Does Trading Volume Predict Stock Returns?
In- and out-of-Sample Evidence from the U.S. and International Markets

Zijun Wang

Texas A&M University

(This draft: July 2013)

* Correspondence: Zijun Wang, Private Enterprise Research Center, Allen Building, Room 3028, Texas A&M University, College Station TX, 77843, USA. Tel: 979-845-4440; Fax: 979-845-6636; E-mail: z-wang@neo.tamu.edu.

Does Trading Volume Predict Stock Returns?
In- and out-of-Sample Evidence from the U.S. and International Markets

Abstract

This paper provides a comprehensive examination of the dynamic relation between trading volume and stock returns by combining three lines of empirical research that feature prominently in the literature: Granger causality test, out-of-sample stock return predictability, and the high volume return premium in the cross section. Evidence from the U.S. suggests that higher returns do follow more intensive trading, especially in the pre-2000 period. However, the ex-ante predictability delivers a small economic gain equivalent to an annual return of 0.73% for a risk-averse investor. This weak out-of-sample predictive power of volume is absent in most of the other major markets.

Keywords: Volume-return relation; Granger causality test; Out-of-sample regression; High volume return premium

JEL number: G12; G15

1. Introduction

The relation between trading volume and stock returns has been an active area of research for many decades. The popularization of high-speed (high-frequency) trading, a conspicuous aspect of financial markets in the last ten years, has attracted increasing attention to the relation from both academicians and practitioners. Not surprisingly, it also figures prominently in the debate about the revived proposals to impose Tobin-type securities transaction taxes to reduce trading volume following the recent financial crisis. Better understanding of the volume-return relation clearly can also shed light on the ongoing debate about whether modern finance is too big (Cochrane, 2013; Greenwood and Scharfstein, 2013).

The questions at issue are whether there is any relation between trading volume and stock returns and, if the answer is yes, whether such a relation is economically significant.[1] The latter is probably more important in the current debate. Market microstructure theory suggests that both trading volume and price changes (returns) are related to the arrival of information to the market. Thus volume and price movement may jointly depend on the intensity of information flow. Much of the early work on the volume-return relation therefore focuses primarily on the contemporaneous relation between volume and price changes (Karpoff, 1987; Gallant, Rossi, and Tauchen, 1992). However, considering the long-standing controversy about the simultaneous determination of price and quantity in economics, it is not surprising that the contemporaneous causality between volume and stock returns has proven very difficult to sort out empirically given the observational nature of data. As a result, more recent empirical research has shifted towards investigating the perhaps less ambiguous, yet more fruitful, dynamic relation between

[1] Like many other papers that study the volume-return relation, here we are interested in whether trading volume Granger-causes stock returns. Also see the interesting papers by Griffin, Harris, and Topaloglu (2003), Statman, Thorley, and Vorkink (2006), and Griffin, Nardari, and Stulz (2007) for complementing evidence on whether investors trade more when the market has performed better (which may be called the return-volume relation).

volume and stock returns via testing Granger (non-) causality (Hiemstra and Jones, 1994). The intent is to determine whether including past volume information can help predict stock returns after controlling for past returns and other relevant information.

As is well known, Granger causality is not causality in a more fundamental sense because of the possibility of omitted common factors. Therefore, even if volume is found to Granger cause stock returns, it does not necessarily imply that investors can arbitrarily trade more (or manipulate volume, to use a jargon from the causality literature) to earn more out of the market.[2] Nevertheless, lack of Granger causality could be an indicator of the absence of an underlying causal relation between volume and returns. The study of Granger causality is thus an important step towards understanding the underlying causal relation from the operational perspective. Furthermore, focusing on the dynamic relation between volume and returns, Granger causality is perhaps more informative than the often elusive contemporaneous causality, as far as prediction and risk management are concerned.

Against this background, we now state our three intended contributions that extend empirical research on the volume-return relation. First, we conduct the Granger causality test in both in- and out-of-sample contexts to provide a more comprehensive picture about the volume-return relation that in-sample evidence alone cannot. Currently, the Granger causality test is implemented using in-sample regressions in most empirical studies, whether or not the concept of Granger causality itself is invoked (e.g., Hiemstra and Jones, 1994; Easley, O'Hara, and Srinivas, 1998; Chordia and Swaminathan, 2000; Malcolm and Stein, 2004). As noted by Granger (1969), and more rigorously documented in Ashley, Granger and Schmalensee (1980), a causality test based on forecasting performance brings the maximum amount of information to

[2] In fact, Odean (1999) and Barber et al. (2009) find that aggressive trading by individual investors results in significant losses, although both aggressive and passive trades of institutions are profitable in the Taiwanese market.

bear on the Granger causation hypothesis. Ashley, Granger and Schmalensee (1980) explicitly stated that “a sound and natural approach to such tests [Granger causality tests] must rely primarily on the out-of-sample forecasting performance of models relating the original (non-prewhitened) series of interest” (p. 1149). Unfortunately, this important insight has been largely ignored in the literature.[3] To the best of our knowledge, applying out-of-sample tests for inferring Granger causal inference on the volume and return relationship is rare, if it has been done at all.

