5-26-2012

On The Efficiency of US Equity Markets

Mikael Carl Erik Bergbrant
University of South Florida, research@bergbrant.com

On the Efficiency of U.S. Equity Markets

by

Mikael C. Bergbrant

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Business Administration
Department of Finance
College of Business
University of South Florida

Major Professor: Delroy M. Hunter, Ph.D.
Daniel J. Bradley, Ph.D.
Patrick J. Kelly, Ph.D.
Christos Pantzalis, Ph.D.

Date of Approval:
May 26, 2012

Keywords:
Asset Pricing, Market Integration, Market Efficiency,
Arbitrage, Idiosyncratic Risk, Risk Prices, EGARCH

DEDICATION

I dedicate this dissertation to my parents, Anders and Christina, and my fiancée Nina. Thank you for always being there, for supporting me, and for providing encouragement when I needed it the most. I love you and I could not have done it without you!

ACKNOWLEDGMENTS

I thank the entire faculty, students, and administrative staff of the Department of Finance at the University of South Florida. Your positive attitudes and insightful advice have made my student experience exciting and rewarding, both professionally and personally. Thank you for always being available to discuss ideas and provide feedback. A special thank you to my dissertation committee – Delroy Hunter (chair), Patrick Kelly, Daniel Bradley, and Christos Pantzalis – for guidance and support throughout the writing process. I owe additional gratitude to Patrick Kelly for encouraging me to take the academic path in finance and to Delroy Hunter for stimulating my intellectual curiosity as well as for the countless hours spent working with me.

This dissertation has benefited from the advice of my dissertation committee, as well as the advice of the entire Department of Finance at the University of South Florida. It has also benefitted from the comments of conference participants at the Southern Finance Association 2011 annual meeting, the Financial Management Association 2010 annual meeting, the Eastern Finance Association 2010 annual meeting, the Midwest Finance Association 2010 annual meeting, as well as the Southwestern Finance Association 2010 annual meeting. In addition, it has benefitted from the comments of seminar participants at Central Michigan University, St. Johns University, the University of Alabama in Huntsville, the University of Alberta, the University of the Pacific, and the University of Toledo. All remaining errors are mine.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Most papers in empirical finance implicitly or explicitly assume the same price of risk, for each priced systematic risk factor, across all risky assets within a given domestic market. In doing so, they rely on the assumption that markets are domestically integrated and, as such, that the price of risk is determined independently of individual investors attitude towards risk. This is true in frictionless markets where investors have complete information, homogenous beliefs, and hold the mean-variance efficient combination of the market portfolio and a risk-free asset. However, investors might not hold the market portfolio because of exogenous reasons. In fact, several recent papers have provided evidence that US investors do not, holding instead vastly undiversified portfolios. There are two main implications to the above. First, if one group of investors does not hold the market portfolio, then the remaining set of investors will also not be able to hold the market portfolio and will rationally expect to be compensated for bearing idiosyncratic risk. Therefore, idiosyncratic risk will be priced in expected returns. Second, the price of risk need not be the same across all assets in which case domestic markets are not integrated.

In the first essay titled “Is Idiosyncratic Volatility Really priced?” I show that the positive relation between idiosyncratic volatility (IV) and returns found by Fu (2009) only exists for firms that are difficult to arbitrage. The relation between IV and returns is strong for small and illiquid stocks, but decreases with size and liquidity and becomes

non-existent for the largest and most liquid firms. Furthermore, zero-cost portfolios based on IV and size do not yield positive returns when conservative trading costs are considered. This evidence is consistent with an efficient market, in which arbitragers exploit profitable investment opportunities and by doing so they prevent systematic mispricing in financial markets.

In the second essay titled "Are the U.S. Equity Markets Domestically Integrated?" I investigate whether the three main U.S. equity markets are domestically integrated by comparing the price of commonly used risk factors across the NYSE, Amex, and Nasdaq. I find that the markets have significantly different prices of risks for several risk factors, indicating that the markets are segmented. The magnitude of the difference is both statistically and economically significant, and is not due to arbitrage constraints or model misspecification. Instead, I find evidence consistent with the investor-segmentation hypothesis, in which different investors choose to hold different subsets of firms and demand different prices of risk among the different groups of securities. I do not find that segmentation is restricted to a specific time period. On the contrary, it is present in all sub-periods. In contrast to the results regarding the pricing of idiosyncratic volatility, these results highlight the value of diversification and suggest that domestic equity markets are not fully efficient.

IS IDIOSYNCRATIC VOLATILITY REALLY PRICED?

1. INTRODUCTION

Recent literature has found a strong relation between idiosyncratic volatility and returns (see, e.g., Malkiel and Xu (2002), Ang, Hodrick, Xing, and Zhang (2006), and Fu (2009)). The findings have puzzled academics as well as practitioners. If robust, such findings would have several important implications.

First, priced idiosyncratic risk would shake the fundamentals of financial theory, in particular asset pricing, which builds on the notion that diversification eliminates non-systematic risk and, therefore, only systematic risk factors are priced.

Secondly, such a relation would be inconsistent with efficient markets, in which arbitragers and analysts work to eliminate arbitrage opportunities. Inefficient markets would put individual investors at a disadvantage (since prices would no longer be informative of the underlying value of securities) and potentially increase the value of actively managed portfolios. Furthermore, it would suggest that there is an unexploited profit (risky-arbitrage) opportunity available for savvy investors, who would spend vast resources to try to forecast idiosyncratic volatility in order to take advantage of the apparent inefficiency.

Third, it would imply that markets are not allocationally efficient. If capital is not allocated where it has the greatest value, economic growth would be slower, and there would be a need for government intervention in the markets. The main objective of this paper is to re-examine the relation between idiosyncratic volatility and returns.

The appearance of priced idiosyncratic risk can be explained in different ways. One of the most common explanations is that idiosyncratic volatility appears priced because some pervasive systematic risk factor has been omitted from the model. Alternatively, if investors are underdiversified this could make them sensitive to the idiosyncratic risk of returns (see, e.g., Levy 1978; Merton 1987; Malkiel and Xu 2002).

It is a well-known fact that many investors are undiversified. Looking at retail investors, Goetzmann and Kumar (2004) find that more than 25% hold a single stock in their portfolios. Campbell, Lettau, Malkiel, and Xu (2001) argue that investors need approximately 50 stocks in their portfolios to be diversified. Goetzmann and Kumar (2004) find that less than 10% of retail investors' portfolios contain more than 10 stocks. This suggests that there could be a positive relation between idiosyncratic volatility and returns, and several papers have found such a positive relation.

Goyal and Santa-Clara (2003) find a strong positive link between total stock market volatility and returns (contradicting the results by Baillie and DeGennaro (1990)). Furthermore, they find that the lagged variance of the market has no predictive power for market returns implying that it must be the idiosyncratic component of volatility that influences returns. However, Wei and Zhang (2005) and Bali, Cakici, Yan, and Zhang (2005) show that these results are driven primarily by data in the 1990's and that the trading strategy used by Goyal and Santa-Clara (2003) does not yield positive returns for extended samples.

In contrast, Ang, Hodrick, Xing, and Zhang (2006) find that stocks with high idiosyncratic volatility have low expected returns, implying that idiosyncratic risk is negatively priced. They argue that papers that have found a positive link between returns and idiosyncratic volatility have not examined the idiosyncratic volatility at the firm level, or have not sorted their portfolios directly based on idiosyncratic risk. Jiang, Xu, and Yao (2009) argue that the results found by Ang et al. (2006) appear due to an inverse relation between idiosyncratic risk and future earning shocks, and the future earnings shocks are related to stock returns. In another paper Ang et al. (2009) find that high past idiosyncratic volatility is associated with lower future returns both domestically and internationally. Han and Lesmond (2009) counter that Ang et al. (2006) results are inflated due to the fact that they do not account for the occurrence of zero returns. Furthermore, Bali and Cakici (2008) find that the way idiosyncratic volatility is measured impacts the relation between idiosyncratic risk and returns.

Fu (2009) attempts to resolve these mixed findings. He argues that idiosyncratic volatility cannot adequately be described as a random walk process, and that using lagged idiosyncratic volatility therefore incorrectly estimates expected idiosyncratic volatility. He uses Exponential Generalized Autoregressive Heteroskedacticity (EGARCH) models to find the expected idiosyncratic volatility and finds a significant positive relation between the estimated conditional idiosyncratic volatilities and expected returns.

However, in order for idiosyncratic volatility to be priced, it is not sufficient that the majority of investors are underdiversified. If a few rational, diversified investors exist, that would be enough to ensure efficient pricing of securities due to risky arbitrage (Fama (1965)). Hence, in order for idiosyncratic volatility to be priced in efficient markets, it is necessary to have both undiversified investors *as well as* severe arbitrage constraints.

Several papers suggest that small stocks have higher arbitrage constraints than large stocks. Wurgler and Zhuravskaya (2002) find that small stocks tend to have higher arbitrage risk and find a highly significant, negative correlation between their measure of risk of arbitrage and size. Furthermore, Pontiff (2006) argues that small stocks, due to their high transaction costs, are more expensive to arbitrage. This implies that size is a good proxy for arbitrage constraints.

In light of the above, we re-examine the relation between expected idiosyncratic volatility and returns while accounting for arbitrage constraints. Specifically, we interact measures of idiosyncratic volatility with size and find that the relation between idiosyncratic volatility and return is very strong for small firms (the lowest size quartile of stocks traded on NYSE, Amex, and Nasdaq) between 1963 and 2009. However, the relation declines in magnitude and disappears completely for the largest quartile of stocks. This suggests that the relation between idiosyncratic volatility and returns only exists for firms that are difficult to arbitrage. The results hold both for Fu's (2009) measure of expected volatility as well as for realized volatility. Furthermore, we find that zero-cost portfolios based on idiosyncratic volatility only yield high positive abnormal returns for stocks that are risky and costly to arbitrage (small stocks). For large stocks, such portfolios do not yield significant abnormal returns. In addition, we find that even conservative estimates of trading costs make zero-cost portfolios based on idiosyncratic volatility unprofitable. These results are consistent with an efficient market in which arbitrage constraints prevents arbitragers from pricing some assets efficiently.

As mentioned, there are strong reasons to believe that size is a good proxy for the risks and costs to arbitrage. However, in sensitivity tests, we divide the firms into quartiles based on liquidity and find very similar results. For the least liquid stocks there is a strong, positive relation between expected idiosyncratic volatility and returns, but this relation disappears completely (becomes insignificant) for the most liquid firms in the sample.

This paper contributes to the literature, on the pricing of idiosyncratic volatility in several ways. First, it adds to the literature that argues that idiosyncratic volatility appears to be positively related to returns by explicitly considering arbitrage constraints, one of the factors that give rise to the appearance that idiosyncratic risk is priced. Accounting for arbitrage, we find that such a relation is driven by stocks that are difficult to arbitrage and is non-existent for the largest, most frequently traded stocks. Secondly, our results suggest that sample selection influences the relation between idiosyncratic volatility and returns and, as such, might help to explain the mixed results in the literature. In addition, it suggests that the relation between idiosyncratic volatility and returns that has been found is consistent with efficient markets.

The reminder of the essay is organized as follows. Section two discusses the data and method of estimating the idiosyncratic volatility measures that are used in the paper. In Section three, we present and discuss the results. Section four investigates potentially different explanations and provides robustness tests. Finally, Section five concludes.

2. DATA AND SUMMARY STATISTICS

To conduct the various tests, four sets of data are required. The first set contains the test asset, returns. The second set contains the systematic risk factors that have been shown to impact the cross-section of returns. The third set contains measures of idiosyncratic volatility, and the forth includes proxies for the difficulties to arbitrage.

The data are obtained from three sources. Daily and monthly individual security data for all firms traded on NYSE, Amex, and Nasdaq from 1963 to 2009 is obtained from the Center for Research in Security Prices (CRSP). We obtain data on stock returns (RET), prices (P), shares outstanding (SHROUT), volume (VOL) and bid-ask spreads (SPREAD). We also obtain value weighted index returns from CRSP. Accounting data, or more specifically book values, are obtained from Compustats annual fundamentals file. The Fama and French (1993) factors, Small-Minus-Big (SMB) and High-Minus-Low (HML), as well as the market premium and the proxy for the risk-free rate are obtained through Kenneth French's webpage[1].

A. Test asset

The test assets used in this paper are individual stock returns for all firms traded on the three major US exchanges. Monthly holding period returns are obtained through CRSP. Following conventions in the literature (Fu (2009)) on predicting idiosyncratic volatility, we exclude returns in excess of 300%. Holding period returns (RET) include capital gains as well as dividend yields.

[1] We thank Kenneth French for making the data available.

B. Systematic risk factors

Fama and French (1992) show in their seminal paper that size and book-to-market ratio are able to explain the cross-section of stock returns. In this paper, we estimate beta, size, and book-to-market ratio as in Fama and French (1992) for each of our securities. Beta, size, and book-to-market are calculated in June every year, and used from July in that year until June in the following year.

To calculate market beta (BETA), monthly stock returns over the previous 60 months are regressed on the CRSP value weighted index return to estimate individual firm betas. We then create 10 size portfolios based on market capitalization in June for stocks traded on NYSE. For each size decile, we form 10 beta portfolios based on the estimated betas for the individual firms. The procedure rolls forward every year.

For each of the 100 size-β portfolios, we calculate the equally weighted portfolio return every month and regress the monthly portfolio returns on the current and lagged (one period) value weighted market index. The final portfolio beta is computed by summing the two slopes (the coefficient on the contemporaneous and lagged value weighted market index return). Finally, we allocate the beta of a size-β portfolio to each stock that was included in that portfolio in each month. Summing the slopes is suggested by Dimson (1979) to adjust for non-synchronous trading.

Size (ME) is measured by the market value of equity (shares outstanding multiplied by price in June every year). BE/ME is the book-to-market ratio, calculated using December (t-1) market value of equity and the book value of equity for the fiscal year that ended in year t-1. The time lag used is to ensure that the information was available to investors at the time it is used in the analysis.

Since Fama and French developed their factors in 1992, much research has been focused on explaining cross-sectional returns. Liquidity and momentum are two of the most frequently used variables that have been shown to impact cross-sectional returns. Similarly, since Amihud and Mendelson (1986) introduced liquidity as a factor, several papers have since shown that it is important in the cross-section of returns (see, e.g., Brennan and Subrahmanyam 1996; Datar, Naik, and Radcliffe 1998; Chorida, Subrahmanyam and Anshuman 2001; Amihud 2002; Pastor and Stambaugh 2003). Jagadeesh and Titman (1993) show that a momentum effect is present in stock price movements and Jagadeesh and Titman (2001) find that the effect has remained significant long after it was discovered. Therefore, we augment the Fama and French (1992) three-factor model with proxies of these factors.

We measure momentum and liquidity as in Chordia, Subrahmanyam and Anshuman (2001). The measure of momentum, RET (-2,-7) is calculated as the average holding period return between month t-7 to t-2. We exclude the last month's (t-1) return to avoid returns merely due to bid-ask bounce. For liquidity, we use two different measures. Turnover (TURN) is calculated as the average monthly share turnover in the past 36 months. Monthly share turnover is calculated as the monthly volume divided by the number of shares outstanding. The second measure of liquidity, CVTURN, is the coefficient of variation of those turnovers over the same 36 months. As can be seen in the variable definitions, all systematic risk factors are estimated using data up until t-1 and are therefore predetermined at time t.

C. Idiosyncratic volatility

In order to investigate the relation between returns and idiosyncratic volatility, we must first have a good measure of idiosyncratic volatility. Many previous papers looked at the relation between returns and lagged idiosyncratic volatility as a proxy for expected idiosyncratic volatility, implicitly assuming that idiosyncratic volatility follows a random walk (Ang, Hodrick, Xing, and Zhang (2006)). However, Fu (2009) points out that this is problematic. He rejects the null hypothesis of a random walk for the idiosyncratic volatility for 90% of the firms in his sample and argues that forecasting using models in the GARCH family would do better in estimating future volatility. Particularly, he suggests that the EGARCH model proposed by Nelson (1991) would provide a superior model of expected idiosyncratic volatility since it captures the asymmetric properties of volatility (Dennis, Mayhew, and Stivers (2006) show the importance of asymmetric volatility effects and investigate the effect further by explicitly distinguishing between surprises in systematic and idiosyncratic volatility). The EGARCH models are also more flexible than other ARCH and GARCH models and do not restrict the parameters to avoid negative values.

Guo, Ferguson, and Kassa (2010) criticize the expected idiosyncratic volatility measure used by Fu (2009), for being estimated in a way that introduces a look-ahead bias. Fu (2010) re-estimate his model with completely out-of-sample data and reinforces the strong positive relation between idiosyncratic volatility and returns. He further finds that the measure of expected idiosyncratic volatility is qualitatively insensitive to the bias suggested by Guo, Ferguson, and Kassa (2010).

We measure idiosyncratic volatility in two ways. The first (main) measure of idiosyncratic volatility is the expected idiosyncratic volatility (E(IVOL)) obtained from EGARCH models (Nelson (1991)). Pagan and Schwert (1990) compare different GARCH models and they find that the EGARCH model does the best in explaining monthly return volatility. Explicitly, the EGARCH model that we fit is:

$$R_{i,t} - RF_t = a_i + b_i(R_{M,t} - R_{RF,t}) + s_i \mathrm{SMB}_t + h_i \mathrm{HML}_t + \varepsilon_{i,t}, \ \varepsilon_{i,t} \sim \mathrm{N}\left(0, \sigma_{\mathrm{i,t}}{}^2\right) \tag{1}$$

$$\ln \sigma_{\mathrm{i,t}}^2 = a_i + \sum_{l=1}^{p} \mathrm{b}_{\mathrm{i},l} \ln \sigma_{\mathrm{i,t}-l}^2 + \sum_{k=1}^{q} \mathrm{c}_{\mathrm{i,k}} \left(\theta \left(\frac{\varepsilon_{\mathrm{i,t-k}}}{\sigma_{\mathrm{i,t-k}}} \right) + \Upsilon \left(\left| \frac{\varepsilon_{\mathrm{i,t-k}}}{\sigma_{\mathrm{i,t-k}}} \right| - \left(\frac{2}{\pi}\right)^{1/2} \right) \right) \tag{2}$$

Following Fu (2009) we estimate all possible EGARCH (p,q) models in which $1 \leq p \leq 3$ and $1 \leq q \leq 3$ for a total of 9 possible models per firm and month. Thus, each EGARCH model has p+q+3 parameters. The goal with this is to estimate the conditional variance σ^2 at time t. We use the best fitting, according to the Akaike Information Criterion (AIC), converging model for each firm and month. French, Schwert, and Stambaugh (1987), as well as Fu (2009), find that using time-varying parameters do not change their results. However, estimating the model during the entire sample period does induce a look-ahead-bias. We therefore estimate the EGARCH models using expanding windows. We require a minimum of 30 months of prior return data in order to estimate the expected idiosyncratic volatility for the following month.[2] The conditional idiosyncratic volatility, denoted E(IVOL) will be used in the cross-sectional tests as the expected idiosyncratic volatility.

[2] Consistent with French, Schwert, and Stambaugh (1987), as well as Fu (2009), we get qualitatively similar results when estimating the parameters using the entire time period.

The second measure of idiosyncratic volatility is the realized idiosyncratic volatility (IVOL). In order to obtain this, we regress daily stock returns (for each individual security) on the daily Fama and French (1993) factors in every month. Han and Lesmond (2009) raise the concern that infrequent trading drive parts of the idiosyncratic volatility effect. In order to mitigate this issue we require that, in a given month, a stock has at least 15 observations of non-zero trading volume to be eligible for estimation.

For each security, in each month, we estimate the realized idiosyncratic volatility using the following time series regression:

$$R_{i,t} - RF_t = a_{i,t} + b_{i,t}(R_{M,t} - RF_t) + s_{i,t}\text{SMB}_t + h_{i,t}\text{HML}_t + \varepsilon_{i,t} \qquad (3)$$

$R_{i,t}$ is the return (in percent) on security i on day t. RF_t is the proxy for the risk-free rate of return. b, s, and h are factor sensitivities to each of the Fama and French (1993) factors. We calculate IVOL, for a given month, as the standard deviation of the daily residuals from the within-month regression. Monthly idiosyncratic volatilities are calculated by multiplying the daily standard deviation of the residuals by the square root of the number of trading days in that month.

D. Proxies for the difficulties to arbitrage

The main idea in this paper is that if markets are efficient, idiosyncratic volatility should not be related to returns for those firms that are easy to arbitrage. As a proxy for arbitrage costs, we use firm size (ME). Although many proxies have been used in the literature, size is highly correlated with all proxies. Several papers suggest that small stocks have higher arbitrage risk than large stocks. Wurgler and Zhuravskaya (2002) find that small stocks tend to have higher arbitrage risk and find a very highly significant negative correlation between their measure of risk of arbitrage and size. Furthermore, Pontiff (2006) argues that small stocks, due to their high transaction costs, are more expensive to arbitrage. This implies that size is a good proxy for the difficulty to arbitrage. Another proxy for difficulties to arbitrage is liquidity. Less liquid firms are risker to arbitrage due to their higher trading costs, and we would therefore expect the relation between idiosyncratic risk and returns to be more pronounced for firms with low liquidity.

Due to limitations in accounting data prior to July of 1963, we have limited the sample period to be between July of 1963 and December of 2009. This results in 3,177,998 firm-month observations over the 558 month sample window. All variables, except for returns and beta, are winsorized at the .5% and 99.5% in order to avoid giving extreme observations heavy weight. Variables that have skewness greater than 3 are transformed into their natural logarithm.

