	0.840	−0.749	0.726	
ME	−0.607	−0.832	0.836	0.903	0.700	−0.758	0.576	0.916

(2) *Inst#*, the number of institutional owners at the end of year $t-1$, obtained from the Compact Disclosure database, with data availability beginning in 1987. *ME* is market value of equity in millions of dollars at the end of June of year t, obtained from the CRSP database.

For each year, all the observations are ranked in descending order of *Ivolatility*. The first 10% of the observations with the highest value of *Ivolatility* are assigned to the first group, G1, the next 10% to the second group, G2, and so on. The procedure is repeated for all the other variables with the following exceptions. First, for the variables *Price*, *Volume*, *Inst%*, *Analysts*, *Inst#*, and *ME*, group G1 represents firms with the lowest values and G10 firms with the highest values. Second, for *Analysts*, group G1 includes all firms with zero analyst following, and the remaining firms are placed in nine equal groups, in ascending order of analyst following. All of the variables exhibit substantial variation between the lowest (G1) and highest decile groups (G10). For group G1 (G10), the mean value of *Ivolatility* ($\times 10^{-2}$) is 0.661 (0.011), *Bid-Ask* is 0.095 (0.006), *Price* is \$1.250 (\$54.040), *Volume* is \$0.981 (\$3,941.632) million, *Inst%* is 0.313% (73.110%), *Zerofreq* is 0.728 (0.082), *Analysts* is 0.000 (25.236), *Inst#* is 0.939 (297.942), and *ME* is \$4.150 (\$5,588.633) million.

Table 3

Three-year size-adjusted returns (*SRet36*) to Q5–Q1 portfolios, based on magnitude of *B/M*, across deciles of the measures of arbitrage risk, transaction costs, investor sophistication and firm-size

SRet36 is the size-adjusted three-year buy-and-hold return, beginning in July of year t, defined as raw buy-and-hold return less size-decile return, where size deciles are based on NYSE and AMEX firms. *B/M* is book value in year $t-1$ divided by market value of equity at the end of June of year t. *Ivolatility* is obtained by regressing daily returns on a value-weighted market index over a maximum of 250 days ending on June 30 of year t, and then computing the variance of the residuals. *Bid-Ask*, percentage bid-ask spread, defined as (bid-ask)/(0.5*(bid + ask)) averaged over the last hour of the last trading day of each of the 12 months, beginning from July of year $t-1$ and ending in June of year t. *Price* is the closing price of a share of common stock at end of June of year t. *Volume* is the annual volume of trade in the firm's shares ending in June of year t, in millions of dollars. *Inst%* is the percentage of common stock owned by institutions at end of year $t-1$. *Zerofreq* is the frequency of zero daily returns over one year ending in June of year t. *Analysts* is the number of analysts' estimates included in the IBES database in May of year t and is coded as 0 if the firm is not listed in the IBES database. *Inst#* is the number of institutional owners at end of year $t-1$. *ME* is market value of equity in millions of dollars at the end of June of year t. For each year from 1976 to 1997, firms are ranked on the magnitude of *B/M*, and each firm is then assigned to a quintile group (Q1–Q5). Independent of the *B/M* rankings, firms are also ranked each year on each of the variables *Ivolatility*, *Bid-Ask*, *Price*, *Volume*, *Inst%*, *Zerofreq*, *Analysts*, *Inst#*, and *ME* and are then assigned to decile groups (G1–G10). For *Ivolatility*, *Bid-Ask*, and *Zerofreq*, firms are ranked in descending order and for *Price*, *Volume*, *Inst%*, *Inst#*, *Analysts*, and *ME*, firms are ranked in ascending order. For *Analysts*, group G1 includes all firms with zero analyst following, and the remaining firms are placed in nine equal groups, in ascending order of analyst following. Q5–Q1 *SRet36* is then computed for each of the groups (G1–G10), where Q5–Q1 *SRet36* is defined as the difference between three-year size-adjusted returns following portfolio formation for the highest- and lowest-*B/M* quintile portfolios (Q5 and Q1). G1–G10 represents the Q5–Q1 *SRet36* of group G1 less the Q5–Q1 *SRet36* of group G10. Correlations are across-group correlations between group ranks 1 to 10 and the Q5–Q1 *SRet36* of the groups. The t-statistic is computed as mean divided by standard error of the annual estimates and Z-statistic is computed in a similar manner except that rank measures of annual estimates are used (Wilcoxon rank-sum test). To correct for serial correlation in returns induced by overlapping holding periods, we use the Newey and West (1987) procedure. The first and second numbers in parentheses below a Q5–Q1 *SRet36* value represent the number of firm-year observations in Q5 and Q1 portfolios, respectively. Statistical significance is indicated by *** for 1% level, ** for 5% level, and * for 10% level (two-tailed).