In contrast, the literature on stock return predictability using out-of-sample regressions is growing fast (e.g., Welch and Goyal (2008) and Rapach, Strauss, Zhou (2010), to name two recent examples). This is not surprising given that out-of-sample return predictability is economically important in itself from an asset allocation perspective. More importantly, in practice, many commonly used variables have been found to have no or negligible out-of-sample forecasting ability despite their enormous in-sample predictive power. The lack of robustness of the in-sample evidence casts doubt on the real predictive power of these variables. Conceivably, if both in- and out-of-sample evidences derived from a model are consistent, then the empirical model is less likely to be misspecified and the theory on which the model is based is more likely to be credible.[4] Finally, as is evident in the later sections, our emphasis on out-of-sample Granger tests for the volume-return relation is very closely linked to this large literature on out-of-sample return predictability, where the predictive variable of interest is trading volume.

Our second contribution is that we investigate the volume-return relation using both direct aggregate time series estimates and the high volume return premium, which summarizes

[3] See Inoue, A., and L. Kilian (2005) and other papers it has aspired for an in-depth discussion about the pros and cons of in-sample vs. out-of-sample tests.
[4] Rapach, Strauss, Zhou (forthcoming) is a recent example of performing Granger causality tests using both in- and out-of-sample regressions. Those authors’ interests are whether and how U.S. stock returns lead markets in other industrialized countries.

cross-sectional return predictability based on volume. Empirical studies typically either concentrate on time series of market/index volume and return, or examine the cross-sectional variation using data on individual stocks.[5] By combining evidence from both time series and cross-sectional analysis, we hope to present a full picture "that relates the variation through time in expected return to models for the cross section of expected returns" (Fama, 1991). The high volume return premium (or simply high volume premium, HVP), defined here as the return on a zero-cost portfolio that is long on stocks experiencing unusually high trading volume and short on low-volume stocks, has been studied before in regard to the interplay of short-run return autocorrelations and volume at the firm level (Chordia and Swaminathan, 2000; Gervais, Kaniel, and Mingelgrin, 2001; Kaniel, Ozoguz, Starks, 2012). Our contribution here is that we formally explore, again within the Granger-causality test framework, whether this cross-sectionally constructed variable has time series predictive power for future returns both in and out of samples.

While there is a large literature on the cross-sectional return predictability, such as the value premium and return momentum, careful empirical studies on the real out-of-sample forecasting ability of these factors are rare. The premise is that if these return anomalies proxy for investment opportunities (hence can be interpreted as the hedging component in Merton's (1973) intertemporal CAPM), they should help predict future returns.[6] Using the portfolio approach-based high volume return premium also offers an advantage over the raw trading volume in studying the volume-return relation. As we show later, the measurement of the volume premium is straightforward and consistent over time. The return premium is stationary by

[5] Malcolm and Stein (2004) appear to be an exception.
[6] For example, utilizing both price and volume information, Lo and Wang's (2006) two-factor ICAPM model predicts that, among returns of all portfolios, current return on the hedging portfolio (the hedge factor) gives the best forecast for future return on the market portfolio (Proposition 2).

construction. In contrast, total trading volume contains trends and needs to be transformed prior to use in regressions, either by taking log-differences as in Hiemstra and Jones (1994) or by using log-detrended turnover as in Lo and Wang (2000).

Thirdly, like many other empirical studies on the volume-return relation, we concentrate on the U.S. market. However, to mitigate the concerns of data mining and low power of time series tests, we also follow Fama and French’s (1998) and Griffin’s (2002) investigation of the value premium and investigate major international equity markets. We examine evidence from six other developed countries in the Group of Seven (G-7) and an additional 12 countries of both developed and emerging economies which house large stock exchanges by market capitalization. The advantage of performing the same tests for many countries is that empirical findings may no longer be sample-specific. Put another way, because the stock valuation processes vary across countries due to differences in institutions and information flows, they should provide some degrees of independence in tests.

In a notable recent study, Kaniel, Ozoguz, and Starks (2012) provide a thorough documentation of the high volume premium up to June 2001 for 41 international markets. Extending their work, we first offer updated estimates of the volume effect on stock returns for the above-mentioned 18 major international markets by including the most recent decade of data. We then perform in- and out-of-sample Granger causality tests to evaluate the time series predictive power of both aggregate volume and the high volume premium for returns in each stock market. This ten-year sample period is particularly interesting for studying the volume-return relation because it witnessed accelerating changes of the market microstructure resulting from high-frequency trading. These fundamental changes have greatly improved the way liquidity is provided in both the U.S. and many other markets. This period also features one of the most severe global financial crises in the history.