E. Summary statistics

Table 1.1 reports the summary statistics for the complete sample consisting of more than three million firm-month observations. The summary statistics are very close to those presented in Fu (2009), even though the sample used in his paper has been extended by three years to December 2009. The arithmetic mean monthly return is 1.08%, and the mean excess monthly return is .64%. The arithmetic mean realized idiosyncratic volatility (IVOL) is 12.57% while the mean expected idiosyncratic volatility (E(IVOL)) is 11.56%. Both measures of idiosyncratic volatility have a large spread, with standard deviations in excess of 70% of their arithmetic means. The average arithmetic beta for the firm-months is 1.27. The interaction term E(IVOL)×LN(ME) has a mean of 21.94 and a standard deviation of 27.34. The high variation in both the idiosyncratic volatility measures, as well as the interaction term implies that the data is rich in variability, and could do a good job of explaining the variation in the cross-sectional returns.

3. RESULTS

In this section we analyze the data and present the major results of the paper. We start out by looking at univariate results in Section 3.A. This analysis is followed by a more thorough analysis using cross-sectional regressions in Section 3.B. Finally we investigate portfolio strategies based on expected idiosyncratic volatility and size in Section 3.C

A. Univariate tests

Table 1.2 presents the time-series means of the simple cross-sectional Pearson correlation coefficients. In every month we obtain the contemporaneous Pearson correlation coefficients between each pair of variables using the cross-section of stocks. The reported correlation coefficients are the average over the sample period of 558 months. Correlation coefficients that are significant at 1%, given their time-series standard error, are marked with two asterisks (**). As would be expected, and in accordance with Fama and French (1992) as well as Fu (2009), book-to-market has a positive correlation with returns while the relation between returns and beta is insignificant. Furthermore, size is negatively correlated with returns, even though the relation is statistically insignificant in our sample.

Consistent with Fu (2009), we find a significant positive correlation between measures of idiosyncratic volatility (both realized and expected) and returns. In sharp contrast the interaction terms E(IVOL)×LN(ME) and IVOL×LN(ME) are both negatively correlated with returns. Interestingly, we also find that returns and liquidity are negatively correlated. Given that these variables are proxies for arbitrage they suggest that Fu's (2009) results might not be robust. All other correlation coefficients are as expected, and congruent with what has been shown in previous literature. Returns are positively correlated with past returns, indicating the presence of a momentum effect. None of the correlation coefficients are large enough to give rise to a concern about multicollinearity in the regressions to follow.

B. Cross-sectional regressions with Fama-MacBeth (1973) T-stats

The preliminary evidence from the correlation analysis implies that when we account for arbitrage constraints, the positive relation between returns and idiosyncratic volatility is not robust. In this section, we take a more rigorous look at the issue. We follow the methodology of Fama and MacBeth (1973) and run cross-sectional regressions in every month from July of 1963 to December of 2009, a total of 558 months. The dependent variable in the monthly regressions is the holding period return on all firms in our sample that have information on all explanatory variables in that month. The explanatory variables are as defined in Section 2. Note, as stated in the variable definitions, that all the explanatory variables (except for realized idiosyncratic volatility) are predetermined at time t with respect to the return on firm i, $R_{i,t}$.

More explicitly, for every month, we estimate a model that is nested in the following cross- sectional regression,

$$\mathrm{R}_{\mathrm{i,t}} = a_t + \beta_1 BETA_{i,t} + \beta_2 LN(ME)_{i,t} + \beta_3 LN\left(\frac{BE}{ME}\right)_{i,t} + \beta_4 E(IVOL)_{i,t} + \beta_5 RET(-7,-2)_{i,t} + \beta_6 LN(TURN)_{i,t} + \beta_7 LN(CVTURN)_{i,t} + \beta_8 E(IVOL) * LN(ME)_{i,t} + \varepsilon_{i,t}\ , \tag{4}$$

where $\mathrm{R}_{\mathrm{i,t}}$ is the gross return on security i in month t. The explanatory variables are described in Section 2. The variables of interest are E(IVOL) and E(IVOL)×LN(ME).

E(IVOL) is the expected idiosyncratic volatility in the month obtained from EGARCH models, and E(IVOL)×LN(ME) represents the interaction term between the expected idiosyncratic volatility and size. If the appearance of priced idiosyncratic volatility is due to costly arbitrage, we would expect that including the interaction term decreases the effect, since larger firms would be easier to arbitrage.

The coefficients reported are the mean coefficients from the above regression estimated from July of 1963 until December of 2009, for a total of 558 months. Similarly, the reported R-squared is the average R-squared from the monthly regressions. The number of securities included, i, varies every month. We calculate time series standard errors using the estimator in Newey and West (1987) with three lags.

The test-statistics reported in Table 1.3 are the mean slopes $\overline{\beta_k}$ divided by their time series standard error. We first replicate the main findings of Fama and French (1992) in our extended sample and find that size (ME) and book-to-market (BE/ME) help to explain the cross-sectional variation in returns (Model 1). As shown by Fama and French (1992), size is negatively related to returns, while book-to-market is positively related to returns in the cross-section. We also find, in accordance with Fama and French (1992), that the relation between return and market beta is flat (insignificant test statistic).

Since Fama and French developed their factors in 1992, much research has been focused on explaining cross-sectional returns. Liquidity and momentum are two of the most frequently used variables that have been shown to impact cross-sectional returns. Adding these variables (measured as described in Section 2.B) to the model is important since it decreases the chance that a relation between idiosyncratic volatility and return would simply be the result of idiosyncratic volatility being a proxy for some omitted, priced factor. Adding liquidity and momentum to the cross-sectional regressions (Model 2) increases the proportion of explained variation in returns, but does not alter the signs of any of the coefficients, congruent with prior literature. As expected, the relation between liquidity (TURN and CVTURN) and return is negative while the relation between momentum (RET -7, -2) and return in positive.

In Models 3 and 4, we replicate the main finding of Fu (2009) and find that expected idiosyncratic volatility, as estimated by the EGARCH model, is an influential determinant in the cross-section of returns. Like Fu (2009) we find that the sign of the relation is positive, so that higher expected idiosyncratic volatility is associated with higher returns. The coefficients on the measure of expected idiosyncratic volatility are statistically significant and large in magnitude. Further, the explanatory power of idiosyncratic volatility remains after controlling for momentum and liquidity (Model 4). Like Fu, we find that adding idiosyncratic volatility to the test changes the sign of the size coefficient. This implies that the negative relation between size and returns found when omitting IV is due to omitted factors that affect small firms adversely. After including such a factor, expected idiosyncratic volatility, the relation between size and returns is positive. This is consistent with Merton (1987) who suggests that, all else equal, large firms should be outperforming small firms.

The main interest in the cross-sectional regression are the results from models 5 and 6, and particularly the coefficient on the interaction term between size and expected idiosyncratic volatility (E(IVOL)×LN(ME)). We find that the interaction term is negative and significantly different from zero. This implies that the relation between returns and idiosyncratic volatility decreases with size. The coefficient remains strongly significant and economically meaningful in all model specifications. In regressions 7 and 8 we substitute realized idiosyncratic volatility (IVOL) for E(IVOL) in Equation 4, and show that the relation also holds for realized idiosyncratic volatility.

With the interaction term the coefficients on E(IVOL) can no longer be interpreted as their stand-alone figures, but need to be interpreted at a certain level of the variable with which it is interacted (LN(ME)). This gives the "net effect" of E(IVOL) given a certain level of size. We analyze the net effect of E(IVOL) on returns at the average size (LN(ME)):

$$\frac{\partial R_{i,t}}{\partial E(IVOL)} = \widehat{\beta_{4_t}} + \widehat{\beta_{8_t}} \times \overline{LN(ME)_{\iota,t}}. \tag{5}$$

Table 1.4 illustrates the relation between idiosyncratic volatility and return for firms in four size quartiles based on NYSE breakpoints. As is clearly evident, the relation between idiosyncratic volatility and return decreases as size increases. For the smallest NYSE quartile of firms (quartile 1), the models suggest that the net effect of idiosyncratic volatility on returns is about twice as large as for the 2rd NYSE quartile, and about 2-4 times larger than when compared to the 3rd quartile. More interesting, the net impact of idiosyncratic volatility on returns approaches zero for firms in the largest NYSE quartile (slightly negative in Model 5 and slightly positive in Model 6). Since the natural logarithm of size (LN(ME)) has been used in the interaction term in the model specification, this suggests that the relation between expected idiosyncratic volatility and returns decreases logarithmically with size. However, the relation could take other forms. In the robustness Section 4.A, we allow for the relation between idiosyncratic volatility and returns to depend on difficulties to arbitrage in other ways by re-estimating the cross-sectional regressions for different size quartiles.

Given that size is a good proxy for costs of arbitrage, this lends strong support to the idea discussed in the introduction to this paper, that idiosyncratic volatility is not priced for firms that are easy to arbitrage. This further implies that the relation between idiosyncratic volatility and returns can differ for different samples. Samples that are dominated by firms that have low arbitrage costs such as large firms, firms that are well covered in the media, or firms that have a substantial amount of accounting data available, would be likely to display a weaker relation between returns and E(IVOL) than those samples comprised of all stocks. This can explain why several previous authors have found conflicting results on the relation between returns and idiosyncratic volatility. Furthermore, this is an important consideration when researching idiosyncratic volatility internationally. Selection biases in data availability internationally could greatly influence any inferences on the relation between idiosyncratic volatility and returns.

C. Portfolios formed on size and expected idiosyncratic volatility

The previous results indicate that the relation between idiosyncratic volatility and returns depends on arbitrage constraints, as proxied by size. We therefore form portfolios based on size and expected idiosyncratic volatility and re-examine this relation. The advantage of the portfolio approach is that it allows for easily interpreted returns on feasible investment portfolios.

Every year we sort the stocks on size and assign them into five size portfolios. Each month, we sort the stocks in each size portfolio on their expected idiosyncratic volatility for the following month and then assign them into five portfolios. Hence, each month we have 25 portfolios based on size and expected idiosyncratic volatility. We then compute the value weighted monthly time-series returns for each portfolio. If stocks that have high expected idiosyncratic volatility outperform those that do not, then zero-cost portfolios formed by purchasing the portfolio with highest idiosyncratic volatility and shorting the portfolio with lowest idiosyncratic volatility should yield positive returns. Fu (2009) finds that this is the case. In this paper we examine if such a strategy is profitable only for small firms by using the double-sorted portfolios. The results are presented in Panel A of Table 1.5. We find that the zero-cost portfolio returns for portfolios of small sized stocks are very large, but approach zero for large stocks. This strongly supports the idea that there is no relation between idiosyncratic volatility and returns for large stocks. Large stocks are easy to arbitrage, and even if underdiversified retail investors hold these stocks, any non-systematic risk premium would quickly disappear through arbitrage.

Furthermore, regressing the portfolio's returns on the Fama-French (1993) three factors also yield high positive Jensen (1968) alphas for the small stock/high idiosyncratic volatility portfolios (Table 1.5, Panel B). For the large stock and high expected idiosyncratic volatility portfolios the alphas are statistically insignificant and economically small. This lends further support to the idea that any relation between idiosyncratic volatility and return is limited to stocks which are difficult to arbitrage. A closer look at the table shows that the portfolio alphas are negative for all small stock portfolios, except for those in the highest quintile of idiosyncratic volatility. Hence, not only is the relation between idiosyncratic volatility and returns limited to small firms, but also the portfolio alphas imply that the relation is limited to those stocks in the highest quintile of idiosyncratic volatility.

However, it is important to bear in mind that trading costs are not considered for these portfolios. Since trading costs are much higher (as a percentage) for small stocks with high idiosyncratic volatility (Kelly (2005)), it would be interesting to look at a trading strategy including trading costs. Unfortunately, there are no bid-ask spreads available for most of the stocks used in the sample in this paper. However, CRSP provides bid-ask spreads for stocks traded on Nasdaq. Limiting the sample to stocks traded on Nasdaq with non-missing values of the bid-ask spread yields a subsample of 511,485 firm-month observations. Every year we divide those stocks into 4 portfolios based on size and for each size quartile, we break them up in four expected idiosyncratic volatility quartiles (monthly). We then look at the average returns for each of the 16 (4×4) portfolios and find very similar results to those in the complete sample (Table 1.6, Panel A). Every month a stock appears in a portfolio *which it was not prior included in* we calculate the (one-way) trading cost as the bid-ask spread divided by two times the stock price, and add that cost to the new portfolio that the stock is included in (hence, we do not account for a cost at the time of the sale). This is a very conservative measure of trading costs, considering it only takes into account purchasing costs and not selling costs. Furthermore, it ignores any brokerage costs that would be charged. The results, presented in Panel B of table 1.6 indicate that the ME-E(IVOL) double-sorted portfolios do not yield substantial positive returns. The value weighted return for the portfolio with the smallest stocks and the highest idiosyncratic volatility actually becomes negative. We are aware that an unlimited number of potential portfolios could be formed and that the results given the 4×4 portfolios[3] is not exclusive evidence that it is impossible to form profitable portfolios based on size and idiosyncratic volatility. Trading strategies using 2×2 and 3×3 portfolios were also tested, with very similar results (not reported, but

[3] Due to very high variability in expected idiosyncratic volatility over time, approximately 55% of stocks are traded (change portfolios) in a given month using 4×4 portfolios. When using fewer portfolios, the proportion of stock traded every months decrease, but the qualitative results remain.

available from the authors on request). However, considering the extremely conservative measure used for trading costs here, and the negative returns yielded by the portfolios, it seems like it would be difficult for such trading strategies to be beneficial. These results imply that it is difficult, if not impossible, for diversified investors to benefit from the positive relation between idiosyncratic volatility and returns by short term trading. This is consistent with an efficient market.

4. ROBUSTNESS TESTS

In this section we perform three robustness tests. Section 4.A provides tests that allow for size to have an effect on the relation between idiosyncratic volatility and returns that is not linear. In Section 4.B we show that the findings in this paper, that the relation between idiosyncratic volatility and returns depend on difficulties to arbitrage is robust to another measure of the difficulties to arbitrage. Finally, in Section 4.C we show that the results are robust to the choice of time frame for the analysis.

A. Sample broken down by size quartiles

One argument that could be made against the previous tests, especially those in Table 1.3, is that the variation in idiosyncratic volatility for large firms is not sufficient to give power to these tests. Table 1.7 shows summary statistics broken up by size quartiles formed based on NYSE stock breakpoints over the period from 1963 to 2009. NYSE breakpoints are used in order to avoid small firms dominating after the inclusion of Nasdaq stocks in 1973 (for more on this issue, see Fama and French (1992)).

The mean expected idiosyncratic volatility decreases monotonically over the size quartiles, confirming the negative correlation between size and idiosyncratic volatility shown in Table 1.2. However, of primary interest is not the mean, but the standard deviation of the variables across the size quartiles. As the evidence in the table indicates, the standard deviations of E(IVOL) is large across all the size quartiles (greater than 4). This should be more than sufficient to generate reliable coefficients and test statistics for the cross-sectional regressions. Furthermore, and of equal interest, is the standard deviation of the interaction term (E(IVOL)×LN(ME)). The standard deviation is very uniform across the size quartiles, lending further support to the validity of the coefficients produced by the cross-sectional regressions.

As mentioned in Section 3.B, the previous cross-sectional regressions suggest that the relation between idiosyncratic volatility and returns decreases logarithmically with size. However, this does not have to be the case. To further test the validity of the results presented above, we re-estimate Models 6 and 8 from Table 1.3 for each size quartile. The results from the regressions are presented in Table 1.8, and the net E(IVOL) effect is the net effect of E(IVOL) evaluated at the mean size for each size quartile.

Once again, it is important not to focus too much on the significance of the idiosyncratic volatility measure alone, since it has to be interpreted in combination with the interaction term. Looking at the net effect, we see that the relation between both E(IVOL) and returns (Panel A) and IVOL and returns (Panel B) decrease monotonically across the size quartiles and reaches 0 for the largest quartiles[4].

In addition, liquidity (as measured by Turnover) appears to have a much stronger explanatory power for returns among small firms. While liquidity appears to have a strong influence on returns among small firms (both statistically and economically significant), it becomes insignificant and economically small for large firms. These results support and add to the evidence that the relation between idiosyncratic volatility and returns is limited to small firms which are difficult to arbitrage.

[4] Even though the interaction term is insignificant in two of the specifications, it has been included for consistency. We have also estimated the regressions without the interaction term (LN(ME)×E(IVOL)) with the same qualitative results.

B. Different measures of the risk and costs of arbitrage

Yet another argument that could be raised against the above results, in particular against the finding of an insignificant relation between idiosyncratic volatility and returns for large firms, is that the relations using size as a proxy for difficulties to arbitrage do not hold for other proxies for difficulties to arbitrage. Here, we use liquidity as an alternative measure for difficulties to arbitrage. We divide the sample into four quartiles based on liquidity, as represented by turnover, and re-estimate the cross-sectional regressions for each of the liquidity quartiles. The results are presented in Table 1.9. We find that the same results that hold when sorting on size also holds when sorting into liquidity quartiles. For the least liquid stocks, the relation between expected idiosyncratic volatility and returns is very strong (both statistically and economically significant), but it decreases in magnitude over the liquidity quartiles. For the quartile of stocks that are most liquid, the relation between expected idiosyncratic volatility and returns disappear completely (it becomes statistically insignificant).[5]

C. Relation between idiosyncratic volatility and return over time

It has been argued that relations such as those found here are sensitive to the sample period (Wei and Zhang (2005)). We, therefore, split the sample into three sub-periods, 1963-1978, 1979-1994, and 1995-2009, and re-estimate model 6 from Table 1.3 for each sub-period. The results are reported in Table 1.10. We find that the inferences from above are insensitive to the time period used. In fact, the effect of E(IVOL) on

[5] As seen in table 9, we include an interaction term between our liquidity measure and idiosyncratic volatility, to be consistent with previous tests where we included interaction between size and idiosyncratic volatility. Considering that the interaction terms are insignificant in all the model specification, we re-estimate the regressions without the interaction terms and obtain identical inferences regarding the effect of E(IVOL). The coefficient on E(IVOL) is negative and insignificant for the quartile of firms that are most liquid. For the remaining quartiles, the coefficients on E(IVOL) are positive and significant, but decrease with liquidity.

returns as well as the effect of E(IVOL)×LN(ME) remains highly statistically and economically significant in all sub-period. The results also hold for other choices of sub-periods (not reported).

5. CONCLUSION

Many investors seem to hold undiversified portfolios. However, in an efficient market, this would not be sufficient to induce a relation between idiosyncratic volatility and returns. Instead, in an efficient market, diversified traders would take advantage of the opportunity and eliminate such a relation. However, if the market is efficient, but arbitraging is difficult or costly, this could prevent diversified traders from trading in certain assets, allowing a relation to emerge between idiosyncratic volatility and returns. This paper shows this to be the case. Specifically, we find that firms that are easy to arbitrage (such as large firms and liquid firms) exhibit *no* relation between idiosyncratic volatility and returns. The result is robust when we use different approaches to examine the relation, when we use alternate proxies for idiosyncratic volatility, and when we estimate the relation over different time periods.

The tests and evidence in this paper reconcile, at least in part, the opposing findings of the relation between idiosyncratic volatility and returns that currently exist in this fast growing literature on idiosyncratic volatility. We show that a positive (or negative) relation does not exist for all securities, but depend on the cost and risk associated with arbitrage for each security. Hence, the relation depends partly on the sample. Only investigating large stocks shows a much weaker relation than only investigating small stocks. Since previous papers look at different samples, this can help to clarify why researchers have found conflicting evidence on the relation between idiosyncratic volatility and returns. Finally, the findings in this paper are consistent with efficient markets.

6. REFERENCES

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time- series effects. *Journal of Financial Markets* 5, 31–56.

Amihud, Y., Mendelson, H., 1986. Asset pricing and the bid-ask spread. *Journal of Financial Economics* 15, 223–249.

Ang, A., Hodrick, R., Xing, Y., Zhang, X., 2006. The cross-section of volatility and expected returns. *Journal of Finance* 61, 259-299.

Ang, A., Hodrick, R., Xing, Y., Zhang, X., 2009. High idiosyncratic volatility and low returns: international and further U.S. evidence. *Journal of Financial Economics* 91, 1-23.

Bali, T., Cakici, N., 2008. Idiosyncratic volatility and the cross section of expected returns. *Journal of Financial and Quantitative Analysis* 43, 29-58.

Bali, T., Cakici, N., Yan, X., Zhang, Z., 2005. Does idiosyncratic volatility really matter? *Journal of Finance* 60, 905-929.

Baillie, R., DeGennaro, R., 1990. Stock returns and volatility. *Journal of Financial and Quantitative Analysis* 25, 203-214.

Brennan, M., Subrahmanyam, A., 1996. Market microstructure and asset pricing: on the compensation for illiquidity in stock returns. *Journal of Financial Economics* 41, 441–464.

Campbell, J., Lettau, M., Malkiel, B., Xu, Y., 2001. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *Journal of Finance* 56, 1-43.

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

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

Dennis, P., Mayhew, S., Stivers, C., 2006. Stock returns, implied volatility innovations, and the asymmetric volatility phenomenon. *Journal of Financial and Quantitative Analysis* 41, 381-406.

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

Fama, E., 1965. The behavior of stock market prices. *Journal of Business* 38, 34-105

Fama, E., French, K., 1992. The cross-section of expected stock returns. *Journal of Finance* 47, 427-465.

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

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

French, K., Schwert, G., Stambaugh, R., 1987. Expected stock returns and volatility. *Journal of Financial Economics* 19, 3–29.