Variable	*Ivolatility*	*Bid-Ask*	*Price*	*Volume*	*Inst%*	*Zerofreq*	*Analysts*	*Inst#*	*ME*
No. of Obs.	64,486	8,690	64,486	64,486	28,949	64,486	64,486	29,001	64,486
Sample Period	1976–1997	1994–1997	1976–1997	1976–1997	1987–1997	1976–1997	1976–1997	1987–1997	1976–1997
G1	0.513	0.502	0.571	0.266	0.284	0.367	0.395	0.365	0.388
	(1978,1589)	(304,115)	(2488,1248)	(3066,598)	(729,748)	(2612,850)	(10289,7967)	(827,722)	(3114,533)
G2	0.438	0.038	0.534	0.388	0.293	0.365	0.243	0.395	0.426
	(1569,1663)	(229,147)	(2242,1245)	(2264,849)	(704,753)	(1896,1220)	(526,480)	(851,577)	(2087,1094)
G3	0.250	0.019	0.302	0.293	0.380	0.291	0.108	0.375	0.395
	(1410,1599)	(157,165)	(2023,1043)	(1664,1076)	(761,570)	(1672,1253)	(428,579)	(868,510)	(1839,1254)
G4	0.306	0.021	0.246	0.399	0.064	0.301	0.153	0.033	0.451
	(1212,1615)	(149,197)	(1570,1054)	(1398,1192)	(693,484)	(1737,1070)	(251,391)	(696,458)	(1470,1301)

G5	0.172	−0.021	0.362	0.341	0.320	0.386	0.209	0.326	0.293
	(1174,1511)	(162,178)	(1273,1065)	(1174,1305)	(619,483)	(1472,1051)	(248,548)	(655,486)	(1184,1325)
G6	0.203	−0.321	0.210	0.312	0.274	0.350	0.223	0.212	0.149
	(1257,1353)	(136,186)	(1078,1215)	(926,1329)	(556,501)	(1106,1121)	(328,660)	(539,508)	(973,1407)
G7	0.085	−0.337	0.207	0.152	0.254	0.209	0.150	0.023	0.120
	(1192,1106)	(107,172)	(792,1186)	(695,1536)	(517,475)	(853,1198)	(229,554)	(399,599)	(754,1438)
G8	0.106	0.198	0.110	0.119	−0.057	0.188	−0.021	−0.196	0.070
	(1125,955)	(90,192)	(615,1388)	(588,1658)	(362,524)	(711,1352)	(219,551)	(258,671)	(523,1521)
G9	0.146	−0.140	0.130	0.109	0.093	0.154	0.221	0.117	0.188
	(1024,839)	(70,193)	(476,1542)	(523,1622)	(261,619)	(500,1556)	(227,533)	(216,614)	(514,1517)
G10	0.017	0.132	0.174	0.312	−0.119	0.310	0.273	0.054	0.344
	(949,659)	(62,252)	(333,1903)	(531,1702)	(186,697)	(331,2218)	(145,626)	(92,726)	(432,1499)
G1–G10	0.496	0.369	0.396	−0.046	0.403	0.056	0.122	0.311	0.044
t-statistic	5.142***	2.762***	3.379***	−0.310	4.304***	0.667	0.961	1.337	0.354
Z-statistic	4.002***	1.876*	2.734***	−0.340	2.232**	0.434	1.349	1.641	0.176
Correlation	−0.918	−0.376	−0.875	−0.545	−0.715	−0.631	−0.318	−0.722	−0.635
p-value	0.000***	0.283	0.000***	0.102	0.019**	0.050**	0.369	0.018**	0.048**

Our tests examine how the returns to a *B/M*-based Q5–Q1 portfolio vary with each of these partitioning variables. The substantial differences in the values of each of the variables across the decile groups indicate that the tests should offer reasonable power.