Our main findings can be summarized as follows. First, for U.S. data from 1963 to 2010, both aggregate trading volume (detrended turnover is used in later implementation) and the high volume premium show statistically significant in-sample predictive power for stock returns. The effect is much stronger in equal-weighted portfolios than in value-weighted ones. Out-of-sample regressions show mixed evidence of incorporating past volume information to predict returns. There is no gain in forecast accuracy for value-weighted portfolios. In contrast, compared to simple autoregression forecasts and historical average forecasts, models also using trading volume generate better forecasts for stock returns to equal-weighted portfolios by both root mean squared forecast errors and a forecast encompassing test. Nevertheless, for a typical risk-averse investor, the improvement in forecasts transforms to a mere utility gain of 0.73% per annum in terms of return rates. Interestingly, this estimate is close to French's (2008) estimate of 0.67% of the aggregate value of the market each year investors spend searching for superior returns over 1980 to 2006.

Further sub-sample analysis shows that trading volume's out-of-sample forecasting ability declines significantly over the last ten years.[7] One possible explanation for the weaker predictive power is that rapid growth of trading volume has been associated with cost saving and gains in market efficiency in the form of narrowing bid-ask price spreads observed during this period. Alternatively, the declining time series return predictability is also potentially consistent with recent trends in trading activity in which turnover has become more sensitive to return

[7] Interestingly, Griffin, Nardari, and Stulz's (2007) also find that the return-volume relation weakens in recent years for U.S. and some high-income countries.

predictors and increased trading by institutions has been accompanied by decreased cross-sectional return predictability (Chordia, Roll, and Subrahmanyam, 2011).

Our second empirical finding is derived from the 18 international markets. Specifically, the in-sample Granger causality test based on international data provides less consistent evidence than in the U.S. market in support of the predictive power of trading volume. The causal relation is found significant only in a few markets and depends on which volume variable (aggregate turnover or the volume premium) is used and whether portfolios are value-weighted or equal-weighted. This part of analysis also demonstrates a marked difference between ex post and ex ante predictive power of trading volume. Canada is the only other G-7 country than the U.S. in which the high volume premium helps predict returns out of sample for equal-weighted portfolios by all three statistical and economic measures. Value-weighted, aggregate turnover shows out-of-sample predictive power in five out of the other 12 major markets. The high volume premium contains some additional predictive power for value-weighted portfolios only in India, and for equal-weighted portfolios only in China.

We also find that trading volume in the U.S. market in general does not contain additional information for forecasting returns in other markets after controlling for past returns, volume, and volatility information from domestic markets as well as past U.S. market returns. Because many international samples comprise predominantly recent data, the absence of a spillover effect of U.S. trading activity is consistent with its vanishing predictive power for U.S. market returns in the second sub-sample analysis.

As briefly reviewed in the next section, a popular hypothesis in the literature on the volume-return relation is that markets are not perfectly efficient in the sense that the process by which prices adjust to information is not immediate. Therefore, market statistics such as volume impound information that is not yet incorporated into the current market price. Our finding of a quantitatively small but statistically significant return forecastability by volume is consistent with this hypothesis and confirms the existence of a dynamic volume-return relation. However, our support for the published theoretical and empirical work is limited because we find that the out-of-sample predictive power of trading volume largely disappeared in the U.S. market in the recent period and that the evidence is scarce in international markets. Overall, it appears reasonable to conclude that investors cannot gain much financially by "riding the volume curve," at least at the levels of net profits suggested by our findings.

2. Review of the Literature

Whereas the process by which trading brings information to markets remains a bit of a mystery (Cochrane, 2013), several prominent pricing models have long been proposed that help describe the volume-return relation among stocks. In Campbell, Grossman, and Wang's (1993) model, investors are heterogeneous but the market is characterized by symmetric information. The model predicts that price changes accompanied by high volume will tend to be reversed which is less true for prices changes on days with low volume. Wang (1994) develops an equilibrium model of stock trading in which investors have rational expectations but are heterogeneous in their information and private investment opportunities. Therefore, trading volume is always positively correlated with absolute price changes, with the association increasing with information asymmetry. Blume, Easley, and O'Hara (1994) develop an equilibrium model in which aggregate supply is fixed and traders receive signals with differing quality. They show how volume, information precision, and price movements are related to each other. Malcolm and Stein (2004) build a behavior model that seeks to explain why increases in liquidity like spreads, depth, and trading volume, carry information about the extent to which irrational traders are influential in the market, and hence about expected returns at both firm and market levels. Finally, Schneider (2009) constructs a model that offers a closed-form solution of a rational expectations equilibrium where price and volume can be studied using a contemporaneous data approach.

The above theoretical studies have been accompanied by a growing number of empirical studies investigating the volume-return relation and exploring the relationship between trading volume and stock return predictability at short horizons. Karpoff (1987), Schwert (1989), and Gallant, Rossi, and Tauchen (1992) all document a positive correlation between volume and absolute price changes. Chordia and Swaminathan (2000) find that daily or weekly returns of stocks with high trading volume lead returns of stocks with low trading volume, which, they argue, is related to the tendency of high volume stocks to respond rapidly and low volume stocks to respond slowly to marketwide information. While many papers focus on aggregate (index) volume and returns, there has been increasing interest in studying individual stocks and analyzing the dynamic relation between volume and returns in the cross section. A partial list includes Conrad, Hameed, and Niden (1994), Brennan, Chordia, and Subrahmanyam (1998), Cooper (1999), Gervais, Kaniel, and Mingelgrin (2001), Llorente, et al. (2002), Chordia, Huh, and Subrahmanyam (2007), and Kaniel, Ozoguz, and Starks (2012).