Fu, F., 2009. Idiosyncratic risk and the cross-section of expected stock returns. *Journal of Financial Economics* 91, 24-37

Fu, F., 2010. On the Robustness of the Positive Relation between Expected Idiosyncratic Volatility and Expected Return. *Working Paper.*

Goetzmann, W., Kumar, A., 2004. Why do individual investors hold under-diversified portfolios? *Working paper*. Yale University.

Goyal, A., Santa-Clara, P., 2003. Idiosyncratic risk matters! *Journal of Finance* 58, 975-1007

Guo, H., Ferguson, M., Kassa, H., 2010. On the Relation between EGARCH Idiosyncratic Volatility and Expected Stock Returns. *Working Paper*.

Han, Y., Lesmond, D., 2009. Liquidity biases and the pricing of cross-sectional idiosyncratic volatility. *Review of Financial Studies,* Forthcoming.

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

Jegadeesh, N., Titman, S., 2001. Profitability of momentum strategies: An evaluation of alternative explanations. *Journal of Finance* 56, 699-720

Jensen, M., 1968., "The Performance of mutual funds in the period 1945-1964" *Journal of Finance* 23, 389-416

Jiang, G., Xu, D, Yao, T., 2009. The information content of idiosyncratic volatility. *Journal of Financial and Quantitative Analysis* 44, 1-28.

Kelly, P., 2005. Information efficiency and firm-specific return variation. *Working paper*. University of South Florida.

Levy, H., 1978. Equilibrium in an imperfect market: a constraint on the number of securities in the portfolio. *American Economic Review* 68, 643–658.

Malkiel, B., Xu, Y., 2002. Idiosyncratic risk and security returns. *Working paper*. University of Texas at Dallas.

Merton, R., 1987. A simple model of capital market equilibrium with incomplete information. *Journal of Finance* 42, 483–510.

Nelson, D., 1991. Conditional heteroskedasticity in asset returns: a new approach. *Econometrica* 59, 347–370.

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

Pagan, A., Schwert, W., 1990. Alternative models for conditional stock volatility. *NBER Working Paper* # 2955.

Pastor, L., Stambaugh, R., 2003. Liquidity risk and expected stock returns. *Journal of Political Economy* 111, 642–685.

Pontiff, J., 2006. Costly arbitrage and the myth of Idiosyncratic Risk. *Journal of Accounting and Economics* 42, 35-42.

Wei, S., Zhang, C., 2005. Idiosyncratic risk does not matter: A re-examination of the relation between average returns and average volatilities. *Journal of Banking & Finance* 29, 603-621.

Wurgler, J., Zhuravskaya, E., 2002. Does arbitrage flatten demand curves for stocks. *Journal of Business* 75, 583-608

Table 1.1: Summary statistics

Variable	Mean	Median	Std dev.	Q1	Q3	Skewness	N
RET (%)	1.08	0.00	17.00	-6.45	6.85	2.40	3177998
RET-RF (%)	0.64	-0.42	17.01	-6.90	6.42	2.40	3177998
BETA	1.27	1.24	0.36	0.98	1.50	0.19	2462043
LN(ME)	2.10	1.97	2.08	0.59	3.49	0.32	3034560
LN(BEME)	-0.41	-0.39	1.09	-1.02	0.18	0.50	1918200
RET(-2,-7)	1.06	0.41	6.87	-2.59	3.49	2.70	2982871
LN(TURN)	1.46	1.46	1.12	0.72	2.22	-0.11	2215612
LN(CVTURN)	4.13	4.13	0.47	3.82	4.43	0.16	2212543
LN(AMILL)	-1.30	-1.05	3.05	-3.40	0.98	-0.30	2609138
E(IVOL)	11.56	9.36	8.40	6.27	14.19	2.65	2196801
IVOL	12.57	9.49	10.77	5.80	15.65	2.72	2873083
E(IVOL)×LN(ME)	21.94	18.47	27.34	6.65	33.68	1.40	2190425
IVOL×LN(ME)	22.16	18.28	28.27	7.36	32.47	1.88	2731531

Table 1.1 shows descriptive statistics for the pooled sample. RET is the average monthly holding period return. RET-RF is the average holding period return minus the proxy for the risk-free rate. BETA is the average firm/month beta estimated as in Fama and French (1992). ME is market equity updated in June of every year (t) and used from July in that year (t) until June the following year (t+1). BE/ME is book-to-market ratio estimated as in Fama and French (1992); the fiscal year end book value (t-1) is divided by the calendar year end (t-1) market equity. A firm's BE/ME is assigned to the firm from July in the year after it was estimated (t), until June (t+1) of the following year in order to ensure that it was available to investors at the time it is used to predict returns. RET(-2,-7) is the average stock return from month t-7 to t-2. TURN is the average turnover over the last 36 months (eg. T-37 to T-1). CVTURN is the coefficient of variation of the turnover over the last 36 months (eg. T-37 to T-1). E(IVOL) is the idiosyncratic volatility estimate generated by EGARCH(p,q) models where $1<p<3$ and $1<q<3$. IVOL is the realized idiosyncratic volatility. All variables except for RET, RET-RF and BETA have been winsorized at the .5% level. Observations with returns exceeding 300% have been excluded, and variables with skewness greater than 3 are presented as their natural logarithm.

Table 1.2: Cross-sectional correlations

	RET(%)	RET -RF(%)	BETA	LN(ME)	LN (B/M)	RET (-2,-7)	LN (TURN)	LN (CVTURN)	E(IVOL)	IVOL	E(IVOL)× LN(ME)
RET-RF(%)	1.00**										
BETA	-0.01	0.01									
LN(ME)	-0.01	-0.01	-0.28**								
LN(B/M)	0.03**	0.03**	-0.07**	-0.25**							
RET(-2,-7)	0.02**	0.02**	-0.03**	0.04**	0.06**						
LN(TURN)	-0.02**	-0.02**	0.42**	0.06**	-0.11**	0.00					
LN(CVTURN)	0.00	0.00	0.21**	-0.57**	0.18**	0.05**	0.01				
E(IVOL)	0.11**	0.11**	0.37**	-0.37**	-0.07**	-0.03**	0.24**	0.33**			
IVOL	0.12**	0.12**	0.35**	-0.43**	-0.05**	-0.13**	0.19**	0.36**	0.54**		
E(IVOL)× LN(ME)	-0.02**	-0.02**	-0.01	0.72**	-0.33**	0.01**	0.24**	-0.32**	0.01	-0.21**	
IVOL×LN(ME)	-0.03**	-0.03**	-0.02**	0.65**	-0.26**	-0.03**	0.14**	-0.25**	-0.11**	-0.05**	0.75**

This table displays the time-series means of the cross-sectional Pearson correlation coefficients. In every month we obtain the contemporaneous Pearson correlation coefficients between each pair of variables using the cross-section of stocks. The reported correlation coefficients are the average over the sample period of 558 months. Correlation coefficients that are significant at 1%, given their time-series standard error, are marked with two asterisks **. The coefficients relate to a sample of stocks traded NYSE, Amex, and Nasdaq between July 1963 and December 2009.

Table 1.3: Cross-sectional regressions with Fama-MacBeth T-stats

Model	1	2	3	4	5	6	7	8
BETA	0.10 (0.41)	0.23 (1.28)						
LN(ME)	-0.13 (-3.15)	-0.18 (-4.54)	0.34 (9.78)	0.22 (5.98)	0.93 (20.31)	0.71 (15.16)	0.99 (20.87)	0.68 (15.37)
LN(BE/ME)	0.21 (3.88)	0.16 (3.24)	0.63 (11.69)	0.51 (10.55)	0.53 (10.27)	0.44 (9.40)	0.59 (10.98)	0.42 (8.80)
E(IVOL)			0.27 (15.97)	0.29 (18.57)	0.34 (19.25)	0.34 (19.73)		
IVOL							0.39 (15.66)	0.40 (15.98)
RET(-7,-2)		0.03 (2.40)		0.05 (4.60)		0.05 (4.54)		0.10 (8.72)
LN(TURN)		-0.15 (-2.60)		-0.58 (-9.12)		-0.41 (-6.77)		-0.57 (-9.32)
LN(CVTURN)		-0.40 (-6.18)		-0.90 (-12.24)		-0.72 (-9.78)		-0.76 (-9.20)
E(IVOL)×LN(ME)					-0.06 (-14.20)	-0.05 (-11.77)		
IVOL×LN(ME)							-0.05 (-9.29)	-0.03 (-5.27)
R^2	3.69%	5.62%	5.70%	7.88%	6.75%	8.63%	9.52%	11.40%

We estimate cross-sectional regressions for every month from July of 1963 to December of 2009, a total of 558 months. The dependent variable in the monthly regressions is the holding period return on all firms in our sample that have information on all explanatory variables in that month. The explanatory variables are as defined in Table 1.1. All the explanatory variables (except for IVOL) are predetermined at time t, and are used to explain the variation in returns at time t. For every month, we estimate a model that is nested in the following cross- sectional regression:

$$R_{i,t} = a_t + \beta_1 BETA_{i,t} + \beta_2 LN(ME)_{i,t} + \beta_3 LN(BE/ME)_{i,t} + \beta_4 E(IVOL)_{i,t} + \beta_5 RET(-7,-2)_{i,t} + \beta_6 LN(TURN)_{i,t} + \beta_7 LN(CVTURN)_{i,t} + \beta_8 E(IVOL) * LN(ME)_{i,t} + \varepsilon_{i,t}$$

This table summarizes the means of the coefficients from these cross-sectional regressions. The standards errors are calculated as in Fama and MacBeth (1973), but corrected for serial correlation as in Newey and West (1987). The corresponding test-statistics are presented in parentheses below each average coefficient.

Table 1.4: Isolating the net effect of idiosyncratic volatility on returns

Panel A				
	NYSE Quartile			
	1st	**2nd**	**3rd**	**4th**
Mean LN(ME)	0.79	3.05	4.12	5.73

Panel B				
	NYSE Quartile			
Net effect of E(IVOL)	**1st**	**2nd**	**3rd**	**4th**
Net E(IVOL) Effect for Model 5 (table 1.3)[a]	0.28	0.14	0.07	-0.03
Net E(IVOL) Effect for Model 6 (table 1.3)[a]	0.30	0.18	0.13	0.04

[a] The net effect in these tables might deviate slightly from calculations using the interaction terms in table 1.3 due to rounding in the presentation of those numbers.

With the interaction term included, the coefficients on E(IVOL) in table 1.3 can no longer be interpreted as their stand-alone figures, but need to be interpreted at a certain level of the variable with which it is interacted (LN(ME)). This gives the "net effect" of E(IVOL) given a certain level of size:

$$\frac{\partial R_{i,t}}{\partial E(IVOL)} = \widehat{\beta_{4_t}} + \widehat{\beta_{8_t}}\overline{LN(ME)_{\iota,t}}$$

Panel A shows the average size ($\overline{LN(ME)}$) for firms in 4 size quartiles based on NYSE breakpoints. Panel B illustrates the relation between idiosyncratic volatility and return for firms in each of the four size quartiles.

Table 1.5: Zero-cost portfolios based on size and idiosyncratic volatility

PANEL A: Value-weighted portfolio returns

		Size Quintile				
		1	2	3	4	5
	1	0.01	0.39	0.52	0.63	0.80
	2	-0.37	0.32	0.55	0.81	0.69
E(IVOL)	3	-0.11	0.31	0.59	0.76	0.56
	4	0.48	0.55	0.75	0.91	0.57
	5	5.95	3.22	1.80	1.37	0.52
	Diff	5.95	2.83	1.29	0.73	-0.28

PANEL B: Value-weighted portfolio alphas

		Size Quintile				
		1	2	3	4	5
	1	-1.15 (-13.65)	-0.56 (-7.94)	-0.30 (-4.74)	-0.16 (-3.04)	-0.04 (-0.73)
	2	-1.41 (-13.70)	-0.70 (-8.63)	-0.32 (-4.89)	-0.08 (-1.51)	-0.01 (-0.17)
E(IVOL)	3	-1.03 (-7.43)	-0.66 (-7.72)	-0.37 (-0.37)	-0.18 (-3.03)	0.03 (0.67)
	4	-0.28 (-1.42)	-0.35 (-3.15)	-0.24 (-3.14)	-0.03 (-0.42)	0.04 (0.65)
	5	6.11 (15.60)	2.16 (8.35)	0.72 (4.17)	0.23 (1.66)	0.03 (0.26)

Each year in June, all sample stocks are assigned into 5 portfolios based on their market equity (with each portfolio having the same number of firms). Each month, the stocks in each of the size portfolios are sorted on expected idiosyncratic volatility and assigned to 5 portfolios. This double sort results in 25 portfolios, each with a time-series of monthly returns between July of 1963 and December of 2009, for a total of 558 months. Panel A presents the value-weighted average returns for the 25 portfolios. The monthly value-weighted portfolio returns for each of these 25 portfolios are then regressed on the Fama and French (1993) three factors (Rm, SMB, and HML).

$$R_{i,t}^{Size-E(IVOL)} = \alpha + \beta_i Rm_t + s_i SMB_t + h_i HML_t + \varepsilon_{i,t},$$

where i=1 to 25 Size-E(IVOL) portfolios and t=1 to 558 months. Panel B presents the Jensen (1968) alphas from those regressions, as well as their test statistic (in parentheses).

Table 1.6: Portfolio returns with trading costs considered

Panel A: Value-weighted portfolio returns

		Size Quartile			
		1	2	3	4
E(IVOL)	1	-0.20	0.15	0.47	1.02
	2	-0.44	0.42	0.45	1.19
	3	0.00	-0.47	0.42	1.19
	4	4.58	1.27	0.33	1.97
	Diff	4.79	1.13	-0.14	0.95

Panel B: Value-weighted portfolio returns considering trading costs

		Size Quartile			
		1	2	3	4
E(IVOL)	1	-5.59	-2.61	-1.39	0.24
	2	-7.38	-3.97	-2.28	0.32
	3	-7.52	-4.23	-2.43	0.25
	4	-0.09	-0.49	-0.31	0.60

Each year in June, all Nasdaq stocks with available data on bid-ask spread are assigned into four portfolios based on their market equity (with each portfolio having the same number of firms). Each month, the stocks in each of the size portfolios are sorted on expected idiosyncratic volatility and assigned to four portfolios. This double sort results in 16 portfolios, each with a time-series of monthly returns between Jan of 1973 and December of 2009, for a total of 444 months. Panel A presents the value-weighted average returns for the 16 portfolios. Panel B presents the average returns for the portfolios when one-way trading costs (based on half of the bid-ask spread) are considered.

Table 1.7: Summary statistics by size

	Quartile 1		Quartile 2		Quartile 3		Quartile 4	
Var	**Mean**	**Std**	**Mean**	**Std**	**Mean**	**Std**	**Mean**	**Std**
RET (%)	1.16	19.15	0.97	14.08	0.98	12.61	0.93	10.54
RETRF (%)	0.70	19.16	0.55	14.08	0.57	12.61	0.51	10.55
BETA	1.35	0.35	1.25	0.38	1.16	0.34	1.04	0.30
LN(ME)	0.79	1.27	3.05	0.78	4.12	0.85	5.73	1.22
LN(BEME)	-0.24	1.16	-0.56	0.99	-0.63	0.93	-0.75	0.87
RET(-2,-7)	0.93	7.62	1.27	6.19	1.24	5.39	1.18	4.53
LN(TURN)	1.32	1.05	1.65	1.13	1.67	1.19	1.56	1.21
LN(CVTURN)	4.33	0.42	4.04	0.39	3.90	0.37	3.67	0.37
E(IVOL)	13.63	9.58	10.12	6.64	8.87	5.35	7.55	4.08
IVOL	15.25	12.33	10.04	7.28	8.61	6.13	7.11	4.82
E(IVOL)×LN(ME)	9.76	22.03	31.66	24.19	37.25	25.56	43.77	26.66
IVOL×LN(ME)	11.96	23.80	31.03	25.52	35.90	28.49	41.06	31.03

This table presents descriptive statistics for a sample of NYSE, Amex, and Nasdaq stocks between July of 1963 and December of 2009 divided into size quartiles. Four size portfolios are formed based on NYSE breakpoints in June and the procedure rolls every year. The table presents summary statistics for each of the portfolios. Variable definitions are as in Table 1.1. All variables except for RET, RET-RF and BETA have been winsorized at 99.5% and .5%. Observations with returns exceeding 300% have been excluded.

Table 1.8: Cross-sectional Fama-MacBeth regressions by size quartiles

	Panel A				Panel B			
	Size quartile				Size quartile			
	1	**2**	**3**	**4**	**1**	**2**	**3**	**4**
LN(ME)	1.91 (4.95)	0.06 (0.23)	0.25 (1.09)	0.18 (1.92)	1.36 (5.17)	0.74 (2.54)	0.54 (2.53)	0.56 (6.26)
LN(BE/ME)	0.65 (8.73)	0.28 (4.58)	0.23 (3.68)	0.07 (1.09)	0.59 (7.36)	0.29 (4.90)	0.23 (3.91)	0.09 (1.43)
E(IVOL)	0.40 (11.73)	0.14 (1.51)	0.23 (1.98)	0.25 (2.91)				
IVOL					0.44 (10.85)	0.38 (4.03)	0.42 (3.28)	0.59 (6.05)
RET(-7,-2)	0.04 (2.87)	0.07 (4.68)	0.03 (1.67)	0.03 (1.58)	0.14 (5.48)	0.10 (6.34)	0.04 (2.53)	0.02 (1.16)
LN(TURN)	-0.71 (-9.71)	-0.28 (-3.37)	-0.14 (-2.22)	-0.08 (-1.08)	-0.92 (-10.85)	-0.53 (-7.40)	-0.23 (-3.96)	-0.03 (-0.47)
LN(CVTURN)	-0.92 (-8.35)	-0.71 (-6.55)	-0.39 (-4.05)	-0.28 (-2.91)	-0.93 (-6.84)	-0.81 (-7.62)	-0.44 (-5.35)	-0.16 (-1.85)
E(IVOL)×LN(ME)	-0.12 (-2.83)	0.01 (0.13)	-0.05 (-1.32)	-0.04 (-2.79)				
IVOL×LN(ME)					-0.01 (-0.25)	-0.03 (-0.76)	-0.07 (-2.23)	-0.10 (-5.83)
R^2	10.12%	9.99%	10.37%	12.63%	13.37%	13.31%	13.50%	15.36%
Net Effect Of E(IVOL)[a]	0.31	0.15	0.04	0.00	0.43	0.29	0.12	0.00

[a] The net effect in these tables might deviate slightly from calculations using the interaction terms and mean size for each quartile due to rounding in the presentation of those numbers.

Panel A presents the average coefficients for the main time-series regression model by size quartiles (based on NYSE stocks). In every month, regressions were run with monthly returns (RET) as the dependent variable and LN(ME), LN(BE/ME), E(IVOL), RET(-2,-7), LN(TURN), LN(CVTURN), and E(IVOL)×LN(ME). All the explanatory variables (except for IVOL) are predetermined at time t, and are used to explain the variation in returns at time t. This table summarizes the means of the coefficients from these cross-sectional regressions by size quartiles. The standards errors are calculated as in Fama and MacBeth (1973), but corrected for serial correlation as in Newey and West (1987). The corresponding test-statistics are presented in parentheses below each average coefficient. The variables relate to a sample of NYSE, Amex, and Nasdaq stocks between July of 1963 and December of 2009. Panel B presents the same regression, but using the realized measure of idiosyncratic volatility (IVOL) instead of the expected idiosyncratic volatility measure (E(IVOL)).

Table 1.9: Alternative measures of trading costs

	Liquidity quartile			
	1	**2**	**3**	**4**
LN(ME)	0.26 (7.87)	0.30 (7.24)	0.19 (3.68)	-0.35 (-0.71)
LN(BE/ME)	0.44 (7.91)	0.53 (9.00)	0.38 (5.03)	-1.10 (-0.89)
E(IVOL)	0.35 (13.80)	0.33 (7.82)	0.27 (2.64)	1.04 (0.74)
RET(-7,-2)	0.05 (3.55)	0.04 (3.08)	0.07 (4.15)	-0.26 (-1.09)
LN(TURN)	-0.17 (-1.12)	-0.73 (-1.73)	-1.29 (-1.70)	10.55 (1.17)
LN(CVTURN)	-0.66 (-7.22)	-0.73 (-8.81)	-0.86 (-7.66)	-0.859 (-0.64)
E(IVOL)×LN(TURN)	0.02 (0.93)	0.05 (0.88)	0.11 (0.70)	-1.28 (-1.10)
R^2	10.01%	9.65%	10.30%	10.14%

This table presents the average coefficients for the main time-series regression model (Model 6, table 1.3) by liquidity quartiles. Liquidity quartiles are based on the liquidity measure (TURN). In every month, for each liquidity quartile, we run regressions with monthly returns (RET) as the dependent variable and LN(ME), LN(BE/ME), E(IVOL), RET(-2,-7), LN(TURN), LN(CVTURN) and E(IVOL)×LN(TURN). All the explanatory variables (except for IVOL) are predetermined at time t, and are used to explain the variation in returns at time t. This table summarizes the means of the coefficients from these cross-sectional regressions. The standards errors are calculated as in Fama and MacBeth (1973), but corrected for serial correlation as in Newey and West (1987). The corresponding test-statistics are presented in parentheses below each average coefficient. The variables relate to a sample of NYSE, Amex, and Nasdaq stocks between July of 1963 and December of 2009.