Table 2 reports Spearman correlations among our measures of arbitrage risk, transaction costs, investor sophistication, and firm size. The correlation of *ME* with *Ivolatility* is -0.607, and the correlation of *ME* with the other measures ranges from 0.576 to 0.916. While the magnitude of each of these correlations is significantly different from zero, each correlation magnitude is also significantly less than one at the 1% level. This suggests that in order to understand the effects of arbitrage risk, transaction costs, and investor sophistication on the ability of *B/M* to predict returns, it is not sufficient to examine only the effect of firm size. The correlations among our measures of arbitrage risk, transaction costs, and investor sophistication range in magnitude from 0.311 to 0.708. All these magnitudes are significantly less than one, but they are also not close to zero. Thus, when using univariate tests, caution is warranted in drawing conclusions about which specific arbitrage cost or investor sophistication factor is contributing to the existence of the *B/M* effect.

3.3. Portfolio tests

As before, firms are ranked each year on the magnitude of *B/M*, and each firm is then assigned to a quintile group (Q1–Q5) based on this ranking. Independent of the *B/M* rankings, firms are ranked each year in descending order of *Ivolatility* and are then assigned to a decile group (G1 to G10) based on this ranking. The result of this procedure is 50 portfolios in each year, with each portfolio composed of firms with similar *B/M* and *Ivolatility*. Size-adjusted returns of the Q5–Q1 portfolio for the three years following portfolio formation (Q5–Q1 *SRet36*) are then computed for each of the groups (G1–G10).

The first column of Table 3 reports the average of the annual estimates of *SRet36* for the Q5–Q1 portfolios for each of the ten *Ivolatility*-based groups G1 to G10. The mean Q5–Q1 *SRet36* of the high-*Ivolatility* group G1 is significantly greater than that of the low-*Ivolatility* group G10. The difference is 0.496 (t-statistic$=5.142$ and Z-statistic$=4.002$). This difference is also economically significant. The 10% of our sample stocks with the greatest past return volatility exhibit Q5–Q1 *SRet36* of 51.3%. On the other hand, the 10% of our sample stocks with the lowest return volatility exhibit Q5–Q1 *SRet36* of 1.7%. Recall that for the full sample Q5–Q1 *SRet36* is 29.7% (see Table 1). The above tests consider only the extreme decile groups of *Ivolatility*. To examine the relation across all the decile groups, we compute a correlation between group ranks and Q5–Q1 *SRet36* of the groups. This correlation is negative and highly significant (-0.918, p-value$=0.000$) and is consistent with the results based on the extreme decile groups. The table also reports for each group, G1 to G10, the number of observations that fall in the *B/M* quintile 5 and *B/M* quintile 1. For example, for G1, these numbers are 1,978 and 1,589, respectively. Note that for none of the portfolios, there seems to be a concern that a small number of observations are driving the results. To check the sensitivity of our results to significant bid-ask bounce in low priced stocks, we repeat our analysis after dropping stocks with price less than \$1 (Lee and Swaminathan, 2000). The sample size decreases from 64,486 to 61,623, but the results remain very similar. For example, for the *Ivolatility* measure, the G1–G10 value is 0.508 ($t = 5.256$) and the correlation is -0.916 ($p = 0.000$).

Since Table 3 provides only summary measures for the sample years, we plot in Fig. 1 annual estimates of Q5–Q1 *SRet36* for *Ivolatility* groups G1 and G10. The plot does not indicate any pattern that contradicts our conclusions from Table 3 results. The Q5–Q1 portfolio returns are greater for the high volatility group (G1) than for the low volatility group (G10) in 20 of the 22 sample years. Also, since Table 3 reports results for aggregate returns for the 36 months after the portfolio formation date, we plot cumulative return at monthly intervals (Fig. 2). For the high volatility group, the cumulative returns for the Q5–Q1 portfolio increases almost monotonically throughout the 36-month period. In contrast, for the low volatility group, the portfolio return is close to zero throughout the 36-month period. These results support the market-mispricing explanation and indicate that arbitrage risk is one of the reasons why arbitrageurs do not trade away the systematic *B/M*-related mispricing. Note, however, that based on these univariate results, we cannot say to what extent arbitrage risk alone is responsible for the results.