3. Econometric Methodology

As discussed earlier, rather than seeking to establish a relation between trading volume and stock returns with a causal interpretation in the strict sense, we approach the issue in a less ambitious manner by considering the forecasting relation between the two variables using the popular framework of Granger causality (Granger, 1969). Specifically, if event X is the cause of another event Y, then the event X should precede the event Y. Therefore, past information on X should help predict the current value of Y. Nevertheless, there are well known reasons why it may be misleading to infer fundamental causality from Granger causality. X Granger-causing Y may simply reflect forward-looking behavior in X and does not mean X *causes* Y. In practice, the spurious causal relation may also result from omission of other important causal variables in the test. On the other hand, if X does *cause* Y, Granger-causality tests may not detect the causal relationship due to the issue of data aggregation. For example, if the causal relation exists between X and Y on the hourly basis, we may fail to detect this causal relation using daily data because of feedback effects. With these caveats in mind, we now proceed to outline the empirical steps we use to implement the popular Granger causality test. Consider the following time series model for Y,

$$y_t = \alpha + \sum_{m=1}^{L_y} \beta_m y_{t-m} + \sum_{n=1}^{L_x} \gamma_n x_{t-n} + \sum_{r=1}^{L_z} \lambda_r z_{t-r} + \varepsilon_t,\ t = 1, ..., T, \qquad (1)$$

where y_t, x_t, and z_t are the realized values of Y, X, and Z at time period t, y_{t-m}, x_{t-n}, and z_{t-r} are the corresponding m-, n-, and r-period lagged values, ε_t is the error term, L_y, L_x, and L_z are the numbers of lags on the three variables, and α, β_m, γ_n, and λ_r are free parameters. To test if X Granger-causes Y, one should examine whether the inclusion of past values of X helps improve predictability of the current value of Y, given the inclusion of other relevant series (Z could be a vector of variables). To this end, we form the following hypothesis in terms of parameter restrictions on Model (1):

H_0: $\gamma_n = 0$ for all n =1, 2, …, L_x vs. H_A: $\gamma_n \neq 0$ for some n.

In general this can be tested by constructing a Wald-type statistic which has a limiting distribution of χ^2 with L_x degrees of freedom under the null. In particular, if only one lagged variable of X is used ($L_x = 1$), then the standard t-test can be applied.

In our later application, Y is daily return to the market portfolio, X is a variable measuring trading activity (aggregate turnover or the high volume premium, to be specific), and Z is market volatility. Note that the lagged returns are included in Model (1) to capture generally small but statistically significant serial correlation resulting from, among other things, non-synchronous trading in the daily returns (e.g., Conrad and Kaul, 1988; Lo and MacKinlay, 1988). We thus rule out the possibility that the seemingly predictive power of volume simply reveals the well-documented autoregression in the return series. Model (1) could be motivated by the market microstructure literature, which explicitly takes the sequential nature of the trading process into account. For example, in Blume, Easley, and O'Hara's (1994) learning model, volume provides information on information quality that cannot be deduced from the price statistic, and traders who use information contained in price and volume statistics do better than traders who do not. Models similar to (1) have also been used by Gallant, Rossi, and Tauchen (1992) and Campbell, Grossman, and Wang (1993), among many others, in their empirical analysis.

Most studies that examine the contemporaneous behavior of volume and *absolute* price changes have documented a positive correlation between the two. Research focusing on the intraday patterns has found similar results. However, the exact signs of coefficients (γ's) for the lagged volume in Model (1) are less clear-cut. According to Wang (1994), uninformed investors trade against informed investors' private information. They also trade to revise their positions as the true state of the economy is revealed. These two components in the trading each lead to a different dynamic relation between volume and return. A high return accompanied by high volume implies high future returns if the first component dominates and low future returns if the second component dominates. Llorente et al. (2002) offer similar predictions that the relation of current return, volume, and future returns depends on the relative significance of speculative trade versus hedging trade.

As noted earlier, previous studies in this line of research focus on in-sample Granger causality tests. However, as pointed out by Corradi and Swanson (2006), the choice of out-of-sample versus in-sample Granger causality tests can have a dramatic impact on statistical inference. Here we complement the existing evidence by also performing out-of-sample tests to infer Granger causal inference, which may be more in the spirit of Granger' (1969) original definition of causality. The implementation of the out-of-sample Granger causality is straightforward. Suppose there are a total of T observations. We estimate Model (1) under both the null and the alternative hypotheses using the first R (in-sample) observations. We then generate one-step-ahead recursive forecasts of y_t for the remaining P (out-of-sample) observations ($R + P = T$) from both the H_0 and the H_A models. Denote the corresponding forecast error series as $\hat{\varepsilon}_{0,t}$ and $\hat{\varepsilon}_{A,t}$, respectively. If the H_A model produces more accurate forecasts than H_0, or equivalently, if $\hat{\varepsilon}_{A,t}$ is smaller than $\hat{\varepsilon}_{0,t}$, then X Granger causes Y in an out-of-sample sense.