Table 1.10: Variations over time

	Time period		
	1963-1978	1979-1994	1995-2009
LN(ME)	0.33 (4.50)	0.90 (12.32)	0.90 (13.75)
LN(BE/ME)	0.36 (3.90)	0.54 (6.81)	0.42 (6.03)
E(IVOL)	0.32 (10.12)	0.31 (13.96)	0.39 (11.52)
RET(-7,-2)	0.05 (2.26)	0.06 (3.87)	0.05 (2.14)
LN(TURN)	-0.31 (-2.68)	-0.39 (-4.59)	-0.55 (-4.83)
LN(CVTURN)	-0.63 (-7.73)	-0.41 (-3.78)	-1.13 (-7.14)
E(IVOL)×LN(ME)	-0.04 (-4.00)	-0.06 (-9.28)	-0.06 (-10.36)
R^2	11.45%	6.46%	8.04%

This table presents the average coefficients for the main time-series regression models by time periods. For every month, we run regressions with monthly returns (RET) as the dependent variable and LN(ME), LN(BE/ME), E(IVOL), RET(-2,-7), LN(TURNOVER), LN(CVTURNOVER) and E(IVOL)×LN(ME). All the explanatory variables (except for IVOL) are predetermined at time t, and are used to explain the variation in returns at time t. This table summarizes the means of the coefficients from these cross-sectional regressions over three different time periods. The standards errors are calculated as in Fama and MacBeth (1973), but corrected for serial correlation as in Newey and West (1987). The corresponding test-statistics are presented in parentheses below each average coefficient. The variables relate to a sample of NYSE, Amex, and Nasdaq stocks between July of 1963 and December of 2009.

ARE THE U.S. EQUITY MARKETS DOMESTICALLY INTEGRATED?

1. INTRODUCTION

> *Both the NYSE and Nasdaq maintain that companies should regularly review their listing decision. "To the extent that Corporate America chooses not to do that, it leads to an unhealthy situation," says Robert Greifeld, president and chief executive officer of Nasdaq. […] A spokesman for the NYSE stated that every company "has that duty to the shareholders to make sure they're listed on the right marketplace" and should evaluate it "in a way that benefits their shareholders the best." (CEO.com (2006))*

Most papers in empirical finance implicitly or explicitly assume the same price of risk, for each priced systematic risk factor, across all risky assets within a given domestic market. In doing so, they rely on the assumption that markets are domestically integrated and, as such, the aggregate risk aversion of all investors determine the prices of risk. This is true in frictionless markets where investors have complete information, homogenous beliefs, and hold the mean-variance efficient combination of the market portfolio and a risk-free asset. However, investors might not hold the market portfolio because of exogenous reasons.[6] In fact, several recent papers have provided evidence that US investors do not, holding instead vastly undiversified portfolios (Barber and Odean (2000)).

There are two main implications to the above. First, if one group of investors does not hold the market portfolio, then the remaining set of investors will also not be able to

[6] Exogenous barriers to holding the market portfolio include incomplete information, limits on short sales, imperfect divisibility of securities, lack of liquidity, transactions costs, and taxes (e.g., Merton (1987)).

hold the market portfolio and will rationally expect to be compensated for bearing idiosyncratic risk (e.g., Malkiel and Xu (2002)). Therefore, idiosyncratic risk will be priced in expected returns. Second, the price of risk need not be the same across all assets (e.g., Lintner (1971), Rubinstein (1973)), in which case domestic markets are not integrated. Stated differently, if two identical firms are held by different subset of investors, they could have different required rates of return, which would imply that markets are segmented. Related to the first, several recent papers find that idiosyncratic risk is priced in U.S. equity returns even though the evidence is mixed (see, e.g., Lehmann (1990), Malkiel and Xu (2002), Ang, Hodrick, Xing, and Zhang (2006), Spiegel and Wang (2007), Fu (2009), and Chua, Goh, and Zhang (2010)).

In contrast, there is relatively little empirical work on the issue of whether the price of systematic risk is the same for all domestic assets and, as such, whether the major US exchanges are integrated. Given that priced idiosyncratic risk is inconsistent with domestic integration (see, e.g. King, Sentana, and Wadhwani (1994)), the mixed evidence on whether idiosyncratic risk is priced demands that we take a closer look at the issue of domestic integration. That is, if investors face obstacles to diversification, then "different securities are held by different subsets of investors" with the implication that the "price of risk […] varies inversely with the summation of the risk-tolerances of the investors who have [the stock] in their portfolios" (Lintner (1971)). This holds in the domestic market as strongly as it does in the international setting.

The purpose of this paper is threefold. First, we examine if U.S. equity markets are domestically integrated over the period 1985 to 2009 by examining if commonly used systematic risk factors have different risk prices in the different domestic markets.

Although prior research finds that US markets are domestically segmented[7], since these papers have been written several systematic risk factors have emerged as important in explaining the cross-section of returns. If risk prices depend on the factors included in the asset pricing model and less so on the particular test assets (King, Sentana, and Wadhwani (1994)), then given that any asset pricing test of integration is a joint test of integration and the correct specification of the model, omitting these factors could bias our results in a manner that is not obvious a priori. For instance, momentum and liquidity are currently accepted factors in the asset pricing literature so including these factors in integration tests is extremely important. In particular, if firms that list on US exchanges have different characteristics, then it is likely that these firms differ in liquidity and sensitivity to momentum, which has been found in prior papers (Sagi and Seasholes (2006)). To the extent that these omitted factors are correlated with the included factors, risk factor prices would be distorted.

Second, we investigate whether and how domestic integration has changed over time. This is important for several reasons and theory is not clear on the direction of the change in integration. As investors become more sophisticated, and markets become increasingly liquid, it could be argued that integration should increase. However, as stock markets become more competitive, they are likely to differentiate and attract different subsets of firms that are held by investors with different preferences (see the opening quote). By doing so, segmentation could have increased over time. Moreover, evidence in Campbell, Lettau, Malkiel, and Xu (2001) indicates that idiosyncratic risk has been increasing over time, suggesting that markets might be becoming more segmented.

---

[7] See Chen and Knez (1995), and Naranjo and Propopapadakis (1997). More recently Flood and Rose (2005) test for domestic integration but Parsley and Schlag (2007) demonstrate that their results do not hold.

Finally, and perhaps most importantly, if markets are segmented, the implications of such segmentation would depend on why it arises and, therefore, we address this. A priori, we hypothesize that segmented markets could depend on three things. One, markets could be segmented due to different investor groups preferring to hold assets with particular characteristics and the different exchanges generally list firms with specific characteristics. In turn, each investor group demands a different price of risk for systematic risks. We term this the investor-segmentation hypothesis. Several papers have found results that are consistent with this hypothesis (see section 5). Two, markets could appear segmented due to arbitrage constraints that would prevent rational diversified investors from taking advantage of the apparent mispricing. If so, rational investors could not profitably exploit the mispricing and the results would have few, if any, implications. For example, Bassin (1998) find that the abnormal returns that firms document when changing their listing venues cannot be exploited by arbitrage. If those abnormal returns are due to different prices of risk in the markets, then this suggests that investors are unable to exploit such opportunities due to arbitrage constraints. Three, the markets could appear to be segmented due to the bad-model problem. Since any integration test using a factor model is subject to the joint-hypothesis problem, these tests are as well.

Using an expanded time frame (1985-2009), relative to previous work, we find strong evidence of market segmentation between the NYSE, Amex, and Nasdaq. We find that the price of several common risk factors differ between the markets, some having more than twice as high a price in one market as in the other markets. Secondly, we find that integration has changed over time, but not in a clear linear fashion. More importantly, we find that segmentation is present in all our sub periods and, hence, is not driven by any specific time period.

Finally, we find evidence consistent with our investor-segmentation hypothesis. Given that certain characteristics (size and B/M) are likely to be used by different investor groups to screen their preferred stocks, we compare the price of risk for firms that are at different ends of the continuum of these characteristics (e.g., small versus large firms) and find that they have significantly different (statistically and economically) prices of risk. We also find that these results are not driven by arbitrage constraints. In addition to excluding firms with a stock price of less than $5 (throughout the paper), we also exclude firms in the bottom 10% of market value, and find that the segmentation results remain. In another test, we exclude the firms with the highest 10% of idiosyncratic volatility in the previous month (since these firms arguably would be the most expensive to arbitrage). Even so, we find that the segmentation results are robust. Since the prior tests are subject to the dual-hypothesis problem, we also test for segmentation using the model-free method suggested by Bekaert, Harvey, Lundblad, and Siegel (2011) which is based on industry-level earnings yield differentials across the markets. Consistent with the model-specific tests, the results using this approach also indicate that the markets are segmented. Hence, our results point to investor segmentation as being the driver for the observed market segmentation.

Our work is not the first to examine the question of domestic segmentation of the US markets. The current paper contributes to, but is different from, others in this line of research in several ways. Chen and Knez (1995) examine if the mean squared distance between stochastic discount factors for firms on the NYSE and Nasdaq is zero (strong-form integration), or some minimum amount (weak-form integration). They reject strong-form integration, but fail to reject weak-form integration. Naranjo and Propopapadakis (1997) compare NYSE, Amex, and Nasdaq stocks and reject the hypothesis that the NYSE, Amex, and Nasdaq are integrated. Both Chen and Knez (1995) and Naranjo and Propopapadakis (1997) conclude that the domestic equity markets are segmented but are silent on the cause of segmentation. Our work builds on these previous papers not only because our more recent sample period reflects the fact that market integration might have evolved over time in a direction that is not known a priori, but also because we test and eliminate the most likely causes of segmentation, and then present a potential reason for our findings.

Recently, Goyal, Perignon, and Villa (2008) find that NYSE/Amex and Nasdaq have three priced factors each, but that only two of those factors are common across both markets. Our work addresses the complementary question of whether these common factors are commonly priced across markets, because in markets where the priced risks are different the price of risk is unlikely to be the same (Bekaert and Harvey (1995)). Ang, Shtauber, and Tetlock (2011) compare return premiums for stocks on listed markets compared to the over-the-counter market and find differences in the factors that are priced in the markets, as well as differences in the premiums of the factors that are priced. Our work differs from theirs in that while they focus on comparing the systematic risk price of OTC stocks with listed stocks, we compare the price differences for stocks in different listed stock markets where, a priori, the factors that might cause segmentation are less obvious.

Our work has implications for studies that attribute the significant abnormal announcement-day return for firms that list on a larger/national exchange to the investor-recognition and superior-liquidity hypothesis (see, e.g., Kadlec and McConnell (1994)). If different clienteles display a preference for firms listed on the different exchanges, thus leading to differences in the price of risk, then firm value is expected to change even without an expected change in future cash flows.

Furthermore, our work contributes to the literature on style investing (see e.g., Barberis and Shleifer (2003)), which argues that commonly accepted risk “factors” are likely priced due to investors allocating their assets based on the characteristics underlying these “factors”. Furthermore, this strand of research finds a higher co-movement in prices among assets in a given style category. We expand the style and co-movement literature by investigating whether such style investment can lead to different risk prices across assets and/or markets.

Likewise, our work is relevant to the research that finds greater co-movement between stock returns of firms listed on the same stock exchange (see e.g., Barberis, Shleifer, and Wurgler (2005) and Kaul, Mehrotra, and Stefanescu (2006)). The co-movement literature finds that changes in fundamentals cannot explain the co-movement, suggesting instead ‘‘friction-based’’ and ‘‘sentiment-based’’ explanations. However, the literature does not make a distinction between whether a change in the quantity of risk (beta), the risk factors, or different prices of risk are driving the co-movement. If the same systematic risks have different risk prices on the different exchanges, then we would expect co-movement even without a change in the quantity of risk.

The paper is organized as follows. The second section introduces the data and presents summary statistics. The third section investigates if systematic risk is priced differently in the domestic equity markets. The fourth section investigates time trends in integration. The fifth section explores reasons for segmented domestic markets, and the sixth section concludes.

2. DATA AND SUMMARY STATISTICS

To conduct the various tests, three sets of data are required. The first set contains returns on the test assets. The second set contains the systematic risk factors that have been shown to impact the cross-section of returns. The third set contains the firm characteristics that have been shown to impact the cross-section of stock returns. The data are obtained from three sources. Daily and monthly individual security data for all firms traded on NYSE, Amex, and Nasdaq from 1973 to 2009 are obtained from the Center for Research in Security Prices (CRSP). We obtain data on stock returns (RET), prices (P), shares outstanding (SHROUT), and volume (VOL). We also obtain value-weighted index returns from CRSP. Accounting data or, more specifically, book values, are obtained from Compustat's annual fundamentals file. Due to the limited availability of volume data for stocks traded on the Nasdaq before 1983, together with the fact that we need monthly volume data for at least two years in order to estimate one of our explanatory variables, we limit the final sample to the period from 1985 to 2009. We exclude firms with a price below $5. This eliminates approximately 26% of firms (or, on average, 1031 firms per month), who make up approximately 0.7255% of total market value[8]. Hence, the screen works to eliminate small, illiquid firms that are likely to have significant arbitrage constraints from the sample.

A. Test asset

The test assets used in this paper are individual stock returns for all firms traded on the three major US exchanges. Monthly holding period returns are obtained through CRSP. Holding period returns (RET) include capital gains as well as dividend yields.

[8] These statistics are not reported in the tables of the paper

B. Systematic risk factors

To estimate the sensitivity to the systematic risk factors for each firm, we follow the methodology of Ang, Shtauber, and Tetlock (2011) and regress individual firm returns from month *t-24* to *t-1* on the proxies for the factors. To estimate the firm sensitivities (betas) to the Fama and French (1993) three-factor model, we regress monthly firm returns on the CRSP value weighted index, the SMB, and the HML factor[9]. We roll the regressions monthly. To be eligible for estimation, we require that a firm has returns in all the months of the estimation period (e.g., 24 months of data).

Since Fama and French developed their factors in 1993, much research has been focused on explaining cross-sectional returns. Liquidity and momentum are two of the most frequently used variables that have been shown to impact cross-sectional returns. Since Amihud and Mendelson (1986) introduced liquidity as a factor, several papers have shown that it is important in the cross-section of returns (see, e.g., Brennan and Subrahmanyam 1996; Datar, Naik, and Radcliffe 1998; Chorida, Subrahmanyam and Anshuman 2001; Amihud 2002; Pastor and Stambaugh 2003). Similarly, Jagadeesh and Titman (1993) show that a momentum effect is present in stock price movements and Jagadeesh and Titman (2001), among others, find that the effect has remained significant long after it was discovered. Therefore, we augment the Fama and French (1993) three-factor model with proxies for these factors.

[9] The SMB and HML factors are created as in Fama and French (1993). A detailed description on how to form those factors is available from Ken French's website at Dartmouth. The correlation between the factors available from Ken French's website and the ones we create exceed .97 for all factors.

In order to estimate the sensitivity to Momentum, we create the Carhart (1997) momentum factor, as described by Ken French[10]. Every month we sort firms, independently, on size and the prior year's return (excluding the last month). We form two portfolios on size (using NYSE median) and three on the prior year's return (using the 30th and 70th percentile on NYSE stocks). The momentum factor, UMD is then calculated as the mean return on the two portfolios with the highest prior yearly return, minus the mean return on portfolios with the lowest prior yearly return. As a proxy for liquidity, we develop a factor based on the Amihud (2002) illiquidity measure. Every month we sort firms, based on the prior month's illiquidity (see equation 2 below). The illiquidity factor is then calculated as the equal-weighted average return of firms with the highest 30 percent illiquidity lagged one month minus the equal-weight average of firms with the lowest 30 percent illiquidity return lagged one month. Both the momentum factors, as well as the liquidity factor, are updated monthly.

In order to estimate firms' sensitivity to these factors, we follow the methodology of Ang, Shtauber, and Tetlock (2011) and regress firm returns on the Market, SMB, and HML in addition to each of these factors, one at a time. Hence, a firm's sensitivity with respect to the momentum factor is calculated by regressing returns on the three Fama and French (1993) factors in addition to the momentum factor. Similarly, the sensitivity to the liquidity factor is estimated by regressing returns on the Fama and French (1993) factors in addition to the liquidity factor. Explicitly, for every month t we estimate a model nested in the following time-series model for each firm i, using data from month t-24 until t-1, and record the betas associated with each firm/month for each factor:

[10] A detailed description on how to form the Momentum factors is available from Ken French's website at Dartmouth

$$\mathrm{R_{i,t}} = \alpha_i + \beta_i^{MKT}\mathrm{Mkt}_t + \beta_i^{SMB}\mathrm{SMB}_t + \beta_i^{HML}\mathrm{HML}_t + \beta_i^{UMD}\mathrm{UMD}_t + \beta_i^{LIQ}\mathrm{LIQ}_t + \varepsilon_i. \quad (1)$$

Fama and MacBeth (1973) point out that betas estimated in this fashion are likely to be estimated with error, since the high beta estimations are likely to be overestimated, and the low beta estimates are likely to be underestimated. We note that this would bias against finding both that the factor premia are significant (since additional noise is introduced and our standard errors are likely to be inflated) and that the markets are segmented (since relatively small factor premia are less likely to be statistically significantly different).

C. Firm characteristics

Fama and French (1992) show in their seminal paper that size and book-to-market ratio are able to explain the cross-section of stock returns. Even though Fama and French (1993), along with several other papers, argue that the characteristics proxy for underlying systematic risk factors, other authors (Daniel and Titman (1997)) argue that the firm characteristics themselves often do better in predicting stock returns than the underlying factors. Acknowledging that it might be the characteristics and not the underlying factors that are priced, we measure the characteristics of interest for use in cross-sectional regressions as a complement to the sensitivity to the risk factors measured above.

We estimate size and book-to-market ratio as in Fama and French (1992) for each of our securities. Size and book-to-market are calculated in June every year, and used from July in that year until June in the following year. Size (ME) is measured by the market value of equity (shares outstanding multiplied by price) in June of year *t*. The calculated size is then assigned to a firm for the following twelve months, starting in July of year *t* until June of year *t+1*. BE/ME is the book-to-market ratio, calculated using December (*t-1*) market value of equity and the book value of equity for the fiscal year that ended in year *t-1*. BE/ME is, like size, updated in July. The time lag used is to ensure that the information was available to investors at the time it is used in the analysis. We take the natural logarithm of both the market value (ME) and the book-to-market value (BE/ME).

We calculate illliquidity as in Amihud (2002), but on a monthly basis. Every day in a month, we calculate the ratio of the daily absolute returns to the dollar trading volume. The illliquidity measure used for month *t* is then the average ratio for the days in the previous month (*t-1*).

$$\mathrm{ILLIQ}_{i,t} = {}^{1}\!/_{D_{i,t-1}} \sum_{d=1}^{D_{i,t-1}} \frac{|R_{i,t-1,d}|}{\mathrm{volume}_{i,t-1,d} \times \mathrm{share\ price}_{i,t-1,d}}, \tag{2}$$

where $D_{i,t-1}$ is the number of days for which data are available for stock *i* in month *t-1* and $R_{i,t-1,d}$ is the return on stock *i* on day *d* in month *t-1*. $\mathrm{Volume}_{i,t-1,d}$ and $\mathrm{share\ price}_{i,t-1,d}$ are the volume and share price, respectively, for stock *i* on day *d* in month *t-1*. We require at least 15 trading days for a firm in month *t-1* with reported values of volume and share price in order to calculate the Amihud illiquidity measure for time *t*.

The measure of momentum, RET (-12, -2) is calculated as the compounded holding period return between month *t-12* to *t-2*. We exclude the last month's (*t-1*) return to avoid returns merely due to bid-ask bounce. As can be seen in the variable definitions, all systematic risk factors are estimated using data up until *t-1* and are therefore predetermined at time *t*.

Due to the need of volume data to calculate illiquidity, and the fact that Nasdaq does not consistently report volume until the end of 1982, we limit our sample to between January of 1985 and December 2009 for our cross-sectional regressions, so that we have data on all our explanatory variables.

This results in a total of 1,247,583 firm-month observations over the 300 month sample window. All explanatory variables are winsorized monthly at the 5% and 95% level in order to avoid giving extreme observations a heavy weight.

D. Summary statistics

Summary statistics for firms listed on the NYSE, Amex, and Nasdaq are presented in Table 2.1. The average (arithmetic) returns is quite different across the markets. NYSE has average monthly returns of 1.41%, while the average return is 2.23% for Amex and 2.38% for Nasdaq. As we know, the average size (market capitalization) of firms on the NYSE is a lot larger than those on the other two exchanges. The same holds true for the share price, which is approximately three times higher on the NYSE, than on Amex and Nasdaq. The average share price for NYSE stocks is $69.17 while the corresponding price is $19.94 on Amex and 19.66 on Nasdaq. Firms on the Nasdaq appear to be more growth oriented (lower BE/ME) than the firms trading on the NYSE and Amex. Our measures of liquidity show that NYSE stocks are much more liquid (less illiquid) than Nasdaq and Amex stocks.

E. Correlation between the variables for each market

Every month we estimate the Pearson correlation coefficients for the variables in each market. We report the time series mean of those correlation coefficients, as well as their significance based on their time-series standard errors. Most correlation coefficients are very similar across the three markets. The mean betas related to the market, SMB, HML, and liquidity are positively correlated to returns in all three markets (even though the correlation is not always statistically significant). The mean momentum (UMD) beta is negatively correlated to returns in all markets. There does not seem to be any concern regarding multi-collinearity between our variables of interest.