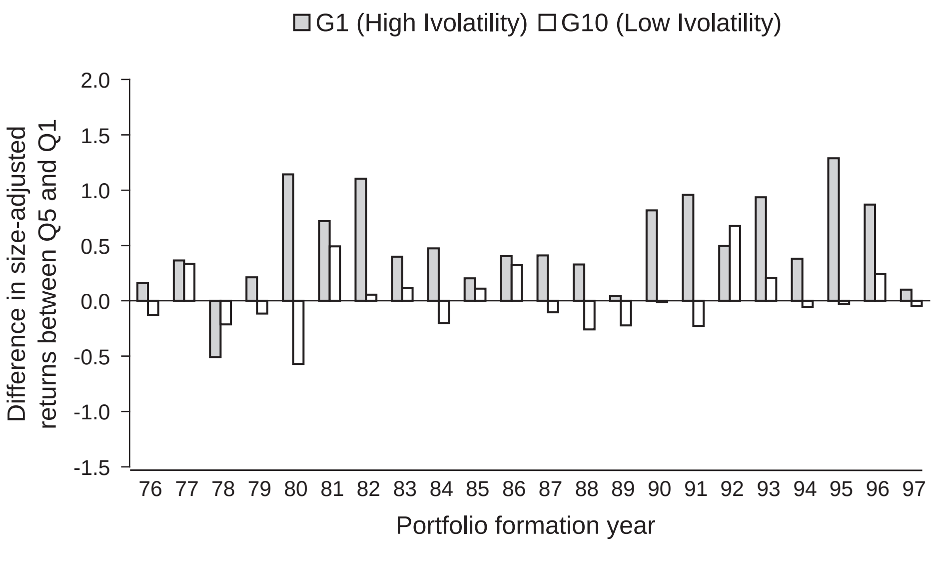

Fig. 1. Year-by-year size-adjusted three-year returns (*SRet36*) to the Q5–Q1 portfolios based on *B/M*, for high versus low arbitrage risk. *B/M* is calculated as book value in year $t-1$ divided by market value of equity at the end of June of year t. Q5–Q1 *SRet36* is defined as the difference between the three-year size-adjusted returns following portfolio formation for the highest- and lowest-*B/M* quintile portfolios (Q5 and Q1). *Ivolatility* is obtained by regressing daily returns on a value-weighted market index over a maximum of 250 days ending on June 30 of year t, and then computing the variance of the residuals. Group G1 (G10) represents 10% of the firms with highest (lowest) *Ivolatility* values in a given year.

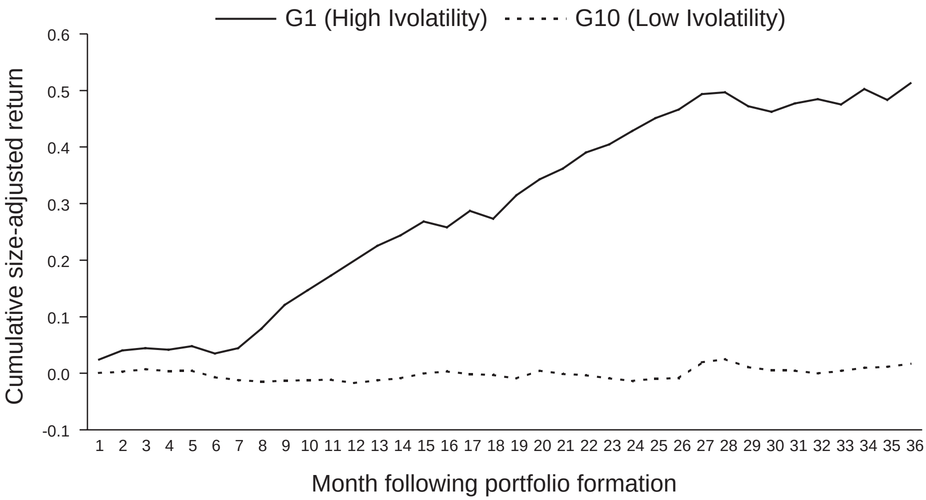

Fig. 2. Cumulative monthly size-adjusted returns to the Q5–Q1 portfolios based on *B/M*, for high versus low arbitrage risk. *B/M* is calculated as book value in year $t-1$ divided by market value of equity at the end of June of year t. Q5–Q1 size-adjusted return is defined as the difference between the size-adjusted returns following portfolio formation for the highest- and lowest-*B/M* quintile portfolios (Q5 and Q1). *Ivolatility* is obtained by regressing daily returns on a value-weighted market index over a maximum of 250 days ending on June 30 of year t, and then computing the variance of the residuals. Group G1 (G10) represents 10% of the firms with highest (lowest) *Ivolatility* values in a given year.