Determining whether there exists a statistically significant difference in forecasting accuracy between the H_A and H_0 models is a fundamental aspect of the out-of-sample Granger causality test. It is possible that, although two sets of forecasts are visually (in)different from each other, they may (not) differ statistically due to sample variability. This can be a problem for studying daily stock returns, a large component of which are unexplained even at the aggregate level. Since the models we consider above are nested and they all may be misspecified, in this study we follow Corradi and Swanson (2006)'s recommendation to apply Clark and McCracken's (2001) encompassing test (ENC-NEW test). Under the null hypothesis that the H_0

model forecast encompasses that of the H_A model, Clark and McCracken (2001) show that the following statistic has an asymptotic nonstandard distribution

$$\text{ENC}-\text{NEW} = P\frac{P^{-1}\sum_{t=(R+1)}^{T}(\hat{\varepsilon}_{0,t}^{2} - \hat{\varepsilon}_{0,t}\hat{\varepsilon}_{A,t})}{P^{-1}\sum_{t=(R+1)}^{T}\hat{\varepsilon}_{A,t}^{2}} \rightarrow_{d} \Gamma_1, \tag{2}$$

where $\Gamma_1 = \int_{\lambda}^{1} s^{-1}B'(s)$, $\lambda = (1+\pi)^{-1}$, π is the limit of P/R, the ratio of the out-of-sample size over the in-sample size, and $B(s)$ is a vector Brownian motion whose dimension equals that of x_t (namely L_x). If H_0 forecasts encompass Model H_A, then H_A forecasts do not provide useful information absent from forecasts from H_0. The encompassing test (2) has seen increasing application in finance (e.g., Butler, Grullon, and Weston, 2005; Welch and Goyal, 2008).

Like other statistical measures for forecast evaluations, the root-mean squared forecast errors and the encompassing test do not explicitly account for the risk borne by an investor in following portfolio recommendations from statistically preferred models (Rapach, Strauss, and Zhou, 2010). To address this limitation, in this paper we also provide evidence of economic significance of forecasting ability of trading activity for market returns. Specifically, following Marquering and Verbeek (2004), Campbell and Thompson (2008), Welch and Goyal (2008), and more recently Rapach, Strauss, Zhou (2010), we calculate realized utility for a mean-variance investor on a real-time basis for the out-of-sample period. Specifically, we assume that the investor allocates her/his investment daily between stocks and risk-free bills. The standard portfolio allocation rule then stipulates that, conditional on information available at period t, the optimal weight of such an investor's portfolio on stocks at period (t + 1) is

$$s_{i,t+1} = (\frac{1}{\gamma})\frac{E_t(R_{i,t+1})}{E_t(\sigma_{i,t+1}^{2})}, \tag{3}$$

where γ is the investor's relative risk aversion parameter, $E_t(R_{i,t+1})$ and $E_t(\sigma^2_{i,t+1})$ are the forecast for stock return and its variance based on forecasting model *i*. Correspondingly, the rest of the portfolio $(1 - s_{i,t+1})$ is invested in the risk-free bills. The realized average utility of the investor is given by

$$\hat{u}_i = \hat{\mu}_{i,p} - \frac{1}{2}\gamma\hat{\sigma}^2_{i,p}, \tag{4}$$

where $\hat{\mu}_{i,p}$ and $\hat{\sigma}^2_{i,p}$ are the out-of-sample mean and variance of the returns to the dynamic portfolio formed based on the above rule. Intuitively, the investor's utility increases with the average return but decreases with its volatility.

4. Data

We study stock return data from both U.S. and international markets. In this section, we describe the U.S. sample in detail. The international data are described briefly and analyzed in Section 6. For the U.S. sample, we consider all NYSE, Amex, and NASDAQ non-financial stocks for the period of July 1, 1963 through December 31, 2010 with July 1962-June 1963 reserved for pre-sample selection. They are obtained from CRSP monthly and daily stock securities files and events files.[8]

Volume can change simply because of (reverse) stock splits. Following Lo and Wang (2000), Griffin, Nardari, and Stulz (2007), and many other studies in the literature, we use turnover rather than raw volume data to measure trading intensity. We consider two types of aggregate time series estimates of turnover, one that is weighted by market capitalization of

[8] Following the literature, we use daily data. The market microstructure literature also studies the relations between trading volume, stock returns, and market volatility. However, as Andersen (1996) points out, the focus of this area of research typically is on intraday rather than interday dynamics as we study here. Therefore, its theoretical predictions regarding the relations among these variables may not hold at the daily frequency due to the complicating effect of temporal aggregation on causality testing (Granger, 1988).

stocks (VWVOL) and the other that is equally weighted (EWVOL).[9] We exclude observations with missing price or volume data and stocks with less than one year of trading history. Also discarded are those delisted from the exchanges within one year due to merger, (partial) liquidation, and other capital events.[10] Figure 1 plots 100-day moving averages of both value-weighted and equal-weighted measures of volume (turnover), which clearly show an upward trend in both measures of aggregate trading activity over the sample period. In particular, turnover increased significantly starting early 2003 and appears to have tailed off by the end of the sample, largely coinciding with the development of high-frequency trading (e.g., Chordia, Roll, and Subrahmanyam 2011). As pointed out by Griffin, Nardari, and Stulz (2007), turnover may be influenced by trends in bid-ask spreads, commissions, availability of information, and other factors that might contribute to the general increase in trading activity through time. To remove this slowing moving average component, in all later analysis, we de-trend turnover by first taking its natural log and then subtracting its 100-trading-day (20 calendar weeks) moving average (see, for example, Chen, Hong, and Stein (2001) for a similar treatment).