3. MARKET INTEGRATION TESTS

In this section we test which factors are priced in each of the markets (sub-section A) and whether the price of risk is different for the different markets (sub-section B). The question of whether or not U.S. equity markets are domestically integrated is important for several reasons. First, if the reward for each unit of systematic risks is not the same in each market, then it would be possible to increase the *total* reward without altering total risk by shifting investments to markets with higher reward for the given risk. Hence, it would present obvious arbitrage opportunities for investors. Second, investment and financing decisions would not be independent as the cost of capital would depend on the market in which projects are financed. Hence, markets would be allocationally inefficient, which would impose a social cost, and lead to a suboptimal growth rate. Third, these tests have implications for, and the results cannot be inferred from, tests of international integration[11] which implicitly assume that markets are domestically integrated. This is because it is possible that two national markets appear internationally integrated because particular subsectors (such as that containing the largest firms, which are typically held by a common set of global investors) are internationally integrated but within each domestic market these subsectors are not fully integrated with the rest of the market that is typically held by domestic investors.[12]

While the NYSE, Amex and Nasdaq provide the *same service*, there are several important differences between the exchanges, that could influence *which firms chose to list in each market and, in turn, create a clientele effect whereby different investor groups*

[11] Bekaert and Harvey (1995), Bekaert, Hodrick, Zhang (2009), and others.

[12] An example of this is that investable stocks in emerging markets appear fairly highly integrated internationally, but are not integrated with non-investable stocks (see, e.g., Carrieri, Chaieb, and Errunza (2010)).

have distinct preferences for the firms trading on one exchange and not the others. First, listing requirements and costs are very different between the markets, suggesting that different firms might chose to list on the two markets. Technology and growth companies have historically been more prominent on the Nasdaq, while large blue chip firms have been predominantly listing on the NYSE. Corwin and Harris (2001) analyze the factors that affect firms' listing decisions and find that firms are more likely to list on the same exchange as their industry peers. Further, they find that smaller, riskier firms are more likely to list on the Nasdaq, consistent with avoidance of delisting costs. Baruch and Saar (2009) find that listing decisions influence the liquidity of assets. More specifically, they find that a firm that chose to list on a market where similar stocks trade has greater liquidity.

Second, Fama and French (2004) find that after 1972 approximately 90% of new listings chose to list on Nasdaq. When taken together with the finding that new listings are dominated by younger firms (Fama and French (2000)), these differences are likely to lead to the situation where firms of broadly similar characteristics list on a particular exchange and the firms listing on the different exchanges have different characteristics which could attract a specific clientele of investors.[13]

A. Testing which risk factors are priced in the domestic markets

In Table 2.3, we present the results from Fama and MacBeth (1973) type regressions. In every month *t*, we estimate a model that is nested in the following cross-sectional regression:

[13] Moreover, as seen from the opening quote, exchange executives themselves think that there is an optimum match between the listing exchange and the firm. As such, they might be explicitly or implicitly contributing to this clientele effect.

$$\mathrm{R_{i,t}} = \lambda_0 + \lambda_1\beta_{i,t}^{Mkt} + \lambda_2\beta_{i,t}^{SMB} + \lambda_3\beta_{i,t}^{HML} + \lambda_4\beta_{i,t}^{UMD} + \lambda_5\beta_{i,t}^{LIQ} + \lambda_6 LN(ME)_{i,t} + \lambda_7 LN\left(\frac{BE}{ME}\right)_{i,t} + \lambda_8 RET(-12,-2)_{i,t} + \lambda_9 AMIILL_{i,t} + \epsilon_{\mathrm{i,t}}, \quad (3)$$

where $\mathrm{R_{i,t}}$ is the return in month *t* for stock *i* and the betas are the estimated firm sensitivities to the systematic risk factors in month *t* for each firm *i* from equation (1). The model is estimated separately for firms on the different exchanges. The price of risk, λ_k, for risk factor *k* will, therefore, be specific to a particular market. This allows us to determine if commonly used systematic risk factors are priced on the individual exchanges. Table 2.3 summarizes the means of the coefficients from these cross-sectional regressions over our estimation period between Jan of 1985 and Dec of 2009, a total of 300 months. The *t*-statistics, calculated as in Fama and MacBeth (1973)) but corrected for three lags of serial correlation as in Newey and West (1987), are presented underneath each coefficient.

Model 1 shows the results of the basic model containing the betas with respect to the market, SMB and HML factors. The results indicate that systematic risk, both with respect to the market and the size factors, are significantly priced in all three markets. Surprisingly, the systematic risk represented by HML only seems to be significantly priced on the NYSE. A further look shows that the magnitude of the price of market risk differs substantially across the markets. Several prior papers have found the relation between beta and returns to be flat in the cross-section of stock returns (Fama and French (1992, 1993)), which is in stark contrast to what we find here. While the premium for one unit of market beta risk is a modest .326% per month in the NYSE, it is approximately 50% higher (.473%) in the Nasdaq and twice as large (.667%) in Amex. With regards to the size factor, the premium also appears much larger for stocks traded on the Nasdaq and

Amex than for NYSE stocks. The size factor premium is .218% on the NYSE, and approximately 50% higher for stocks trading on the Amex (.317%) or Nasdaq (.308%).

The second model specification augments the three-factor model of Fama and French (1993) by adding the momentum and liquidity factors. The qualitative results from Model 1 hold after including these additional risk factors. In our sample, momentum is not significantly priced in any of the markets. Given that our sample excludes stock with a price below $5, and empirical research have shown that momentum seem to be most pronounced for small stocks (see Sagi and Seasholes (2006) and Bergbrant (2011)), this might explain why the risk represented by momentum does not appear significantly related to returns in our sample. In contrast, the premia for liquidity risk is significant in all markets. Similarly to the factor premium for market risk and SMB, the premium for liquidity is also much larger on the Amex and Nasdaq than on NYSE.

The finding that there are factors that are priced across all markets but other factors are priced in only a subset of markets is consistent with the results in Goyal, Perignon, and Villa (2008) and Ang, Shtauber, and Tetlock (2011). Overall, the above results indicate that there are economically large differences in the price of risks across exchanges, tentatively indicating that the markets are domestically segmented.

In the last two models, we explain the cross-section of returns using characteristics, as Daniel and Titman (1997) and others argue that characteristics do a better job in explaining the cross-section of stock returns than factor loadings. If investors hold only a certain subset of stocks, which ultimately leads to domestic segmentation, then investors might choose the subset of stocks in which to invest based on certain stock characteristics. If so, this could explain why characteristics have been shown to better price assets than factors formed on the basis of these characteristics (see e.g., Daniel and Titman (1997)). Following this explanation, characteristics such as size could appear priced due to the fact that the subset of investors that hold large stocks require different prices of risk than those investors who hold small stocks.

When using characteristics to explain the variation in stock returns, as opposed to factor loadings, the results are similar. Size appears important (statistically significant) in all specifications[14], while B/M is significant and positive only for NYSE firms in Model 3. Turning to Model 4 there are some interesting results. When adding momentum and liquidity to the model, B/M emerges as significantly negatively related to returns on the Amex (it remains positive and significantly related to returns on the NYSE). As in Model 2, momentum remains insignificantly related to return, and liquidity remains important in the extended model.

As these regressions reveal, there appear to be several differences between the markets with respect to the risk factors. The differences in the premia between the markets suggest that the prices of systematic risk may differ across the three markets, which would imply that the markets are not integrated. In the next section, we investigate whether the markets are, in fact, integrated.

[14] The negative sign on the coefficient on size in the characteristics regressions (larger firms' earnings lower returns) is consistent with the positive coefficient in the factor models on SMB, where firms more sensitive to the small-minus-large factor earn higher returns.

B. Are the U.S. equity markets integrated?

In order to investigate if the three domestic US markets are integrated we proceed by comparing the risk prices of commonly used risk factors across the markets (subsection i) as well as conduct time-series asset pricing tests (subsection ii).

i) Prices of risk across the markets

Since our aim is to investigate whether the prices of systematic risk differ across the markets, we estimate a model that allows for the price of risks to differ across the markets and allows us to test for statistical differences in these prices of risk. The model is a modification of that presented above (equation 3), as it includes a dummy variable defined as one for stocks that are listed on Amex and zero otherwise (*Amex*). It also includes a second similarly defined dummy variable for firms listed on Nasdaq (*Nasdaq*). The dummy variables are included in the model as standalone variables and interacted with the risk factors. As standalone variables they allow for different intercepts across the three markets, while the interaction terms allow us to determine if the prices of risk differ across the markets. Specifically, we estimate:

$$\mathrm{R_{i,t}} = \mathrm{a_t} + \mathrm{a_{at}}\,\mathrm{Amex} + \mathrm{a_{nt}}\mathrm{Nasdaq} + \sum_{k=1}^{5} \lambda_{\mathrm{k,t}} \beta_{i,t}^{k} + \sum_{k=1}^{5} \lambda_{\mathrm{ak,t}} \left(\beta_{i,t}^{k} \times \mathrm{Amex}\right) + \sum_{k=1}^{5} \lambda_{\mathrm{nk,t}} \left(\beta_{i,t}^{k} \times \mathrm{Nasdaq}\right) + \varepsilon_{\mathrm{i,t}}, \tag{4}$$

where $\lambda_{\mathrm{k,t}}$ is the price of risk for factor *k* in time *t* on the NYSE, and $\lambda_{\mathrm{xk,t}}$ is the incremental price of risk for factor *k*, in time *t*, on exchange *x* (where *x* = *a* represents Amex and *x* = *n* represents Nasdaq). As can be inferred from the dummy variables in equation (4), NYSE is the omitted market. Hence, the coefficient estimate on the interaction terms represents the incremental price of risk for stocks trading on Amex and

Nasdaq, relative to those trading on NYSE. The test statistic for these interaction terms indicates whether the price of risk on the Amex and Nasdaq, respectively, is statistically significantly different from that on the NYSE.

The results are reported in Table 2.4, where the coefficients on the various risk factors for Amex and Nasdaq stocks are the incremental price of risk, relative to the price of risk for NYSE stocks. Hence, for Amex or Nasdaq the total price of risk for a particular factor is the sum of the price of risk for NYSE stocks plus the incremental price of risk for Amex or Nasdaq. They indicate that several of the systematic factors have a different price of risk in the different markets. Focusing on Model 1, and comparing Amex and NYSE, the results indicate that the price of market risk is significantly higher on Amex than on NYSE. Further, the price of HML risk is significantly lower on both the Amex and Nasdaq than on NYSE. Adding the momentum and liquidity factor does little to change the relations just described. In addition, liquidity appears to be significantly higher priced on both the Amex and Nasdaq as compared to the NYSE. Momentum remains insignificant in all models.

Turning to the models using characteristics, they show that, if the characteristics are priced, the premia would differ between the NYSE and Nasdaq (and Amex) for both size, book-to-market, and liquidity. Hence, the segmentation results are even stronger using the characteristic models than they are using the factor loading models.

Strong-form integration requires the risk prices to be the same for the markets in every period. In our tests, we investigate whether the average risk prices have been different over our entire sample period. As such, these tests would be considered conservative, so the results strongly imply that the domestic markets are segmented.

These results are consistent with the evidence in Chen and Knez (1995) and Naranjo and Protopapadakis (1997)) over a sample period that ends nearly two decades prior to ours.

ii) Time-series asset pricing tests

We also test the hypothesis of segmentation using time series tests. In addition to the "global" factors created previously created using data from all the three exchanges, we also create exchange-specific factors (independently created for each exchange) using each exchanges individual breakpoints. We then form 25 portfolios (sorted on book-to-market and size) for each exchange and use those (3×25=) 75 portfolios as test portfolios[15]. If the markets are integrated, we expect that the global factors would explain the variation in the portfolios over time and that exchange-specific factors would be redundant and not add any explanatory power.

We then regress the monthly returns on each of the (75) portfolios on the global factors and record the R^2s. We then include the exchange-specific factors and investigate whether the exchange-specific factors significantly improve the fit of the model. We find that for 96% of the three factor models, the exchange-specific factors are important (72 out of the 75 portfolios).[16] This provides support to the previous results that the US markets are domestically segmented. Furthermore, adding the additional two factors, momentum and liquidity, does not change the results. The exchange-specific factors remain important in 96% (72 out of 75) of the portfolios when we estimate the five-factor models. This strongly implies that the domestic US markets are segmented.

[15] The results are very similar when using independently sorted portfolios.
[16] These tests are not reported for brevity, but are available upon request.

4. CHANGES IN INTEGRATION OVER TIME

Having shown that the domestic markets over the entire sample period, 1985 to 2009, are segmented naturally raises the question of how market integration has evolved over time. At first, it might seem natural that markets should be increasingly integrated, due to the fact that technology has made information more readily available over time. With advanced computer power, we now have the ability to research more stocks, diversify easier, and trade at a lower cost. At the same time, computerized trading allows investors to exploit arbitrage opportunities quickly. Furthermore, increasing similarities between the stock exchanges could have caused greater similarity between the characteristics of firms that chose to list on the three exchanges. That is, in the early 1970s most firms that listed on the Nasdaq did not have the option to list on the NYSE. Today, more firms have the opportunity to choose which exchange they want to list on.

However, the investment climate has also changed. Exchanges that were complementary a few decades ago, such as NYSE and Nasdaq, have grown into competitors. In an increasingly competitive industry, which has lately been characterized by several acquisitions and mergers, the exchanges are being forced to differentiate themselves. This differentiation could attract both specific subset of firms, and subset of investors. As Nasdaq has grown to become known for technology firms (often growth firms), NYSE has continued to focus on staple firms (often focused on value). Hence, it would be likely that firms that previously could not chose where to list, now make an active choice based on which market best fits their needs. To the degree that this is true, we would expect markets to become less integrated over time. Moreover, evidence in

Campbell, Lettau, Malkiel, and Xu (2001) indicates that idiosyncratic risk has been increasing over time, suggesting that markets might be becoming more segmented.

Furthermore, it is possible that markets change in their levels of integration over time without a clear direction. If markets are segmented because different groups of investors hold different securities, then market segmentation could increase if the risk aversion of the different investor groups changes at different times (or changes by different amounts at a particular time). For example, if individual investors hold Nasdaq stocks and institutions own NYSE stocks, then segmentation would increase in the event that individual investors become more risk averse relative to institutional investors. Hence, changing levels of segmentation over time, without a clear time trend would be consistent with investor- driven segmentation.

To test the time trends of segmentation, we split the sample into subperiods, the first spanning the years 1985 through 1992, the second covers 1993 to 2000, and the third time period covers the years 2001 until 2009.

To determine how integration changes (if it changes) over time we re-estimate the model with the interaction terms, equation (4), for Nasdaq and Amex, where NYSE is the omitted exchange, over the three subperiods. The results from these regressions are presented in Table 2.5.

There are three interesting things to note in Table 2.5. First, there does not seem to be a clear time-trend in integration (e.g. integration does not change linearly over time). Secondly, the risk factors that seems to imply segmentation changes over time and the magnitude of those differences changes as well. For example, the price of HML risk is significantly *higher* on Amex than on NYSE in the first sub-period, but significantly *lower* in the subsequent sub-periods. Turning to market risk, we find that its risk price is significantly *higher* on Nasdaq as compared to NYSE in the first and last sub-period, but appears to be *lower* in the middle sub-period. This is consistent with the investor-driven segmentation hypothesis described briefly above. Third, and perhaps most importantly, segmentation is present in all sub-periods and, hence, is not driven by any particular period.

To provide the reader with a picture of the time variation of domestic segmentation we present a plot of the risk prices from the base model (Figure 2) and their differences (Figure 3), where the latter depicts the individual systematic risk prices for Amex or Nasdaq stocks minus the same for NYSE. Figure 2[17] is informative as to the variation in the prices of risk over time, which is consistent with asset pricing models. For instance, it is noticeable that prices of risk tend to be larger in magnitude in periods when the economy is performing poorly as investors are worried about future consumption and, as such, utility of consumption is high. Hence, investors demand higher risk prices. It also appears that there is less variation in the market price of risk and less difference in risk prices over time. In contrast, we observe greater time variation and greater differences in the prices of HML and SMB risks between the exchanges. With respect to the latter, the risk prices tend to be larger on Amex and Nasdaq than on NYSE. This is possibly because the latter exchanges are, for the most part, the preferred habitat of retail investors who are more prone to vary their demand for compensation for exposure to risk. If markets were perfectly domestically integrated, then the difference in risk prices would be zero, on a period-by-period basis. While this is not likely in practice,

[17] The series have been smoothed using Hodrick-Prescott filters, with the recommended smoothing (14400) for monthly data

large differences, especially for lengthy periods, are indicative of segmentation. Figure 3 indicates that the differences are large in the 1999-2001 periods and then starts declining over the next few years. Interestingly, segmentation seems to increase again during the last years of the sample, which coincides with the financial crisis. Given that there was a crisis in both the 1999-2001 and the 2007-2009 periods, it seems plausible that segmentation is in fact driven by a differential rate of change of the risk aversion and, in turn, the expected compensation for bearing a unit of systematic risk, of the different investor groups.

5. WHY ARE THE MARKETS SEGMENTED?

In order to test why the prices of risk differ across the markets we first need to form a set of hypotheses as to why the prices of risk could differ. As stated in the introduction, different investors might prefer to hold stocks with specific characteristics, leading to investor segmentation and, therefore, potentially different prices of risk for different subsets of stock. Several papers suggest that this could be the case (see below). We term this the investor-segmentation hypothesis. From a market efficiency standpoint, segmented investors would be a necessary, but not sufficient, condition to cause different prices of risk across different securities. If a few rational, diversified investors exist, that would be enough to ensure efficient pricing of securities due to risky arbitrage (Fama (1965)). Hence, in order for domestic markets to be segmented, it is necessary to have both domestic securities held by different investors who demand different prices of risk for the securities they hold *and* arbitrage constraints. In this scenario, the appearance of differences in risk prices across the markets could be due to stocks with high arbitrage constraints and, therefore, be an artifact of the data. Bergbrant (2011) finds that idiosyncratic volatility, as measured by either expected idiosyncratic volatility from generalized autoregressive conditional heteroskedasticity models or realized idiosyncratic volatility, is unrelated to returns for large firms, consistent with the idea that small firms are arbitraged constraint. If arbitrage constraints are the reason that risk prices remain different, it would not be exploitable, and therefore would have limited real implications for investors. We term this the *arbitrage-driven segmentation* hypothesis.

Finally, as with all asset pricing tests, there is the risk of model misspecification. Since all asset pricing tests of integration are joint tests of both the model being used is well specified and the hypothesis of integration, differences in risk prices across the markets could be due to model misspecification, such as an omitted, priced factor. We attempt to control for this in several ways. First, we specify several different asset pricing models, based both on firms' sensitivities to systematic risk factors as well as firm characteristics. In addition, we employ tests that are not structured asset pricing tests and, therefore, less susceptible to the joint-hypothesis problem.

A. Investor-Segmentation Hypothesis

Several papers provide evidence that is consistent with the idea that different types of domestic securities are held by different investors. For instance, several papers find that institutional investors prefer certain stocks. Pruitt and Wei (1989) find that changes in institutional holdings are positively correlated with additions and deletions to the S&P500, suggesting that institutional investors prefer to hold stocks that are in the index. Grinstein and Michaely (2005) find that institutions have a preference for large stocks and stocks that pay dividends.[18]

Further, many funds such as index funds, by definition, hold a subset of firms. Shleifer (1986) and Lynch and Mendenhall (1997) show that firms' price rise substantially upon inclusion in the S&P500 index. Such a gain is likely due to mechanical rebalancing of index funds, but suggests that the firm becomes a part of a group of stocks owned by a different subset of investors, who possibly demand a different price of systematic risks. Given that Standard and Poor's explicitly states that inclusion in the index does not reflect a change in the firms' future prospects and, hence, the inclusion contains no new information the increase in the price of the stock is consistent with the

[18] See, also, Badrinath, Kale, and Noe (1995) who argue that some stocks are institutionally favored and some are institutionally unfavored.

new investors demanding a lower price of risks. Petajisto (2010) find that index funds *incur additional costs* by being mechanically tied to holding the index at all times, providing further evidence that investors care about certain subsets of firms. Considering their choice to incur these additional costs, they must enjoy greater benefits from the shares than other investors, suggesting that they might price the risks of the stock differently from other investors.

Other research has shown that there is a home bias in equity returns. Coval and Moskowitz (2002) find that investors have a preference for investing in firms that are local. Hong, Kubik, and Stein (2008) confirm these results domestically and find that in regions with few company headquarters stock prices are higher for those companies. This preference for local firms shows that investors prefer certain subsets of firms, which could lead to different prices of risk for different firms.

In addition, it is common to refer to mutual fund “style drift” as a negative phenomenon, partly because it alters the risk structure that funds claim to follow, and partly because the *managers stray away from their area of expertise*, which is detrimental to performance in the long run. Since investors often specialize in one area, such as growth stocks, or large stocks, they would likely hold a diversified portfolio with regards to the subset of assets that they specialize in. If these stocks are predominantly held by the same type of investors, they would evaluate systematic risk compared to their portfolio holdings, and could demand different prices of risks than other investors specializing in different stocks.

Furthermore, it has become a well-known fact that many investors are undiversified. Looking at retail investors, Goetzmann and Kumar (2004) find that more than 25% hold a single stock in their portfolios. Campbell, Lettau, Malkiel, and Zu (2001) argue that investors need approximately 50 stocks in their portfolios to be diversified. Goetzmann and Kumar (2004) find that less than 10% of retail investors' portfolios contain more than 10 stocks. If retail investors chose to invest in a different subset of firms than other investors, this could manifest itself in the form of different prices of risk for that subset of firms. In addition, it could lead to priced idiosyncratic volatility (if calculated with respect to the market portfolio). This has been documented by among other, Ang, Hodrick, Xing, Zhang (2006)) and Fu (2009). Bergbrant (2011) makes the point that the relation between idiosyncratic volatility and returns, expected as well as realized, is limited to firms which are difficult to arbitrage (low liquidity), and stocks more likely to be held by individual investors (small stocks).