The above analysis is repeated for each of our other measures of arbitrage costs and investor sophistication. For all the measures, the correlation between Q5–Q1 *SRet36* and the group rank for G1 to G10 is negative, consistent with the mispricing explanation. Moreover, these correlations are significant for most of the variables: *Bid-Ask* is -0.376 ($p = 0.283$), *Price* is -0.875 ($p = 0.000$), *Volume* is -0.545 ($p = 0.102$), *Inst%* is -0.715 ($p = 0.019$), *Zerofreq* is -0.631 ($p = 0.050$), *Analysts* is -0.318 ($p = 0.369$), *Inst#* is -0.722 ($p = 0.018$), and *ME* is -0.635 ($p = 0.048$). The correlation for *Bid-Ask* is not significant, but the difference in Q5–Q1 *SRet36* between groups G1 and G10 is significant, 0.369 ($t = 2.762$, $Z = 1.876$). Also note that for *Bid-Ask* variable, data are available to us for just four years. The lack of significance of the correlation for *Analysts* could be because group G1 (zero analysts following) has most of the observations, leaving relatively few observations in the other groups. We compare the Q5–Q1 *SRet36* for stocks with zero versus non-zero analysts following and find the difference, 0.205, to be significant ($t = 2.267$ and $Z = 2.124$).[7] Thus, either the correlation result or the G1–G10 result is significant for

[7] Earlier, we argue that if *Inst%* variable proxies for short selling costs, then the difference in *SRet36* between high versus low *Inst%* group would be greater for the low *B/M* portfolio than the high *B/M* portfolio. We find that the difference in *SRet36* across the highest and lowest decile of *Inst%* for Q1 (low *B/M*) portfolio is -0.278, $t = -4.258$, and for Q5 (high *B/M*) portfolio is 0.125, $t = 2.028$. The result is stronger for the low *B/M* portfolio, consistent with *Inst%* proxying for short-selling costs and that short-selling costs contribute to the existence of the *B/M* effect.

each of the variables, suggesting that the *B/M* effect increases with arbitrage costs and decreases with investor sophistication, consistent with the mispricing explanation.

3.4. Incremental role of arbitrage risk

In the previous section, we use a univariate analysis and find that the *B/M* effect is cross-sectionally correlated with the arbitrage costs and investor sophistication measures, and conclude that the *B/M* effect is likely to be due to market mispricing. For this conclusion, it does not matter that a measure presumed to represent one type of arbitrage cost or investor sophistication also captures the effect of other types of arbitrage cost and/or investor sophistication. However, we are also interested in determining the incremental role of arbitrage risk in the existence of the *B/M* effect and therefore use multiple regression tests. We estimate the following regression model:

$$\begin{aligned} SRet36 \;=\; & b_0 + b_1 Beta + b_2 B/M + b_3 B/M \times Ivolatility^{-1} + b_4 B/M \times Price \\ & + b_5 B/M \times Ln(Volume) + b_6 B/M \times Zerofreq^{-1} + b_7 B/M \times Analysts \\ & + b_8 B/M \times Ln(ME) + b_9 Ivolatility^{-1} + b_{10} Price + b_{11} Ln(Volume) \\ & + b_{12} Zerofreq^{-1} + b_{13} Analysts + b_{14} Ln(ME) + \mathrm{e}. \end{aligned} \tag{1}$$

The coefficients on the interaction terms in Eq. (1) capture how the *B/M* effect varies cross-sectionally with the arbitrage cost/investor sophistication measures. To maintain consistency of the expected signs on the coefficients on all the interaction terms, we use inverse of *Ivolatility* and *Zerofreq*. We use the log transformation of *Volume* and *ME* because of the distributional properties of these variables (see e.g., Brennan et al., 1993). All the arbitrage cost and investor sophistication measures are also included by themselves to capture their main effects on future abnormal returns. Otherwise, the coefficients on the interaction terms could be biased. The model also controls for the relation between *Beta* and returns. A significant negative coefficient on $B/M \times Ivolatility^{-1}$ would suggest that after controlling for the other arbitrage cost and investor sophistication measures, *Ivolatility* explains cross-sectional variation in the *B/M* effect.

Following the Fama and MacBeth (1973) procedure, we estimate Eq. (1) by sample years, and report the means of the annual estimates. We also report the *t*-statistics of the mean of annual slope estimates, using the Newey and West (1987) corrected standard error of the means. Given that data for *Bid-Ask* are available for only four years and for *Inst%* and *Inst#* are available for only one-half of the sample years, we do not include these variables in the above model. Note, however, that Eq. (1) does include alternative measures of transaction costs and investor sophistication.