Panel A of Table 1 provides descriptive statistics of the value- and equal-weighted log-detrended trading volume (turnover) (VWVOL and EWVOL), and the corresponding excess returns on the market portfolio without dividends (VWMKT and EWMKT). The averages of the (detrended) volumes are positive for both measures because of the generally upward trend in the raw series. They are also serially correlated with first-order autocorrelations of 0.60 and 0.73, respectively. The value-weighted market portfolio returns are averaged at 0.010% on a daily

[9] For U.S. data, turnover is defined as trading volume divided by the number of outstanding shares (multiplied by 1,000 for presentation). For international data from Datastream, it is defined as traded value (price times volume) scaled by market capitalization. For convenience of exposition, we sometimes also refer turnover simply as volume throughout the paper.

[10] This is to follow Thornton and Valente (Forthcoming) in a partial attempt to control for possible reverse causality that investors anticipating better/worse future stock performance could be more likely to trade.

basis and the equal-weighted returns at 0.061%.[11] Both return series also feature statistically significant serial correlation which is more evident in EWMKT than in VWMKT. The two measures of trading volume are contemporaneously correlated to the market returns (the details are not reported in the table).

To be comparable to popular return anomalies such as the value premium (HML), we slightly modify the construction of the high volume return premium (HVP) as implemented by Gervais, Kaniel, and Mingelgrin (2001) and Kaniel, Ozoguz, Starks (2012). We set the last trading day of each month as the portfolio formation period and define a stock as a low- (high-) volume stock if its trading volume on the one-day formation period is among the lowest (highest) ten percent out of its 50 daily volumes prior to the formation period (inclusive). We eliminate stocks for which the price or volume data are missing on the portfolio formation day. Stocks which are not traded for nine or more days or whose prices fall below $5 out of the 50 trading days are also removed from the sample to alleviate the microstructure concerns associated with these securities.

We exclude the stocks with less than one year of trading history to mitigate backfilling biases, and those delisted from the exchanges one year prior to the formation date. We also delete observations with an earnings or dividend announcement during a three-day window around the formation date because the volume-return relation during announcement periods may be different than in non-announcement periods (e.g., Kandel and Pearson, 1995). The portfolios are rebalanced monthly by sorting all remaining stocks into ten low- and high-volume portfolios based on their volume classification at the end of each month (t). We then compute both value- and equal-weighted returns for each of the ten portfolios for all trading days in month (t + 1).

[11] The averages of the two corresponding CRSP market portfolio returns without dividends are 0.009% and 0.052%. The correlations between our market portfolio returns and those of CRSP are 0.994 and 0.993, respectively, for the value and equal-weighted estimates.

The value-weighted high volume return premium (VWHVP) is the difference between the value-weighted portfolio return on the highest volume decile and the return on the lowest volume decile. The equal-weighted high volume return premium (EWHVP) is similarly defined.[12]

Rows 5 and 6 of Panel A provide descriptive statistics of the daily value- and equal-weighted high volume premiums (VWHVP and EWHVP), which are estimated using the sample period of July 1963 through December 2010. The daily average of the value-weighted HVP is 0.027%, which is slightly higher than the CRSP value-weighted market portfolio returns of 0.021%. The bottom line of Panel A shows that, not surprisingly, the alternative equal-weighted volume premium is more than twice as high as the value-weighted one (0.061%).[13]

In empirical studies, many return anomalies, such as the value and momentum premiums, are typically quoted on a monthly basis. Therefore, we report in Panel B of Table 1 monthly HVP estimates for the full sample as well as two sub-sample periods. As a comparison, we also include the value premium estimates defined as the difference between the highest and lowest deciles of portfolio returns based on a one-dimensional sort on the book/market ratio.[14] The average monthly value-weighted HVP is a statistically significant value of 0.57%, which is close to the value-weighted value premium of 0.54%. The estimate for the equal-weighted HVP is 1.29%, reasonably close to the 20-day holding period returns of 1.12% for similarly constructed portfolios by Kaniel, Ozoguz, and Starks (2012, Table 2).

[12] This volume classification follows from Gervais, Kaniel, and Mingelgrin (2001) and Kaniel, Ozoguz, and Starks (2012), although our methodology does differ from these two studies in the way portfolios' formation period and test period are chosen.