To test the investor-segmentation hypothesis we look at subsets of firms which are likely to be predominantly traded by different subsets of investors. As mentioned above, mutual funds often focus on stocks with certain characteristics, such as large stocks or growth-oriented stocks. If segmented investors truly drive the different risk prices between the markets, then we would expect firms with different levels of the characteristics most likely to be used by different investor groups to screen their preferred stocks to have different risk prices even if they are listed on the same exchange. Alternatively, we would expect no differences in the prices of risk between different portfolios of firms with similar levels of these characteristics even if the portfolios are from different exchanges. We, therefore, test the investor-segmentation hypothesis by comparing the prices of risk for firms based on size (market capitalization), as well as based on growth prospects (book-to-market). We select size and book-to-market because they represent characteristics used by investors to screen stocks for their portfolio. It is widely believed, especially in the United States, that institutional investors make investment decisions based on categories.[19] Froot and Teo (2008) find evidence consistent with this belief, and find that institutions reallocate fund across groupings on size, book-to-market, as well as sectors. Kumar (2009) finds that this holds true also for individual investors. He finds that individual investors shift their preferences across extreme portfolios based on size and growth prospects. Bekaert, Hodrick, and Zhang (2009) note that "the style of a portfolio, value versus growth or small versus big, has long been a main organizing principle in the U.S. asset management industry."

Similarly, the behavioral finance literature also stresses the potential importance of style classification for stock return comovements. Boyer (2011) extends the co-movement literature by showing that arbitrary style categories (book-to-market) influences co-movement in asset prices. Hence, if the segmented investor hypothesis is the reason that the exchanges are segmented, and investors segment along the lines of style categories such as size and book-to-market, then we would expect the prices of risk to be different along these styles. We, therefore, test if the prices of risk are different for firms in groups based along these characteristics.

We re-estimate the regressions represented by equation (4) and reported in Table 2.4, but instead of allowing risk prices to differ across the three exchanges, we allow risk prices to differ for firms with different size and book-to-market ratios. It should be noted that the size portfolios are based on all the stocks that are listed on the three exchanges. In this way, the size-sorted portfolios are essentially equally affected by other characteristics

---

[19] See e.g. "Curtains Coming Down on the Sensational Small-Cap Show," The Financial Times, August 3, 2006, "Smart Money Stock Screen/Bargain Growth," The Wall Street Journal, May 18, 2006.

that might differentiate, say, Nasdaq and NYSE stocks and potentially cause segmentation. We sort firms on size and on book-to-market, respectively, into three equal groups (based on the number for firms). The results are presented in Table 2.6. The results indicate that the prices of risk differ, both statistically and economically, for the different size groups (Table 2.6, Panel A) as well as for the different book-to-market groups (Table 2.6, Panel B). Focusing on Model 1 (Panel A), the evidence indicates that market beta risk is significantly (*t*-stat of 5.14) priced for small firms and that the risk price is significantly larger than that for large firms. As indicated by the sum of the price of risk for small firms and the incremental risk price for large firms, the price of market risk for large firms is economically small. Thus, it appears that market beta risk is unrelated to returns for the largest stocks. The results are highly similar for the SMB factor. The risk price is economically large (0.581% per month) and significant for small firms, but are significantly smaller for medium firms and essentially zero for large firms. Interestingly, for the HML factor, the evidence indicates that the price of risk for small firms is of the opposite sign (negative) to that for medium and large firms. Assuming that HML reflects growth opportunities, this implies that greater growth opportunities reduce the cost of equity for small firms (see Fama and French (1992)).

Including a momentum and liquidity factor, Model 2, leads to essentially the same results. Turning to Models 3 and 4, the same results also hold when using characteristics instead of factor loadings. In these models, the prices of risk for all priced factors are economically larger for small firms than for medium or large firms. This is consistent with Sagi and Seasholes (2006) who find that momentum strategies yield higher returns for small firms.

Turning to Panel B, where we allow prices of risk factors to differ across firms with different levels of book-to-market, the inferences are highly similar. We find that the prices of risk are significantly larger (in magnitude) for firms with high book-to-market than for firms with low book-to-market ratios.

Taken together, these results suggest that investors with strong preferences for small (low book-to-market) firms demand a different compensation for exposure to a unit (beta) of the same systematic risk relative to investors that prefer large (high book-to-market) firms. The unwillingness of investors to change their portfolio allocation, therefore, lends support to the investor-segmentation hypothesis.

B. Arbitrage Constraints

We now consider the role, if any, of arbitrage constraints in the segmentation previously reported. It is important to note that it is undesirable for segmentation to be solely driven by micro-capitalization firms, which would make any investment strategies arising from segmentation difficult, if not impossible to implement. Therefore, for all tests in this essay, firms with a price of less than $5 have been excluded from the sample.

However, we are still concerned that arbitrage constraints could be the main driver of the segmentation that has been found in the preceding tests. In order to test our second hypothesis, the arbitrage-driven segmentation hypothesis, we exclude firms that are likely to have severe arbitrage constraints. Several papers suggest that small stocks have higher arbitrage constraints than large stocks. Wurgler and Zhuravskaya (2002) find that small stocks tend to have higher arbitrage risk and find a highly significant, negative correlation between their measure of risk of arbitrage and size. Furthermore, Pontiff (2006) argues that small stocks, due to their high transaction costs, are more expensive to arbitrage. This implies that size is a good proxy for arbitrage constraints. We, therefore, exclude the 10% of firms in our sample with the lowest market value and re-estimate the tests. The choice of excluding 10% of stocks is twofold. First, we require a large enough sample, so that we do not limit the stocks to those held by the same subset of investors. Secondly, we need to ensure that we eliminate all stocks that have severe arbitrage constraints, which could influence our results. Another common proxy for difficulties to arbitrage is idiosyncratic volatility. In additional tests (sub-section ii), we compare the prices of systematic risk factors excluding firms in the decile of highest lagged idiosyncratic volatility (see Ang, Hodrick, Xing, Zhang (2006)).

i) Excluding small firms

In order to determine if the differences in the prices of risk across the markets are due to arbitrage constraints we re-estimate the model in equation (4) excluding the smallest 10% of firms judged by market capitalization in the prior month. The results are presented in Table 2.7 (Panel A). Excluding the small firms does little to change the results shown earlier. Segmentation is still present between the exchanges. This suggests that arbitrage constraint is not the main driving force behind the segmentation.

ii) Excluding high idiosyncratic volatility firms

Table 2.7 (Panel B) shows the results when the 10% of firms with the highest idiosyncratic volatility in the prior month have been excluded. Once again, the results are very similar to those shown before. Segmentation is present even with these firms excluded. Hence, it does not appear as if arbitrage constraints are the reason that markets remain segmented.

C. Bad-Model Problem

There are strong indications to believe that the market segmentation found in this paper is not due to model misspecification. Several papers have provided evidence consistent with domestic market segmentation, using methods that are not subject to the joint-hypothesis problem. In a recent paper Goyal, Perignon, and Villa (2008) show that NYSE/Amex firms and Nasdaq firms share two common risk factors, but each have an additional risk factor, not shared by the other exchange. Even though this is in itself not evidence of segmentation, it is consistent with segmentation. The alternative explanation, that risk is present in the NYSE/Amex market which Nasdaq firms are not exposed to, is less compelling. In another recent paper, Hegde, Lin, and Varshney (2010) show that firms that chose to dual list on both the NYSE and Nasdaq benefit from higher liquidity and lower spreads. Further, studies investigating firms that chose to change their listing venue from Nasdaq to NYSE (Baker and Edelman (1992), Kadlec and McConnell (1994)) or dual list (Baker and Khan (1994)) find that firms that chose to do so experience abnormal returns. This, too, is consistent with, but not proof of segmentation. The abnormal returns can be the result of decreasing risk (lower betas), lower risk premia (segmentation) or a combination thereof. However, since merely changing listing location should not have an effect on the quantity (beta) of most systematic risks (perhaps with the exception of liquidity and investor recognition risks (see Kadlec and McConnell (1994)) significant changes in returns seem more consistent with changing price of risks and, hence, segmentation. This, as previously argued, would be the case if different investor groups have a strong preference for firms listed on a particular exchange and demand a different level of compensation for exposure to the same systematic risk than other investor groups that prefer firms listed on other exchanges.

Nonetheless, to ensure that our results are robust, we test for integration using the method suggested by Bekaert, Harvey, Lundblad, and Siegel (2011) which is model free and, therefore, is not susceptible to model misspecification.

It is important to note that even though these tests allow us to determine if the markets are statistically integrated, the economic implications of the results are less straightforward. However, if the domestic markets are segmented using this method, then it is less likely that our previous results are due to model misspecification.

Any model-specific test of integration is by definition a joint test between the asset pricing model used and integration. In order to make sure that the segmentation results are not driven by the specific asset pricing model, we employ a model-free test of integration. The segmentation measure, developed by Bekaert, Harvey, Lundblad, and Siegel (2011) rely on the notion that the earnings yield should be similar for similar firms, regardless of the exchange on which they trade. The exchange-level measure is therefore based on industry-level earnings yield in excess of the average earnings yield across the three exchanges. These differentials are aggregated across all industries in a given exchange. If the markets are integrated, then we would expect the earnings yield to be similar for firms in the same industry regardless of the exchange on which they trade, resulting in a value close to zero. The measure is constructed from firm level data from CRSP and Compustat. Every year, reported earnings (net income) are aggregated for each GIC industry group[20] in each US exchange (NYSE, Amex, Nasdaq). The earnings yield

[20] While Bekaert, Harvey, Lundblad, and Siegel (2011) use 38 industries according to DataStream industry classifications, we use 25 industries based on GIC industry groups.

(EY) for the particular industry (*j*) in each year (*t*) for each exchange (*i*) is then calculated by dividing the aggregate income by the aggregate market value for the included firms.[21] The measure of segmentation (SEG) is then calculated as the industry weighted difference between the absolute earnings yield in each exchange, industry, and year as compared to a global earnings yield measure (based on combining the firms in all three exchanges). Explicitly, SEG is calculated as

$$SEG_{i,t} = \sum_{j=1}^{N} IW_{i,j,t} \left| EY_{i,j,t} - EY_{w,j,t} \right|, \tag{5}$$

where $EY_{i,j,t} - EY_{i,j,t}$ is the difference in earnings yield (EY) between an industry (*j*) in a specific exchange (i) and that of the entire domestic market (w) in every year (t). The segmentation measure (SEG) is the industry weighted (IW) average of the absolute earnings yield differentials. This measure of segmentation only requires industry level earnings yield ratios, which are observed and not estimated, and therefore the measure is not subject to critiques of model misspecification. Since the integration measure is calculated solely for domestic firms, it is not subject to many of the shortcomings that would be present when applied internationally.[22] However, it is still subject to the bias arising from using net income reported by firms, and not true economic earnings (see Black (1980)). If the three exchanges are integrated, we would expect that our measure of integration, SEG, is relatively small and constant over time (Bekaert, Harvey, Lundblad, and Siegel (2011)). One major drawback with calculating the measure for the domestic exchanges is that the NYSE makes up a very large part of the entire US market. Hence, we would not expect any segmentation between the NYSE and the total US market.

[21] Following Bekaert, Harvey, Lundblad, and Siegel (2011) we set negative earnings equal to 0.
[22] For a discussion of these shortcomings, see Bekaert, Harvey, Lundblad, and Siegel (2011).

Therefore, instead of reporting the segmentation per exchange, we report the average segmentation score of the three exchanges over time.[23] This could be seen as a measure of how integrated the Amex and Nasdaq exchanges are with the NYSE.

The results from the model-free segmentation measure are presented in Figure 1. As can be seen, US domestic segmentation has decreased over time.[24] This is interesting for at least two reasons. First, the fact that domestic segmentation has decreased is evidence that the domestic markets have been segmented during the period investigated. As discussed earlier, if the markets were not segmented, we would expect the measure of segmentation to be fairly constant over time. Secondly, the decrease in segmentation is consistent with the qualitative results presented in Table 2.4. It is also interesting to note that there is an increase in segmentation during the most recent financial crisis.

Overall, the tests in this section provide strong support for the hypothesis that US stock markets are domestically segmented. More important, they support the hypothesis that the main cause of segmentation is the investor-segmentation hypothesis–stock markets attract firms with certain general characteristics and certain investors have a stronger preference for these firms and not for other firms with different characteristics and which are listed on other exchanges. Given that these investors only hold a part of the universe of stocks their demand for compensation for exposure to a unit of systematic risk varies with the portfolios they hold, thus leading to segmentation.

[23] The individual exchange segmentation scores are available from the authors upon request.
[24] The time trend in segmentation is statistically significant, with a t-statistic of -3.11.

6. CONCLUSION

We investigate whether the three main U.S. equity markets are domestically integrated by comparing the price of commonly used risk factors across the NYSE, Amex, and Nasdaq. We find that the markets have significantly different prices of risks for several risk factors, indicating that the markets are segmented. The magnitude of the difference is both statistically and economically significant. Further, we find that the segmentation has become *less pronounced* over the last decade, but that the markets remain segmented. The segmentation holds both for specific asset pricing model specifications, as well as for model-free integration tests. We also find evidence against the hypothesis that the segmentation is due to arbitrage constraints. Instead, we find evidence consistent with the investor-segmentation hypothesis, in which different investors choose to hold different subsets of firms and demand different prices of risk among the different groups of securities. These results highlight the value of diversification and suggest that domestic equity markets are not fully efficient. Further, it raises a concern with respect to testing for integration of international equity markets. If the characteristics of the stocks on a particular market impact the price of systematic risks, then testing integration from a standpoint where the price of risks is the same for all stocks would be troublesome, given that firm characteristics differ widely across international markets. Further, we open up the discussion of whether investors can profitably exploit the segmentation that has been revealed, both domestically and internationally.

7. REFERENCES

Ang, A., Hodrick, R., Xing, Y., Zhang, X., 2006. The cross-section of volatility and expected returns. *Journal of Finance* 61, 259-299.

Ang, A., Shtauber, A., Tetlock, P., 2011. Asset pricing in the dark: The cross section of OTC Stocks. *Working Paper*, Columbia University.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time- series effects. *Journal of Financial Markets* 5, 31–56.

Amihud, Y., Mendelson, H., 1986. Asset pricing and the bid-ask spread. *Journal of Financial Economics* 15, 223–249.

Badrinath, S., Kale, J., Noe, T., 1995. Of Shepherds, Sheep, and the Cross-autocorrelations in Equity Returns. *Review of Financial Studies* 8, 401-430.

Baker, K., Edelman, R., 1992. AMEX to NYSE transfers, market microstructure, and shareholder wealth. *Financial Management* 21, 60-72.

Baker, K., Khan, W., 1994. The post dual listing anomaly. *Journal of Economics and Business* 46, 287-297.

Barber, B., Odean., T., 2000. Trading is hazardous to your wealth; the common stock investment performance of individual investors. *Journal of Finance* 55, 773-806.

Barberis, N., Shleifer, A., 2003. Style investing. *Journal of Financial Economics* 68, 161-199.

Barberis, N., Shleifer, A., Wurgler, J., 2005. Comovement. *Journal of Financial Economics* 75, 283-317.

Baruch, S., Saar, G., 2009. Asset returns and the listing choice of firms. *Review of Financial Studies* 22, 2239-2274.

Bassin, W., 1998. An evaluation of a trading strategy for stocks that switch exchanges from the Nasdaq to the NYSE. Working paper.

Bekaert, G., Harvey, C., 1995. Time-varying world market integration. *Journal of Finance* 50, 403– 444.

Bekaert, G., Hodrick, R., Zhang, X., 2009. International stock return comovements. *Journal of Finance* 64, 2591-2626.

Bekaert, G., Harvey, C., Lundblad, C., Siegel, S., 2007. Global growth opportunities and market integration. *Journal of Finance* 62, 1081-1137.

Bekaert, G., Harvey, C., Lundblad, C., Siegel, S., 2007. What segments equity markets? *Forthcoming in the Review of Financial Studies*, 2012.

Bergbrant, M., 2011. Is idiosyncratic volatility really priced? *Working paper*, University of South Florida.

Black, F., 1980. The magic in earnings: Economic earnings versus accounting earnings. *Financial Analyst Journal* 36, 19-24

Boyer, B., 2011. Style-Related Comovement: Fundamentals or Labels? *Journal of Finance* 66, 307-332.

Brennan, M., Subrahmanyam, A., 1996. Market microstructure and asset pricing: on the compensation for illiquidity in stock returns. *Journal of Financial Economics* 41, 441–464.

Campbell, J., Lettau, M., Malkiel, B., Xu, Y., 2001. Have individual stocks become more volatility? An empirical exploration of idiosyncratic risk. *Journal of Finance* 56, 1-43.

Carhart, M., 1997. On persistence in mutual fund performance. *Journal of Finance* 52, 57-82.

Carrieri, F., Chaieb, I., Errunza, V., 2010. Do implicit barriers matter for globalization? *Working paper*.

Chen, Z., Knez, P., 1995. Measurement of market integration and arbitrage. *Review of Financial Studies* 8, 287-325.

Chen, L., Novy-Marx, R., Zhang, L. 2010. An alternative three-factor model. *Working Paper.*

Chua C., Goh, J., Zhang, Z. 2010. Expected volatility, unexpected volatility, and the cross-section of stock returns. *Journal of Financial Research* 33, 103-123.

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

Corwin, S., Harris, J., 2001. The initial listing decisions of firms that go public. *Financial Management*, 30, 35-55.

Coval, J., Moskowitz, T., 2002. Home bias at home: local equity preference in domestic portfolios. *Journal of Finance* 54, 2045-2073.

Daniel, K., Titman, S., 1997. Evidence on the characteristics of cross sectional variation in stock returns. *Journal of Finance* 52, 1-33.

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

Fama, E., 1965. The behavior of stock market prices. *Journal of Business* 38, 34-105.

Fama, E., French, K., 1992. The cross-section of expected stock returns. *Journal of Finance* 47, 427-465.

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

Fama, E., French, K., 2000. Disappearing dividend: changing firm characteristics or lower propensity to pay? *Journal of Financial Economics* 60, 3-43.

Fama, E., French, K., 2004. New Lists: fundamentals and survivor rates. *Journal of Financial Economics* 73, 229-269.

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

Flood, R., Rose, A., 2005. Financial integration: A new methodology and an illustration. *Journal of the European Economic Association* 3, 1349-1359.

Froot, K., Teo, M., 2008. Style investing and institutional investors. Journal of Financial and Quantitative Analysis 43, 883-906.

Fu, F., 2009. Idiosyncratic risk and the cross-section of expected stock returns. *Journal of Financial Economics* 91, 24-37

Goetzmann, W., Kumar, A., 2004. Why do individual investors hold under-diversified portfolios? *Working paper*. Yale University.

Goyal, A., Perignon, C., Villa, C., 2008. How common are common return factors across the NYSE and Nasdaq. *Journal of Financial Economics* 90, 252-271.

Grinstein, Y., Michaely, R., 2005. Institutional holdings and payout policy. *Journal of Finance* 60, 1389-1426.

Hegde, S., Lin, H., Varshney, S., 2010. Competitive stock markets: Evidence from companies' dual listings on the NYSE and Nasdaq. *Financial Analysts Journal* 66.

Hong, H., Kubik, J., Stein, J., 2008. The only game in town: stock-price consequences of local bias. *Journal of Financial Economics* 90, 20-37

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

Jegadeesh, N., Titman, S., 2001. Profitability of momentum strategies: An evaluation of alternative explanations. *Journal of Finance* 56, 699-720

Kaul, A., Mehrotra, V., Stefanescu, C., 2006. Habitats and return comovement: evidence from firms that switch exchanges. *Working Paper*.

Karolyi, A., 1995. A multivariate GARCH model of international transmissions of stock returns and volatility: The case of the United States and Canada. *Journal of Business and Economic Statistics* 13, 11-25.

Kadlec, G., McConnell, J., 1994. The Effect of Market Segmentation and Illiquidity on Asset Prices: Evidence from Exchange Listings. *Journal of Finance* 49, 611-636.

King, M., Sentana, E., Wadhwani, S., 1994. Volatility and links between national stock markets. *NBER working paper* No. 3357.

Kumar, A., 2009. Dynamic style preferences of individual investors and stock returns. Journal of Financial and Quantitative Analysis 44, 607-640.

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

Lintner, J., 1971. Capital market equilibrium with divergent borrowing and lending rates. *Journal of Financial and Quantitative Analysis* 6, 1197-1205.

Lynch, A., Mendenhall, R., 1997. New evidence on stock price effects associated with changes in the S&P 500 index. *Journal of Business* 70, 351-383.

Malkiel, B., Xu, Y., 2002. Idiosyncratic risk and security returns. *Working paper*. University of Texas at Dallas.

Naranjo, A., Protopapadakis, A., 1997. Financial market integration test: an investigation using US equity markets. *Journal of International Financial Markets, Institutions, and Money* 7, 93-135.

Parsley, D., Schlag, C., 2007. Measuring financial integration via idiosyncratic risk: What effects are we really picking up? *Journal of Money, Credit, and Banking* 39, 1267-1273.

Pastor, L., Stambaugh, R., 2003. Liquidity risk and expected stock returns. *Journal of Political Economy* 111, 642–685.

Petajisto, A., 2010. The index premium and its hidden cost for index funds. *Journal of Empirical Finance* 18, 271-288.

Pontiff , J., 2006. Costly arbitrage and the myth of Idiosyncratic Risk. *Journal of Accounting and Economics* 42, 35-42.

Pruitt, S., Wei, J., 1989. Institutional ownership and changes in the S&P500. *Journal of Finance* 44, 509-513.

Rubinstein, M. 1973. A mean-variance synthesis of corporate financial theory. *Journal of Finance* 28, 167-181.