Table 4 reports the regression estimates of Eq. (1) and two reduced versions of the equation. The first column presents the results of a model with *Beta* and *B/M* as the explanatory variables. The mean value of the slope coefficient on *B/M* is positive and significant 0.115 ($t = 3.585$), suggesting the presence of the *B/M* effect, consistent

Table 4

Regression tests of incremental role of arbitrage risk beyond transaction costs in explaining the *B/M* effect
The dependent variable is *SRet36,* the size-adjusted three-year buy-and-hold return, beginning in July of year *t*, defined as raw buy-and-hold return less size-decile return, where size deciles are based on NYSE and AMEX firms. *Beta* is systematic risk estimated using monthly returns over a maximum of 36 months beginning July of year *t*. *B/M* is book value in year $t-1$ divided by market value of equity at the end of June of year *t*. *Ivolatility* is obtained by regressing daily returns on a value-weighted market index over a maximum of 250 days ending on June 30 of year *t*, and then computing the variance of the residuals. *Price* is the closing price of a share of common stock at end of June of year *t*. *Volume* is the annual volume of trade in the firm's shares ending in June of year *t*, in millions of dollars. *Zerofreq* is the frequency of zero daily returns over one year ending in June of year *t*. *Analysts* is the number of analysts' estimates included in the IBES database in May of year *t* and is coded as 0 if the firm is not listed in the IBES database. *ME* is market value of equity in millions of dollars at the end of June of year *t*. Regression models are estimated for each year from 1976 to 1997 using 63,855 available observations, and means of the annual estimates are reported. To avoid giving extreme observations heavy weight in the regressions, we eliminate observations with the smallest and largest 0.5% of *SRet36*. We also set the smallest and largest 0.5% of the observations on each of the independent variables to its 0.005 and 0.995 fractile values. The *t*-statistic is computed as mean divided by standard error of the annual estimates. To correct for serial correlation in returns induced by overlapping holding periods, we use the Newey and West (1987) procedure. Statistical significance is indicated by *** for 1% level, ** for 5% level, and * for 10% level (two-tailed).

	(1)	(2)	(3)
Intercept	−0.251***	−0.348***	−0.355***
	(−10.010)	(−7.021)	(−4.865)
Beta	0.126***	0.148***	0.174***
	(2.995)	(3.797)	(4.879)
B/M	0.115***	0.136***	0.078***
	(3.585)	(3.874)	(2.799)
$B/M \times Ivolatility^{-1}$ ($\times 10^{-2}$)		−0.138***	−0.147***
		(−2.797)	(−2.589)
$B/M \times Price$ ($\times 10^{-2}$)			0.119
			(0.787)
$B/M \times Ln(Volume)$ ($\times 10^{-2}$)			−0.054
			(−0.049)
$B/M \times Zerofreq^{-1}$ ($\times 10^{-2}$)			−0.413
			(−0.581)
$B/M \times Analysts$ ($\times 10^{-2}$)			−0.357
			(−1.305)
$B/M \times Ln(ME)$ ($\times 10^{-2}$)			2.152*
			(1.746)
$Ivolatility^{-1}$ ($\times 10^{-2}$)		0.323***	0.183
		(2.699)	(1.680)
Price ($\times 10^{-2}$)			0.450***
			(3.160)
Ln(Volume) ($\times 10^{-2}$)			−7.758***
			(−4.855)
$Zerofreq^{-1}$ ($\times 10^{-2}$)			−0.287
			(−0.842)
Analysts ($\times 10^{-2}$)			0.655***
			(2.498)
Ln(ME) ($\times 10^{-2}$)			4.429***
			(2.978)
Average Adj. R^2	0.024	0.034	0.049

with the portfolio results in Table 1. The second column presents the results of a model with *Ivolatility* variables. The variable $B/M \times Ivolatility^{-1}$ has a significantly negative coefficient (-0.138, $t = -2.797$), suggesting that the *B/M* effect increases with the increase in return volatility, consistent with the portfolio results in Table 3. The third column presents the results of complete Eq. (1). The coefficient on $B/M \times Ivolatility^{-1}$ remains negative and significant (-0.147, $t = -2.589$), suggesting that return volatility incrementally explains, beyond the other arbitrage cost/investor sophistication measures, the cross-sectional variation in the *B/M* effect.