[13] Following Kaniel, Ozoguz, and Starks (2012), we also find that the high volume premiums are similar to the reported ones if a stock is eliminated from the portfolio if its price falls in the lowest 5 percent of the market during the 49-day reference period. The average premiums are 0.029% and 0.075% for the value- and equal- weighted high volume portfolios, respectively. Both are higher than their counterparts in Table 1. The in-sample and out-of-sample Granger causality test results based on these estimates are also very close to those benchmarks reported later in tables 2-4.

[14] The ten book/market portfolios are obtained from Kenneth French's personal website http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html.

To provide an intuitive description of their dynamics, we plot in Figure 2 the two premium series where NBER-dated recessions are shaded in gray to illustrate cyclical movements in them. Two points are worth pointing out. First, the volume premium edged lower in the last decade. The averages of value- and equal-weighted HVP are 0.62% and 1.34% from July 1963 to December 1999 (Panel B of Table 1). The estimates decreased by 30.42% and 15.74% to 0.43% and 1.13%, respectively, for VWHVP and EWHVP in the period January 2000-December 2010. The differential changes imply that the decrease in the high volume premium is more significant for large stocks than for small stocks. Second, the graph suggests a significant countercyclical pattern in the volume premium. The premium tends to be high around macroeconomic downturns, which is particularly evident around two recessions, one in the period of December 1973-March 1975 and the other in April 2001-November 2001. Numerically, the average value-weighted premium is 1.62% during economic recessions, much higher than the average of 0.34% in the expansion stages of the economy. The difference is less striking for the equal-weighted portfolios, 1.82% vs. 1.03%, suggesting that the volume premium is more countercyclical for large stocks.[15]

5. Empirical Results

In-sample Granger noncausality test

As a preliminary but intuitive way to evaluate the volume-return relation, we first sort all daily market portfolio returns into ten groups with equal numbers of observations based on the

[15] The countercyclical behavior means that the volume premium is high exactly when the risk price is also high during recessions. This piece of new evidence adds to the understanding of the high volume premium itself. Namely, the premium might be related to economic risks. Gervais, Kaniel, and Mingelgrin (2001) argue that the volume premium is consistent with the stock's visibility hypothesis and cannot be explained by market risk. Nevertheless, Rouwenhorst (1999) finds in emerging markets that stocks with high beta, small market cap, high past medium-term return, or high book-to-market have higher average turnover than stocks with low beta, large market cap, poor past performance, or low book-to-market. Rouwenhorst's (1999) evidence suggests that the volume premium may be linked to the Fama-French factors and the momentum factor.

one-period lagged trading volume. Figure 3 plots the simple average returns of each of these ten groups from low to high. There is no clear lead-lag relationship between trading volume and subsequent stock returns for the value-weighted portfolios, possibly with the exception of the two high-volume ones. In contrast, the positive volume-return relation is nearly monotonic among the ten equally-weighted portfolios. Quantitatively, stocks in the lowest-volume decile on average have a next-day return of −0.086%, and those in the highest-volume decile have a returm of 0.360%. The spread is a statistically significant 44.6 basis points.

To control for the small serial correlation in daily market returns and factors other than trading volume which may also affect returns, we complement the evidence in Figure 3 with the regression results from in-sample Granger causality tests using various specifications of Model 1. The regression results are summarized in Table 2, where the first three columns give the selected lag orders of the predictive variables. The next three columns report the point estimate, the *t*-statistic, and the heteroskedasticity-corrected *t*-statistic (*HC*-*t*-statistic) for coefficient γ(s) associated with the lagged trading volume terms. The last two columns of the table report the regression adjusted-R^2 statistic and another popular model fitness measure, the Schwarz's Bayesian information criterion (BIC). Trading volume is proxied by aggregate turnover in Panels A-F, and by the volume premium in Panels G-J.

We first test whether lagged values of trading volume have additional predictive power for excess stock returns while controlling for past returns. To determine the lag orders L_y and L_x in Model (1), we use a model selection approach via BIC. For the daily sample, we assume that the maximum lag order is 22 (approximately the number of trading days in a month). Row 1 of Table 2 summarizes the test results on whether the past trading volume of value-weighted market portfolio (VWVOL) helps predict current daily returns (VWMKT). The optimal model contains one-and two-day lagged returns and one-day lagged volume ($L_y = 2$ and $L_x = 1$). The coefficient for VWVOL (γ_1) is 0.119×10^{-2}, implying that a one-standard deviation increase in trading volume today (which is 0.217 from the sample) predicts a higher return of 0.026% tomorrow, *ceteris paribus*. The *p*-value associated with the *t*-test for non-causality is less than 0.01. However, when correcting for heteroskedasticity in the data, we can only reject the null hypothesis at a marginal 10% level. In Panel B we test if the trading volume of the equally weighted portfolio (EWVOL) helps predict the market returns. The Granger non-causality is strongly rejected based on the model with three autoregressive lags and one lag of trading volume EWVOL. The volume shows more significant predictive content in this case. A one-standard deviation increase in volume (0.237) is associated with a higher return of 0.086%, more than three times larger than the impact of VWVOL.[16]

Because daily returns feature significant time varying volatility (variance), we have reported both standard and heteroskedasticity-corrected *t*-statistics in panels A and B. As an alternative, we repeat the above tests after scaling raw returns by their conditional variance. We estimate the conditional variance by using the sum of the squared daily returns of asset *i* in the past three months.[17] The results in panels C and D show that the trading volume's predictive power for standardized returns is significant at the 5% level for the value-weighted portfolio and 1% for the equal-weighted portfolio.