Rubinstein, M. 1973. The fundamental theorem of parameter-preference security valuation. *Journal of Financial and Quantitative Analysis* 8, 61-69.

Sagi, J., Seasholes, M., 2006. Firm-specific attributes and the cross-section of momentum. *Journal of Financial Economics* 84, 389-434.

Shleifer, A., 1986. Do demand curves for stocks slope down? *Journal of Finance* 41, 579-590.

Spiegel, M., Wang, X., 2007. Cross-sectional variation in stock returns: liquidity and idiosyncratic risk. *Working paper*.

Wurgler, J., Zhuravskaya, E., 2002. Does arbitrage flatten demand curves for stocks? *Journal of Business 75*, 583-608

Table 2.1: Summary statistics

	NYSE		Amex		Nasdaq	
	Mean	Std	Mean	Std	Mean	Std
N	388065		82926		562314	
RET(%)	1.41	11.31	2.23	14.63	2.38	16.56
PRC($)	69.17	1710.68	19.94	50.06	19.66	33.39
VOL	14.86	68.46	0.68	3.24	7.17	47.51
beta_mkt	1.04	0.75	0.86	0.88	0.92	0.96
beta_smb	0.47	1.04	0.84	1.23	0.84	1.24
beta_hml	0.32	1.19	0.27	1.39	0.12	1.49
beta_umd	-0.08	0.87	-0.04	1.07	-0.07	1.09
beta_liq	0.02	0.85	0.28	1.09	0.17	1.12
LN(ME)	6.84	1.43	4.19	1.26	5.01	1.50
LN(BE/ME)	-0.57	0.68	-0.37	0.74	-0.68	0.76
RET(-12,-2)	0.13	0.37	0.19	0.43	0.20	0.51
AMIILL	0.07	0.32	0.79	1.30	0.71	2.06

Table 2.1 shows descriptive statistics for the pooled sample. RET is the average monthly holding period return, reported in percent. PRC($) is the average stock price. VOL is the average monthly volume (in millions of shares). Beta_x are the firm sensitivity to the systematic risk factor x. Every month we regress individual firm returns from *t-24* to *t-1* on the proxies for the risk factors and record the beta. To estimate the firm sensitivities (betas) to the Fama and French (1993) three factors, we regress monthly firm returns on the CRSP value weighted index, the SMB, and the HML factor. In order to estimate the betas for the momentum and liquidity factors, we regress the returns on the Fama and French (1993) factors as well as the other factor which beta we want to estimate. We roll the regressions monthly. To be eligible for estimation, we require that a firm has returns in all the months of the estimation period (e.g., 24 months of data). ME is defined as shares outstanding times the share price and is updated in June of every year (t) and used from July in that year (*t*) until June the following year (*t+1*). BE/ME is book-to-market ratio estimated as in Fama and French (1992); the fiscal year end book value (*t-1*) is divided by the calendar year end (*t-1*) market equity. A firm's BE/ME is assigned to the firm from July in the year after it was estimated (*t*), until June (*t+1*) of the following year in order to ensure that it was available to investors at the time it is used to predict returns. RET(-12,-2) is the average stock return from month *t-12* to *t-2*. AMIILL is Amihuds (2002) illiquidity measure estimated in the previous month. All explanatory variables have been winsorized at the 5% level.

Table 2.2: Cross-sectional correlations

		RET(%)	beta_mkt	beta_smb	beta_hml	beta_umd	beta_liq	LN(ME)	LN(BE/ME)	RET(-12,-2)	AMIILL
	RET(%)	1.000	0.010	0.018***	0.015***	-0.009	0.014***	-0.034***	0.025***	0.006	0.029***
	beta_mkt	0.010	1.000	0.063***	0.354***	-0.100***	0.051***	0.036***	0.001	-0.032**	-0.058***
	beta_smb	0.018***	0.063***	1.000	0.227***	-0.079***	0.054***	-0.357***	0.089***	-0.006	0.090***
	beta_hml	0.015***	0.354***	0.227***	1.000	-0.018*	0.005	-0.072***	0.226***	-0.022*	0.029***
NYSE	beta_umd	-0.009	-0.100***	-0.079***	-0.018*	1.000	-0.371***	0.042***	-0.079***	0.110***	-0.023***
	beta_liq	0.014***	0.051***	0.054***	0.005	-0.371***	1.000	-0.088***	0.061***	-0.044***	0.079***
	LN(ME)	-0.034***	0.036***	-0.357***	-0.072***	0.042***	-0.088***	1.000	-0.322***	0.046***	-0.395***
	LN(BE/ME)	0.025***	0.001	0.089***	0.226***	-0.079***	0.061***	-0.322***	1.000	0.008	0.166***
	RET(-12,-2)	0.006	-0.032**	-0.006	-0.022*	0.110***	-0.044***	0.046***	0.008	1.000	-0.076***
	RET(%)	1.000	0.015**	0.010	0.013**	-0.012**	0.034***	-0.061***	0.018***	0.001	0.090***
	beta_mkt	0.015**	1.000	0.054***	0.394***	-0.028***	0.027***	0.141***	-0.113***	0.053***	-0.093***
	beta_smb	0.010	0.054***	1.000	0.208***	-0.017***	0.040***	0.003	-0.075***	0.034***	-0.005
	beta_hml	0.013**	0.394***	0.208***	1.000	0.027***	0.022***	-0.019***	0.096***	-0.010	0.042***
Amex	beta_umd	-0.012**	-0.028***	-0.017***	0.027***	1.000	-0.334***	0.005	-0.062***	0.058***	-0.026***
	beta_liq	0.034***	0.027***	0.040***	0.022***	-0.334***	1.000	-0.171***	0.075***	0.000	0.096***
	LN(ME)	-0.061***	0.141***	0.003	-0.019***	0.005	-0.171***	1.000	-0.352***	0.003	-0.496***
	LN(BE/ME)	0.018***	-0.113***	-0.075***	0.096***	-0.062***	0.075***	-0.352***	1.000	0.018**	0.243***
	RET(-12,-2)	0.001	0.053***	0.034***	-0.010	0.058***	0.000	0.003	0.018**	1.000	-0.155***
	RET(%)	1.000	0.008	0.013**	0.003	-0.013***	0.035***	-0.073***	0.028***	-0.005	0.053***
	beta_mkt	0.008	1.000	0.059***	0.259***	-0.049***	0.024***	0.174***	-0.157***	-0.006	-0.129***
	beta_smb	0.013**	0.059***	1.000	0.165***	-0.035***	0.059***	-0.021***	-0.072***	0.019*	-0.046***
	beta_hml	0.003	0.259***	0.165***	1.000	0.011	0.013*	-0.109***	0.210***	-0.014	0.064***
Nasdaq	beta_umd	-0.013***	-0.049***	-0.035***	0.011	1.000	-0.345***	0.007	-0.065***	0.071***	-0.014***
	beta_liq	0.035***	0.024***	0.059***	0.013*	-0.345***	1.000	-0.171***	0.083***	0.024**	0.091***
	LN(ME)	-0.073***	0.174***	-0.021***	-0.109***	0.007	-0.171***	1.000	-0.385***	-0.023***	-0.442***
	LN(BE/ME)	0.028***	-0.157***	-0.072***	0.210***	-0.065***	0.083***	-0.385***	1.000	0.030***	0.204***
	RET(-12,-2)	-0.005	-0.006	0.019*	-0.014	0.071***	0.024**	-0.023***	0.030***	1.000	-0.092***

This table displays the time-series means of the cross-sectional Pearson correlation coefficients. In every month, for each exchange, we obtain the contemporaneous Pearson correlation coefficients between each pair of variables using the cross section of stocks. The reported correlation coefficients are the average over the sample period of 300 months. The coefficients relate to a sample of stocks traded NYSE, Amex, and Nasdaq between Jan 1985 and December 2009. Stars denote significance at the 10% (*), 5% (**), and 1% (***) respectively.

Table 2.3: Cross sectional regressions with Fama-MacBeth (1973) T-stats

		beta_mkt	beta_smb	beta_hml	beta_umd	beta_liq	LN(ME)	LN(BE/ME)	RET(-12,-2)	AMIILL	Adj R-Sq	OBS
Model 1	NYSE	0.326*	0.218**	0.156*							3.60%	388065
		(1.84)	(2.48)	(1.75)								
	Amex	0.667***	0.317***	-0.039							2.40%	82925
		(3.48)	(2.65)	(-0.31)								
	Nasdaq	0.473***	0.308***	-0.019							3.20%	562314
		(2.64)	(2.64)	(-0.17)								
Model 2	NYSE	0.283	0.192**	0.190**	0.031	0.251***					4.60%	388065
		(1.63)	(2.23)	(2.08)	(0.30)	(2.88)						
	Amex	0.655***	0.317***	-0.063	0.187	0.639***					3.20%	82925
		(3.48)	(2.64)	(-0.50)	(1.33)	(4.57)						
	Nasdaq	0.465***	0.293**	-0.005	0.118	0.640***					3.80%	562314
		(2.68)	(2.54)	(-0.05)	(1.32)	(6.21)						
Model 3	NYSE	0.336**					-0.245***	0.174**			3.00%	380438
		(2.18)					(-4.25)	(2.09)				
	Amex	0.646***					-0.731***	-0.130			2.20%	78679
		(4.21)					(-11.68)	(-0.99)				
	Nasdaq	0.488***					-0.821***	-0.170			3.10%	510302
		(3.24)					(-12.08)	(-1.08)				
Model 4	NYSE	0.129					-0.190***	0.149*	-0.239	0.614***	4.70%	378834
		(1.28)					(-3.20)	(1.87)	(-0.51)	(4.51)		
	Amex	0.511***					-0.503***	-0.250*	0.111	0.739***	4.20%	65949
		(3.58)					(-6.12)	(-1.82)	(0.31)	(6.97)		
	Nasdaq	0.336***					-0.817***	-0.112	-0.395	0.266***	4.10%	459264
		(2.93)					(-10.00)	(-0.76)	(-1.18)	(5.31)		

We estimate cross sectional regressions for every month from July of 1985 to December of 2009, a total of 300 months. The dependent variable in the monthly regressions is the holding period return on all firms in our sample that have information on all explanatory variables in that month. The explanatory variables are as defined in Table 2.1. All the explanatory variables are pre-determined at time *t*, and are used to explain the variation in returns at time *t*. For every month, we estimate a model that is nested in the following cross- sectional regression:

$$\mathrm{R}_{\mathrm{i,t}} = \lambda_0 + \lambda_1\beta_{i,t}^{Mkt} + \lambda_2\beta_{i,t}^{SMB} + \lambda_3\beta_{i,t}^{HML} + \lambda_4\beta_{i,t}^{UMD} + \lambda_5\beta_{i,t}^{LIQ} + \lambda_6 LN(ME)_{i,t} + \lambda_7 LN\left(\frac{BE}{ME}\right)_{i,t} + \lambda_8 RET(-12,-2)_{i,t} + \lambda_9 AMIILL_{i,t} + \epsilon_{\mathrm{i,t,}}$$

This table summarizes the means of the coefficients from these cross sectional regressions (commonly referred to as risk prices). The standards errors are calculated as in Fama and MacBeth (1973), but corrected for serial correlation as in Newey and West (1987). The corresponding test-statistics are presented in parentheses below each average coefficient. Stars denote significance at the 10% (*), 5% (**), and 1% (***) respectively.

Table 2.4: Cross sectional regressions with interactions

		beta_mkt	beta_smb	beta_hml	beta_umd	beta_liq	LN(ME)	LN(BE/ME)	RET(-12,-2)	AMIILL	Adj R-Sq	OBS
Model 1	NYSE	0.326*	0.218**	0.156*							3.90%	1033304
		(1.84)	(2.48)	(1.75)								
	Amex	0.341**	0.099	-0.195*								
		(2.26)	(0.92)	(-1.71)								
	Nasdaq	0.147	0.089	-0.175**								
		(1.44)	(0.95)	(-2.31)								
Model 2	NYSE	0.283	0.192**	0.190**	0.031	0.251***					4.70%	1033304
		(1.63)	(2.23)	(2.08)	(0.30)	(2.88)						
	Amex	0.371**	0.125	-0.252**	0.156	0.388***						
		(2.32)	(1.17)	(-2.10)	(1.24)	(3.35)						
	Nasdaq	0.181*	0.101	-0.195**	0.087	0.390***						
		(1.76)	(1.10)	(-2.54)	(1.18)	(5.56)						
Model 3	NYSE	0.336**					-0.245***	0.174**			3.70%	969419
		(2.18)					(-4.25)	(2.09)				
	Amex	0.309**					-0.486***	-0.304**				
		(2.49)					(-7.22)	(-2.57)				
	Nasdaq	0.152*					-0.576***	-0.344***				
		(1.68)					(-8.58)	(-2.84)				
Model 4	NYSE	0.129					-0.190***	0.149*	-0.239	0.614***	4.70%	904047
		(1.28)					(-3.20)	(1.87)	(-0.51)	(4.51)		
	Amex	0.382***					-0.313***	-0.399***	0.350	0.126		
		(2.85)					(-3.75)	(-3.21)	(1.03)	(0.76)		
	Nasdaq	0.207**					-0.627***	-0.262**	-0.156	-0.347***		
		(2.14)					(-9.05)	(-2.31)	(-0.63)	(-2.60)		

We estimate cross sectional regressions for every month from July of 1985 to December of 2009, a total of 300 months. This regression is similar to the one in the previous table (Table 2.3), except that we create dummies for the different exchanges and interact those with the explanatory variables in the regression specification. We use NYSE as the base group. The coefficients can then be interpreted as the incremental effect of each explanatory variable on the Amex and Nasdaq exchanges, respectively, as compared to the NYSE. The standards errors are calculated as in Fama and MacBeth (1973), but corrected for serial correlation as in Newey and West (1987). The corresponding test-statistics are presented in parentheses below each average coefficient. Stars denote significance at the 10% (*), 5% (**), and 1% (***) respectively.

$$R_{i,t} = a_t + a_{at}\,\text{Amex} + a_{nt}\text{Nasdaq} + \sum_{k=1}^{5}\lambda_{k,t}\beta_{i,t}^{k} + \sum_{k=1}^{5}\lambda_{ak,t}\,(\beta_{i,t}^{k} \times \text{Amex}) + \sum_{k=1}^{5}\lambda_{nk,t}\,(\beta_{i,t}^{k} \times Nasdaq) + \varepsilon_{i,t}$$

Table 2.5: Sub-Period integration tests

		Beta_Mkt	Beta_SMB	Beta_HML	Beta_UMD	Beta_Liq	LN(ME)	LN(BE/ME)	RET(-12,-2)	AMIILL	Adj R-Sq	OBS
PANEL A: 1985-1992												
Model 1	NYSE	0.136	0.156	0.042							3.00%	301871
		(0.49)	(1.06)	(0.37)								
	Amex	0.117	0.029	0.202**								
		(0.62)	(0.25)	(2.42)								
	Nasdaq	0.321**	0.131	-0.070								
		(2.20)	(1.30)	(-0.92)								
Model 2	NYSE	0.176	0.159	0.023	0.053	0.108					3.60%	301871
		(0.64)	(1.12)	(0.20)	(0.52)	(1.16)						
	Amex	0.102	0.012	0.201**	0.003	0.096						
		(0.56)	(0.11)	(2.44)	(0.03)	(1.21)						
	Nasdaq	0.311**	0.118	-0.079	0.071	0.214***						
		(2.16)	(1.23)	(-1.03)	(0.83)	(2.72)						
Model 3	NYSE	0.164					-0.102	0.160			3.00%	262554
		(0.63)					(-1.00)	(1.00)				
	Amex	0.399**					-0.526***	-0.116				
		(2.03)					(-4.75)	(-0.77)				
	Nasdaq	0.282**					-0.601***	-0.241*				
		(2.07)					(-6.95)	(-1.81)				
Model 4	NYSE	0.099					-0.011	0.154	0.346	0.720***	3.60%	240915
		(0.44)					(-0.11)	(1.00)	(0.91)	(5.10)		
	Amex	0.216					-0.295**	-0.213	-0.063	-0.096		
		(1.05)					(-2.42)	(-1.41)	(-0.19)	(-0.66)		
	Nasdaq	0.213					-0.620***	-0.198	0.396	-0.507***		
		(1.51)					(-8.46)	(-1.57)	(1.31)	(-3.61)		

Table 2.5 (cont.)

		Beta_Mkt	Beta_SMB	Beta_HML	Beta_UMD	Beta_Liq	LN(ME)	LN(BE/ME)	RET(-12,-2)	AMIILL	Adj R-Sq	OBS
PANEL B: 1993-2000												
Model 1	NYSE	0.194	0.037	0.092							3.90%	381638
		(1.23)	(0.18)	(0.52)								
	Amex	0.438	0.424*	-0.536**								
		(1.41)	(1.75)	(-2.02)								
	Nasdaq	-0.060	0.345	-0.255								
		(-0.31)	(1.51)	(-1.52)								
Model 2	NYSE	0.142	0.014	0.106	-0.126	0.121					4.30%	381638
		(0.94)	(0.07)	(0.61)	(-1.05)	(1.03)						
	Amex	0.501	0.416*	-0.515*	0.020	0.511***						
		(1.53)	(1.68)	(-1.88)	(0.13)	(3.61)						
	Nasdaq	-0.013	0.345	-0.262	0.003	0.367***						
		(-0.07)	(1.53)	(-1.57)	(0.03)	(3.30)						
Model 3	NYSE	0.201*					-0.148	0.210			3.50%	364024
		(1.81)					(-1.62)	(1.35)				
	Amex	0.082					-0.516***	-0.437**				
		(0.41)					(-4.98)	(-1.99)				
	Nasdaq	-0.272*					-0.804***	-0.586*				
		(-1.67)					(-5.51)	(-1.88)				
Model 4	NYSE	0.125					-0.090	0.186	0.650	0.943***	4.30%	336854
		(1.32)					(-0.89)	(1.23)	(1.59)	(3.72)		
	Amex	0.160					-0.361***	-0.464**	0.270	-0.076		
		(0.80)					(-3.09)	(-2.03)	(0.61)	(-0.24)		
	Nasdaq	-0.207					-0.840***	-0.479*	-0.503	-0.570**		
		(-1.47)					(-4.91)	(-1.69)	(-1.35)	(-2.22)		

Table 2.5 (cont.)

						PANEL C: 2001-2009						
		Beta_Mkt	Beta_SMB	Beta_HML	Beta_UMD	Beta_Liq	LN(ME)	LN(BE/ME)	RET(-12,-2)	AMIILL	Adj R-Sq	OBS
Model 1	NYSE	0.613	0.436***	0.313*							4.80%	349795
		(1.54)	(3.96)	(1.75)								
	Amex	0.453*	-0.128	-0.244								
		(1.73)	(-0.79)	(-1.36)								
	Nasdaq	0.177	-0.175	-0.197								
		(0.99)	(-1.39)	(-1.55)								
Model 2	NYSE	0.504	0.379***	0.411**	0.150	0.493**					5.90%	349795
		(1.27)	(3.45)	(2.27)	(0.61)	(2.47)						
	Amex	0.495*	-0.034	-0.422**	0.415	0.539*						
		(1.70)	(-0.21)	(-2.14)	(1.36)	(1.91)						
	Nasdaq	0.239	-0.132	-0.239*	0.176	0.565***						
		(1.25)	(-1.10)	(-1.81)	(1.06)	(3.89)						
Model 3	NYSE	0.610*					-0.459***	0.155			4.50%	342841
		(1.78)					(-4.96)	(1.28)				
	Amex	0.432*					-0.424***	-0.354				
		(1.81)					(-3.33)	(-1.58)				
	Nasdaq	0.414***					-0.351***	-0.220				
		(2.94)					(-3.74)	(-1.44)				
Model 4	NYSE	0.159					-0.439***	0.113	-1.549	0.227	5.90%	326278
		(0.92)					(-4.92)	(1.04)	(-1.34)	(0.87)		
	Amex	0.727***					-0.286	-0.507**	0.788	0.502		
		(2.79)					(-1.64)	(-2.10)	(0.99)	(1.50)		
	Nasdaq	0.569***					-0.444***	-0.125	-0.339	-0.008		
		(3.21)					(-5.12)	(-0.83)	(-0.65)	(-0.03)		

We estimate cross sectional regressions for every month from July of 1985 to December of 2009, a total of 300 months. We then divide the sample up into three time period; the first time period covers years 1985-1992, the second 1993-2000, and the third 2001-2009. This regression is similar to the one in Table 2.4. We use NYSE as the base group. The coefficients can then be interpreted as the incremental effect of each explanatory variable on the Amex and Nasdaq exchanges as compared to the NYSE. Panel A show the results for the first time period, 1985-1992. Panel B shows the results for the second time period, 1993-2000, and Panel C shows the results for the last time period, 2001-2009. The standards errors are calculated as in Fama and MacBeth (1973), but corrected for serial correlation as in Newey and West (1987). Stars denote significance at the 10% (*), 5% (**), and 1% (***) respectively.

$$\mathrm{R}_{\mathrm{i,t}} = \mathrm{a}_{\mathrm{t}} + \mathrm{a}_{\mathrm{at}}\,\mathrm{Amex} + \mathrm{a}_{\mathrm{nt}}\mathrm{Nasdaq} + \sum_{k=1}^{5}\lambda_{\mathrm{k,t}}\beta_{i,t}^{k} + \sum_{k=1}^{5}\lambda_{\mathrm{ak,t}}\,(\beta_{i,t}^{k}\times \mathrm{Amex}) + \sum_{k=1}^{5}\lambda_{\mathrm{nk,t}}\,(\beta_{i,t}^{k}\times Nasdaq) + \varepsilon_{\mathrm{i,t}}$$

Table 2.6: Comparing prices of risk for firms with different characteristics

PANEL A: Size

		Beta_Mkt	Beta_SMB	Beta_HML	Beta_UMD	Beta_Liq	LN(ME)	LN(BE/ME)	RET(-12,-2)	AMIILL	Adj R-Sq	OBS
	Low	1.007***	0.581***	-0.355***								
		(5.14)	(4.46)	(-2.93)								
Model 1	Medium	-0.523***	-0.491***	0.413***							4.60%	1033304
		(-6.68)	(-6.25)	(5.40)								
	High	-0.745***	-0.543***	0.481***								
		(-6.47)	(-5.49)	(4.96)								
	Low	0.953***	0.552***	-0.317***	-0.633***	0.065						
		(5.07)	(4.35)	(-2.70)	(-5.11)	(0.68)						
Model 2	Medium	-0.501***	-0.470***	0.388***	-0.145*	-0.490***					5.30%	1033304
		(-6.48)	(-6.32)	(5.43)	(-1.85)	(-6.19)						
	High	-0.725***	-0.514***	0.456***	-0.089	-0.598***						
		(-6.25)	(-5.44)	(4.93)	(-0.85)	(-7.18)						
	Low	0.829***					-1.745***	-0.522***				
		(4.67)					(-11.88)	(-3.21)				
Model 3	Medium	-0.469***					1.346***	0.682***			3.90%	969419
		(-5.33)					(9.22)	(7.87)				
	High	-0.643***					1.662***	0.632***				
		(-4.97)					(10.43)	(5.63)				
	Low	0.611***					-2.286***	-0.405***	-7.612***	-1.044***		
		(4.35)					(-11.44)	(-2.60)	(-9.17)	(-2.94)		
Model 4	Medium	-0.387***					1.843***	0.572***	0.671***	0.205	5.30%	904047
		(-4.05)					(10.38)	(6.11)	(3.61)	(0.61)		
	High	-0.600***					2.206***	0.514***	1.082***	7.326		
		(-4.07)					(11.01)	(4.23)	(3.88)	(1.62)		

Table 2.6 (cont.)