Among the other interaction terms in Eq. (1), none has a negative and significant coefficient. Why do the results for some of the other arbitrage costs/investor sophistication measures lack significance, given that in Table 3, the portfolio results were significant for many of these variables? A likely explanation is that these measures are highly correlated with each other and that more than one measure can capture similar underlying effects. For example, we include in the model three measures for transaction costs, *Price*, *Volume*, and *Zerofreq*. Our reason for including all three transaction-cost measures is to try and control for the effect of transaction costs, as best as we can, so that the result for return volatility has a better chance of representing the effect of arbitrage risk.

We also estimate Eq. (1) with rank transformed explanatory variables (see Table 3 for ranking procedure). Consistent with the Table 4 results, the coefficients on $RK(B/M) \times RK(Ivolatility)$, where $RK(\ )$ indicates rank transformation, are negative and significant for both the reduced and full models, -0.911 ($t = -4.755$) and -0.655 ($t = -2.503$), respectively. These results suggest that the finding of significant incremental explanatory power of return volatility is not sensitive to the linearity assumption in Eq. (1).

4. Conclusion

Using the various measures of arbitrage costs and investor sophistication identified in the literature, we find that the ability of book-to-market (*B/M*) ratio to predict future returns is greater for stocks with higher transaction costs and with less ownership by sophisticated investors. These results are consistent with the view that the *B/M* effect is due to market mispricing. We also find that of all the arbitrage costs and investor sophistication measures considered, the ability of *B/M* to predict returns is most strongly and most consistently related to return volatility. Specifically, the 10% of stocks in our sample with the greatest return volatility exhibit three-year size-adjusted return of 51.3% for a portfolio long in the highest quintile of *B/M* and short in the lowest quintile of *B/M*; the corresponding return for the 10% of stocks with least return volatility is negligible, 1.7%. Also, the *B/M* effect for the high volatility group exceeds that for the low volatility group in 20 out of 22 sample years. Furthermore, our results show that return volatility has incremental power beyond all the other measures of transaction costs and investor sophistication in explaining the cross-sectional variation in the *B/M* effect. This finding is consistent with the Shleifer and Vishny (1997) thesis that risk associated with the volatility of arbitrage returns deters arbitrage activity and is an important reason for the existence of the *B/M*-related mispricing. Given that arbitrage risk is a holding cost, which increases with the length of the period the position is held, this factor is likely to be important in explaining other long-term trading strategies identified in the literature.

Even though our measures of arbitrage risk, transaction costs, and investor sophistication are motivated by prior literature, we recognize the limitation of our tests in discriminating between the mispricing and risk explanations for the *B/M* effect. In the absence of a full understanding of asset pricing, it is not possible to test for mispricing without jointly testing some model of expected returns (Fama, 1970). For example, our result that the *B/M* effect is greater for high volatility stocks could be consistent with the following risk-based explanation. Notwithstanding the contradictory findings in the literature (discussed earlier), if *B/M* captures financial distress, then the *B/M* effect may be stronger for firms with high volatility stocks because the financial distress factor may be more sensitive for such firms. We leave this issue for future research. However, the tests in our study provide new evidence that should help readers revise their beliefs about the relative likelihood of the mispricing and risk explanations for the book-to-market anomaly.

References

Admati, A.R., Pfleiderer, P., 1988. A theory of intraday patterns: volume and price variability. Review of Financial Studies 1, 3–40.

Barber, B., Lehavy, R., McNichols, M., Trueman, B., 2001. Can investors profit from the prophets? security analyst recommendations and stock returns. The Journal of Finance 56, 531–563.

Bartov, E., Radhakrishnan, S., Krinsky, I., 2000. Investor sophistication and patterns in stock returns after earnings announcements. The Accounting Review 75, 43–64.

Bhardwaj, R., Brooks, L., 1992. The january anomaly: effects of low share price, transaction costs and bid-asked bias. Journal of Finance 47, 552–576.

Bhushan, R., 1992. Trading costs, liquidity and asset holdings. Review of Financial Studies 4, 343–360.

Bhushan, R., 1994. An informational efficiency perspective on the post-earnings announcement drift. Journal of Accounting and Economics 18, 45–66.

Blume, M.E., Goldstein, M.A.,1992. Displayed and effective spreads by market. Unpublished working paper. Wharton School, University of Pennsylvania.

Brennan, M., Jagadeesh, N., Swaminathan, B., 1993. Investment analysis and the adjustment of stock prices to common information. Review of Financial Studies 6, 799–824.

Chan, L., Hamao, Y., Lakonishok, J., 1991. Fundamentals and stock returns in Japan. Journal of Finance 46, 1739–1789.