		Beta_Mkt	Beta_SMB	Beta_HML	Beta_UMD	Beta_Liq	LN(ME)	LN(BE/ME)	RET(-12,-2)	AMIILL	Adj R-Sq	OBS
PANEL B: B/M												
	Low	0.378**	0.282**	0.052								
		(2.34)	(2.48)	(0.48)								
Model 1	Medium	0.060	0.116**	-0.194***							3.80%	1033304
		(0.93)	(2.16)	(-3.59)								
	High	0.228**	0.220***	-0.298***								
		(2.38)	(2.94)	(-3.64)								
	Low	0.378**	0.265**	0.053	-0.023	0.019						
		(2.36)	(2.36)	(0.49)	(-0.21)	(0.25)						
Model 2	Medium	0.032	0.118**	-0.164***	0.145	-0.032					4.40%	1033304
		(0.50)	(2.32)	(-3.14)	(1.55)	(-0.37)						
	High	0.178**	0.182**	-0.231***	0.187**	0.217**						
		(2.00)	(2.55)	(-3.06)	(2.15)	(2.43)						
	Low	0.219*					-0.584***	-0.207				
		(1.74)					(-7.48)	(-1.20)				
Model 3	Medium	0.204***					0.101**	0.048			3.30%	969419
		(3.10)					(2.54)	(0.29)				
	High	0.423***					0.018	0.955***				
		(3.81)					(0.32)	(4.00)				
	Low	0.101					-0.488***	-0.212	-0.430	-0.400		
		(1.10)					(-6.06)	(-1.35)	(-1.56)	(-1.16)		
Model 4	Medium	0.148**					0.054	0.054	0.220	-0.483***	4.40%	904047
		(2.34)					(1.46)	(0.31)	(1.28)	(-4.46)		
	High	0.372***					-0.024	0.962***	0.093	-0.367***		
		(3.62)					(-0.47)	(4.18)	(0.45)	(-3.14)		

We estimate cross sectional regressions for every month from July of 1985 to December of 2009, a total of 300 months. This regression is similar to the one in Table 2.4, except that we sort all firms listed on the three exchanges into three groups based on characteristics and create dummies for the different groups. We then interact those dummies with the explanatory variables in the regression specification. We use group 0 (small firms or low book-to-market) as the base group. The coefficients can then be interpreted as the incremental effect of each explanatory variable in group 1 and group 2 (large firms or high book-to-market) as compared to the base group. The standards errors are calculated as in Fama and MacBeth (1973), but corrected for serial correlation as in Newey and West (1987). Stars denote significance at the 10% (*), 5% (**), and 1% (***) respectively.

$$\mathrm{R_{i,t}} = \mathrm{a_t} + \mathrm{a_{G^1 t}}\,\mathrm{Group1} + \mathrm{a_{G^2 t}}\mathrm{Group2} + \sum_{k=1}^{5} \lambda_{k,t}\beta_{i,t}^{k} + \sum_{k=1}^{5} \lambda_{G^1 k,t}\,(\beta_{i,t}^{k} \times \mathrm{Group1}) + \sum_{k=1}^{5} \lambda_{G^2 k,t}\,(\beta_{i,t}^{k} \times Group2) + \varepsilon_{i,t}$$

Table 2.7: Comparing prices of risk across exchanges for firms less subject to arbitrage constraints

PANEL A: Excluding smallest 10% of firms

		Beta_Mkt	Beta_SMB	Beta_HML	Beta_UMD	Beta_Liq	LN(ME)	LN(BE/ME)	RET(-12,-2)	AMIILL	Adj R-Sq	OBS
Model 1	NYSE	0.335*	0.199**	0.157*								
		(1.88)	(2.27)	(1.76)								
	Amex	0.169	0.018	-0.184							4.30%	930109
		(1.04)	(0.17)	(-1.52)								
	Nasdaq	0.151	0.024	-0.172**								
		(1.53)	(0.29)	(-2.32)								
Model 2	NYSE	0.294*	0.175**	0.192**	0.037	0.226***						
		(1.67)	(2.04)	(2.09)	(0.35)	(2.63)						
	Amex	0.216	0.056	-0.218*	0.073	0.095					5.00%	930109
		(1.31)	(0.50)	(-1.76)	(0.49)	(0.76)						
	Nasdaq	0.179*	0.042	-0.193**	0.004	0.170***						
		(1.81)	(0.50)	(-2.56)	(0.05)	(2.81)						
Model 3	NYSE	0.330**					-0.217***	0.175**				
		(2.14)					(-3.74)	(2.10)				
	Amex	0.087					-0.083	-0.307**			3.70%	880733
		(0.65)					(-1.02)	(-2.44)				
	Nasdaq	0.029					-0.291***	-0.228*				
		(0.32)					(-5.27)	(-1.91)				
Model 4	NYSE	0.130					-0.183***	0.154*	-0.199	0.489***		
		(1.28)					(-3.07)	(1.93)	(-0.44)	(2.86)		
	Amex	0.238*					-0.177**	-0.249**	0.416	-0.758***	5.00%	850849
		(1.72)					(-2.08)	(-1.99)	(0.91)	(-2.62)		
	Nasdaq	0.114					-0.396***	-0.149	0.059	-0.654***		
		(1.32)					(-6.53)	(-1.39)	(0.23)	(-4.15)		

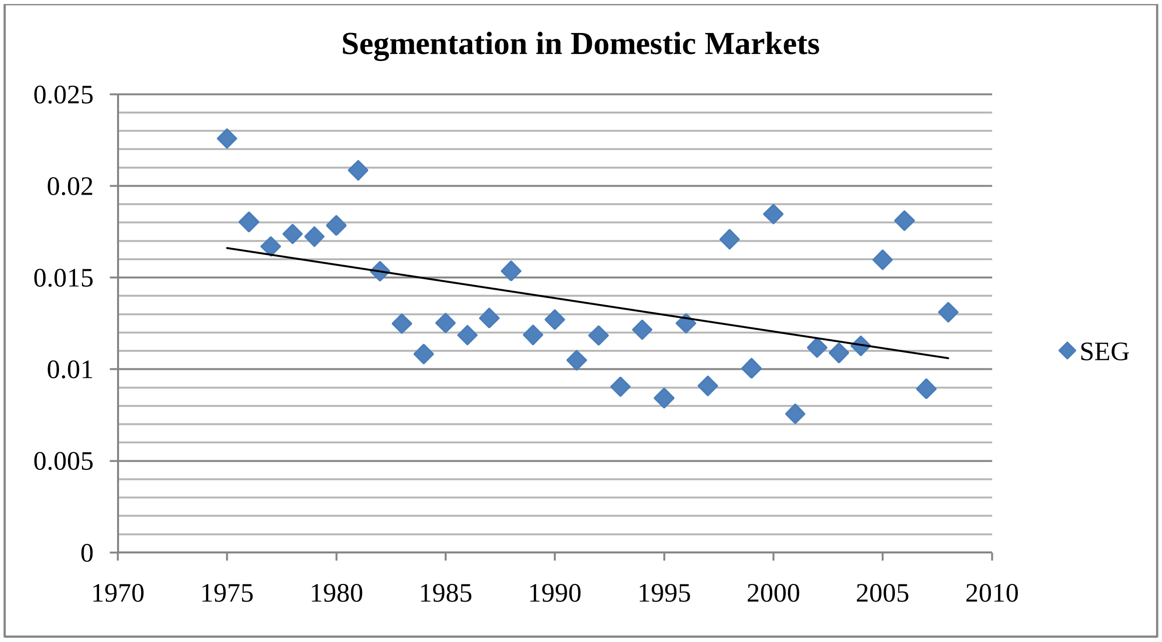
Table 2.7 (cont.)

PANEL B: Excluding 10% of firms with highest lagged IV

		Beta_Mkt	Beta_SMB	Beta_HML	Beta_UMD	Beta_Liq	LN(ME)	LN(BE/ME)	RET(-12,-2)	AMIILL	Adj R-Sq	OBS
	NYSE	0.245	0.180**	0.145*								
		(1.48)	(2.06)	(1.65)								
Model 1	Amex	0.273**	0.070	-0.087							4.00%	938375
		(1.98)	(0.73)	(-0.82)								
	Nasdaq	0.136	0.034	-0.149**								
		(1.35)	(0.38)	(-2.09)								
	NYSE	0.211	0.158*	0.177**	0.029	0.193**						
		(1.30)	(1.83)	(1.97)	(0.28)	(2.37)						
Model 2	Amex	0.285*	0.079	-0.138	0.184	0.420***					4.80%	938375
		(1.97)	(0.83)	(-1.27)	(1.47)	(3.83)						
	Nasdaq	0.173*	0.049	-0.170**	0.112	0.310***						
		(1.72)	(0.56)	(-2.35)	(1.44)	(4.70)						
	NYSE	0.259*					-0.209***	0.134				
		(1.83)					(-3.83)	(1.64)				
Model 3	Amex	0.324***					-0.319***	-0.068			3.60%	881448
		(2.61)					(-4.83)	(-0.57)				
	Nasdaq	0.118					-0.337***	-0.229**				
		(1.44)					(-6.18)	(-2.23)				
	NYSE	0.072					-0.164***	0.115	-0.053	0.538***		
		(0.70)					(-2.90)	(1.44)	(-0.12)	(3.90)		
Model 4	Amex	0.313**					-0.171**	-0.164	0.281	0.077	4.80%	816076
		(2.21)					(-2.25)	(-1.30)	(0.97)	(0.46)		
	Nasdaq	0.170*					-0.419***	-0.158*	0.241	-0.492***		
		(1.87)					(-7.94)	(-1.66)	(0.98)	(-3.87)		

We estimate cross sectional regressions for every month from July of 1985 to December of 2009, a total of 300 months. This regression is similar to the one in the table (Table 2.4), except that we exclude the firms that should have the greatest arbitrage constraints. Panel A excludes the 10% of firms with the lowest MV in the prior month, and Panel B excludes the 10% of firms with the highest idiosyncratic volatility in the prior month. We use NYSE as the base group. The coefficients can then be interpreted as the incremental effect of each explanatory variable on the Amex and Nasdaq exchanges as compared to the NYSE.

$$R_{i,t} = a_t + a_{at}\,\mathrm{Amex} + a_{nt}\mathrm{Nasdaq} + \sum_{k=1}^{5} \lambda_{k,t}\beta_{i,t}^{k} + \sum_{k=1}^{5} \lambda_{ak,t}\,(\beta_{i,t}^{k} \times \mathrm{Amex}) + \sum_{k=1}^{5} \lambda_{nk,t}\,(\beta_{i,t}^{k} \times Nasdaq) + \varepsilon_{i,t}$$



	Intercept	Trend	R-Sq
Parameters	.37607	-0.00018201	23.26%
T-Stat	3.23	-3.11	

Figure 1: Segmentation using non model-specific tests
We estimate segmentation as in Bekaert, Harvey, Lundblad, and Siegel (2011) and apply it to the domestic US markets. Explicitly, we model segmentation as

$$SEG_{i,t} = \sum_{j=1}^{N} IW_{i,j,t} \left| EY_{i,j,t} - EY_{w,j,t} \right|,$$

where $EY_{i,j,t} - EY_{i,j,t}$ is the difference in earnings yield (EY) between an industry (j) in a specific exchange (i) and that of the entire domestic market (w) in every year (t). The segmentation measure (SEG) is the industry weighted (IW) average of the absolute earnings yield differentials. Figure 1 illustrates how the average segmentation (for the three exchanges) has decreased over time for the three US exchanges (NYSE, Amex, and Nasdaq).

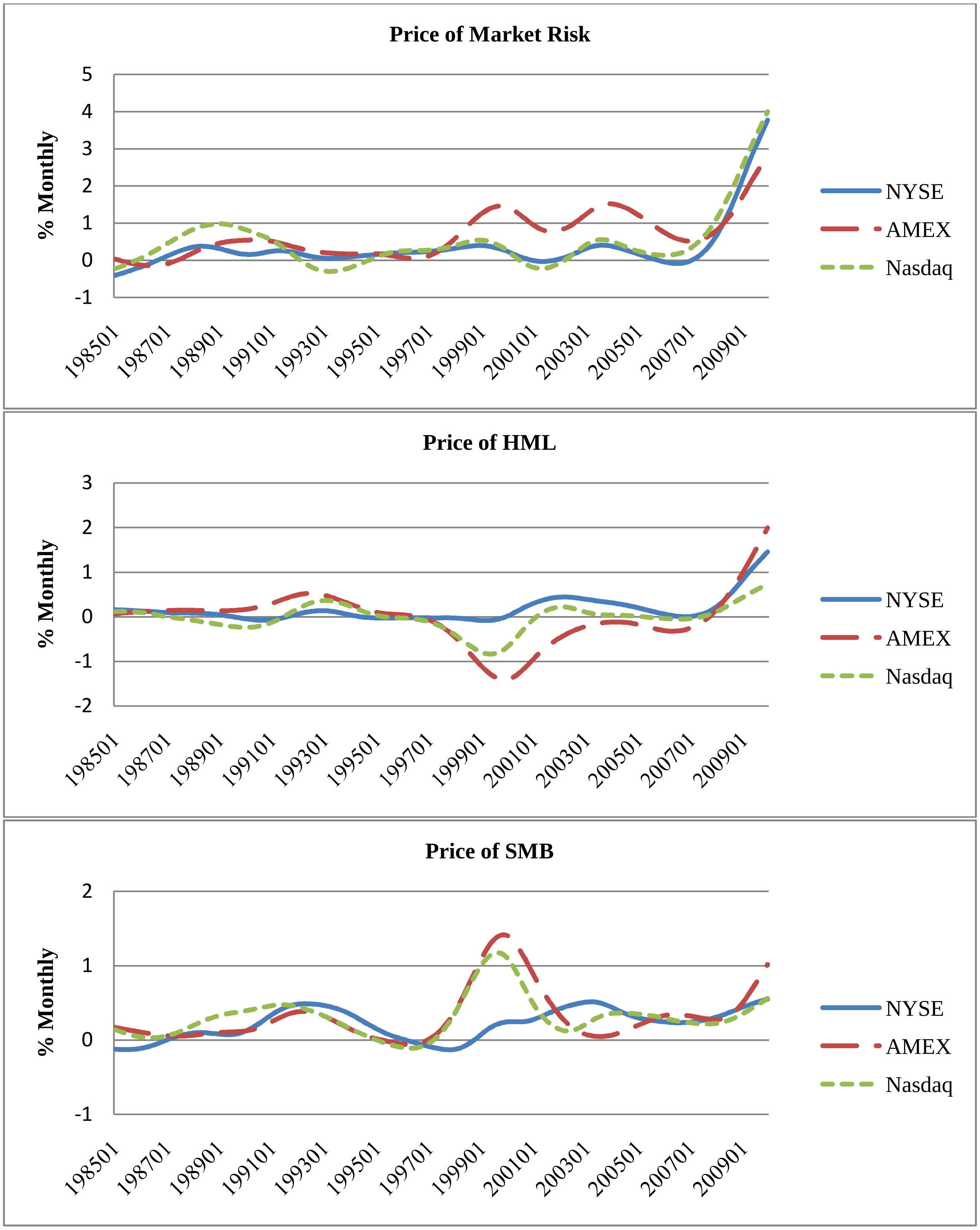


Figure 2: Variation in the prices of systematic risk over time
These graphs shows the prices of risk factors, as estimated monthly using cross sectional regressions of returns on the Fama and French (1993) firm factors loadings beta_mkt, beta_smb, and beta_hml (Model 1, Table 2.3) for the three domestic exchanges NYSE, Amex, and Nasdaq over the period from January of 1985 to December of 2009. The series have been smoothed using HP filters.

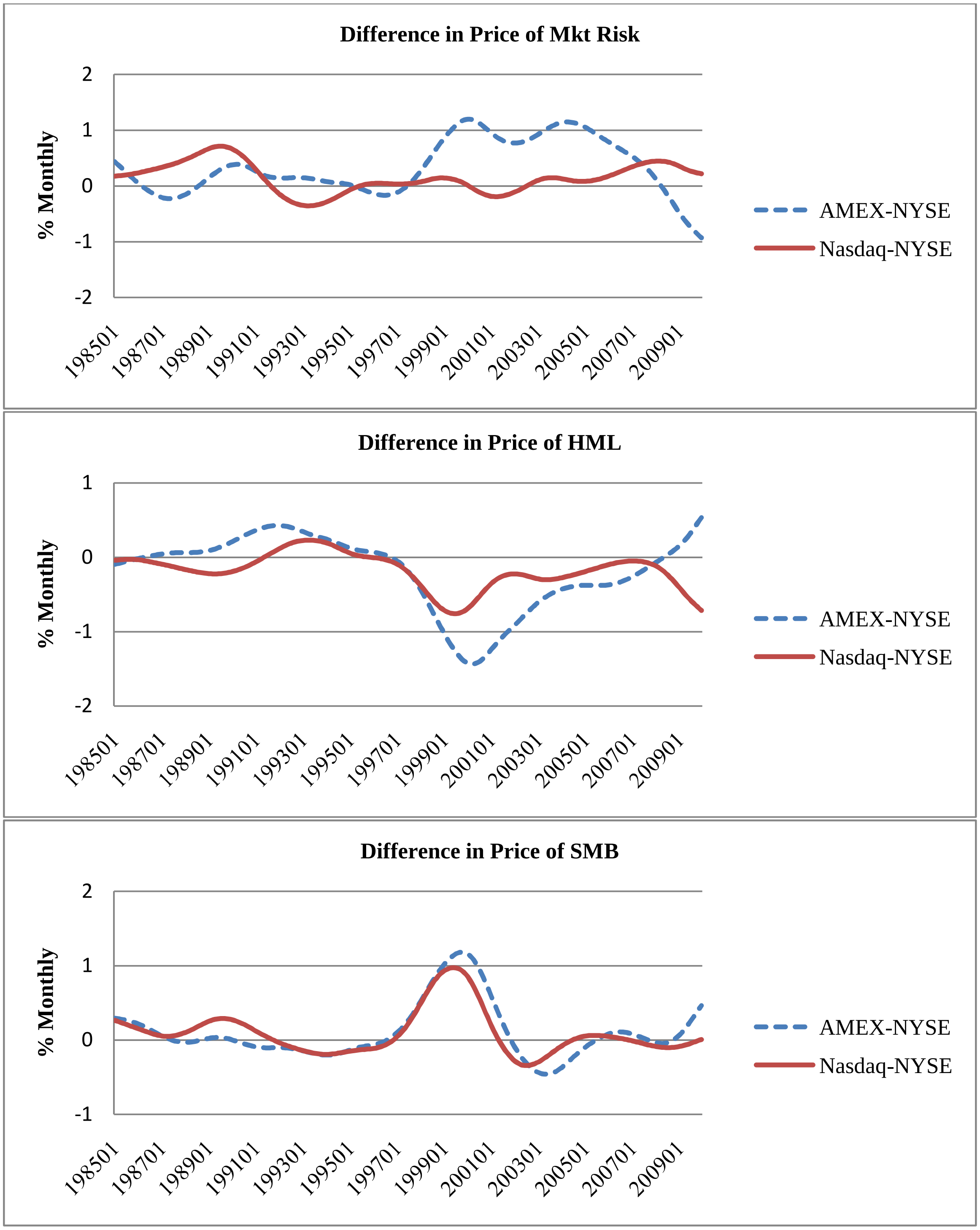


Figure 3: Variation in the differences of systematic risk prices over time
These graphs shows the differences in the prices of risk for the three domestic exchanges over the period from January of 1985 to December of 2009. The series have been smoothed using HP filters.