Chan, L., Jagadeesh, N., Lakonishok, J., 1995. Evaluating the performance of value versus glamour stocks: the impact of selection bias. Journal of Financial Economics 38, 269–296.

Chen, J., Hong, H., Stein, J., 2002. Breadth of ownership and stock returns. Journal of Financial Economics 66, 171–205.

Copeland, T., Mayers, D., 1982. The value line enigma (1965-1978): a case study of performance evaluation issues. Journal of Financial Economics 10, 289–322.

Dechow, P., Hutton, A., Meulbroek, L., Sloan, R., 2001. Short-sellers, fundamental analysis and stock returns. Journal of Financial Economics 61, 77–106.

Dechow, P., Sloan, R., 1997. Returns to contrarian investment strategies: tests of naïve expectation hypothesis. Journal of Financial Economics 43, 3–27.

Dichev, I.D., 1998. Is the risk of bankruptcy a systematic risk? Journal of Finance 53, 1131–1147.
Fama, E., 1970. Efficient capital markets: a review of theory and empirical work. Journal of Finance 25, 383–417.
Fama, E., French, K., 1992. The cross-section of expected stock returns. Journal of Finance 46, 427–466.
Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.
Fama, E., French, K., 1997. Industry costs of equity. Journal of Financial Economics 43, 153–193.
Fama, E., MacBeth, J., 1973. Risk, return and equilibrium: empirical tests. Journal of Politcal Economy 81, 607–636.
Foster, F.D., Viswanathan, S., 1990. A theory of the interday variations in volume, variance and trading costs in securities markets. Review of Financial Studies 3, 593–624.
Hong, H., Lim, T., Stein, J., 2000. Bad news travels slowly: size, analyst coverage, and the profitability of momentum strategies. Journal of Finance 55, 265–295.
Ippolito, R., 1992. Consumer reaction to measures of poor quality: evidence from the mutual fund industry. Journal of Law and Economics 35, 45–70.
Knez, P., Ready, M., 1996. Estimating the profits from trading strategies. Review of Financial Studies 9, 1121–1163.
Kothari, S.P., Shanken, J., Sloan, R., 1995. Another look at the cross-section of expected stock returns. The Journal of Finance 50, 185–224.
Kyle, A., 1985. Continuous auctions and insider trading. Econometrica 53, 1315–1335.
Lakonishok, J., Shleifer, A., Vishny, R., 1994. Contrarian investment, extrapolation and risk. Journal of Finance 49, 1541–1578.
La Porta, R., Lakonishok, J., Shleifer, A., Vishny, R., 1997. Good news for value stocks. Journal of Finance 52, 859–873.
Lee, C., Swaminathan, B., 2000. Price momentum and trading volume. Journal of Finance 55, 2017–2069.
Lesmond, D., Ogden, J., Trzcinka, C., 1999. A new estimate of transaction costs. Review of Financial Studies 12, 1113–1141.
Lesmond, D., Schill, M., Zhou, C., 2001. The illusory nature of momentum profits. Unpublished working paper. University of Virginia.
Lo, A.W., MacKinlay, A.C., 1988. Stock market prices do not follow random walks: evidence from a simple specification test. Review of Financial Studies 1, 41–66.
Newey, W., West, K., 1987. Hypothesis testing with efficient method of moments estimation. International Economic Review 28, 777–787.
Pontiff, J., 1996. Costly arbitrage: evidence from closed end funds. Quarterly Journal of Economics 111, 1135–1152.
Reinganum, M., 1983. The anomalous stock market behavior of small firms in january: some empirical tests for tax-loss selling effects. Journal of Financial Economics 12, 89–104.
Rosenberg, B., Reid, K., Lanstein, R., 1984. Persuasive evidence of market inefficiency. Journal of Portfolio Management 11, 9–17.
Shleifer, A., Vishny, R., 1997. The limits of arbitrage. Journal of Finance 52, 35–55.
Skinner, D., Sloan, R., 2002. Earnings surprises, growth expectations, and stock returns, or, don't let an earnings torpedo sink your portfolio. Review of Accounting Studies 7, 289–312.
Stoll, H., Whaley, R., 1983. Transactions costs and the small firm effect. Journal of Financial Economics 12, 52–79.
Warther, V., 1995. Aggregate mutual fund flows and security returns. Journal of Financial Economics 39, 209–236.
Wurgler, J., Zhuravskaya, E., 2003. Does arbitrage flatten demand curves for stocks? Journal of Business 75, 583–608.