Lamoureux and Lastrapes (1990) show that trading volume, used as a proxy for information arrival time, has significant explanatory power regarding the variance of daily

[16] As a robustness check, in implementing the Granger causality test, we also removed systematic day-of-the-week and month-of-the-year calendar effects from stock returns. The slopes for trading volume are smaller (0.099% and 0.276%, respectively) but remain statistically significant.

[17] We also estimate the time-varying variance by an exponential GARCH(1,1) model. The results are similar to what are reported here. In this paper we focus on realized variance estimates because it is both consistent (e.g., see Ghysels, Santa-Clara, and Valkanov (2005)) and, more importantly, easy to estimate for out-of-sample tests.

returns, and autoregressive conditional heteroscedastity (ARCH) effects tend to disappear when volume is included in the variance equation. Many other studies have also suggested that trading volume is related to volatility (e.g., Andersen, 1996; Gomes, 2005; Li and Wu, 2006; He and Velu, forthcoming). Therefore, market volatility could be a confounding factor in testing for causality between volume and returns. To address this issue, in panels E and F, we extend the two models in panels A and B to include past volatility as an additional predictive variable for returns. The predictive power of the lagged volume remains the same as before in terms of both the magnitude and the significance level of the slope coefficient (γ_1). We also find from the same regressions that past volatility is a strong predictive factor for market returns in the equal-weighted portfolio, but it is insignificant in predicting returns of the value-weighted portfolio. Nevertheless, the coefficient estimate for past volatility is not reported in the table to save space.

Why is the predictive power of trading volume stronger in equal-weighted portfolios than in value-weighted ones? One intuitive statistical explanation is that, as shown in Panel A of Table 1, daily returns to equal-weighted portfolios are more serially correlated than are those to value-weighted portfolios. The same pattern holds for the two corresponding trading volume series, although both are persistent. In addition, although not shown in Table 1, the contemporaneous correlation between equal-weighted portfolio returns and trading volume is 0.15, also larger than the correlation of 0.06 between the two variables for value-weighted portfolios. These two factors combine to cause returns on equal-weighted portfolios to be more predicable by trading volume than are their value-weighted counterparts.

Panels G through J contain results for the Granger causality tests using the two cross-sectionally constructed measures of high volume premiums (VWHVP and EWHVP) as predictive variables in place of two direct measures of trading volume VWVOL and EWVOL.[18] It can be seen from these panels that the volume premium is a significant predictor for market returns whether or not the past volatility is controlled for. As before, the evidence is stronger for the equal-weighted portfolios. Although not directly comparable because different market returns are used, the R^2 metrics of these four models are close to those in Panels A, B, E, and F, suggesting that high volume premium has in-sample predictive content for future returns similar to the two turnover measures.

Out-of-sample Granger causality test

As discussed in the methodology section, to implement the out-of-sample Granger causality test to evaluate the hypothesis that series X causes series Y, we obtain the one-step-ahead recursive forecasts from four competing models: C, R, U, and W. The first model we consider is one with a constant as the sole explanatory variable (the C model). In this simple model, the one-step-ahead forecast for the market return on day (t + 1) is simply the up to day t historical average returns. We include Model C as the benchmark because, as pointed out by Mayfield (2004), although a substantial body of research shows that expected returns vary over time, the naïve static approach of estimating the risk premium as the simple average of historical excess stock returns remains the most commonly employed method in practice. The next three models we use for forecasting have been introduced in the in-sample tests. Briefly, the R model is an autoregressive regression and includes past returns as the sole predictive variables. The U

[18] In studying the predictive power of aggregate trading volume, we have followed the literature and used market portfolio returns without dividends (capital gain only). Furthermore, we use our own estimates rather than CRSP estimates because stocks in NASDQ are not included in our market portfolio until November 1982 due to missing volume. However, literature focusing on stock return predictability especially in the out-of-sample context often use total returns. Therefore, to be comparable to the existing evidence, when examining the predictive power of the high volume premium we use CRSP market portfolio returns with dividends (VWRET and EWRET), assuming that high volume premium estimates based on NYSE/ASE stocks are also representative of those based on NASDQ stocks for the 1982 and earlier period (the summary statistics for VWRET and EWRET are reported in the bottom two rows of Panel A, Table 1). Nevertheless, our basic findings hold if returns without dividends are used instead.

model includes both past returns and past trading volume. And finally, the W model uses as predictive variables lagged values of stock returns, trading volume, and market volatility proxied by realized variance.

As the starting point, we estimate these four models using daily observations from the first ten years in our U.S. sample (July 1963-June 1973). The first set of one-step-ahead out-of-sample forecasts for the market portfolio returns are generated using the estimated coefficients and observed values of the predictive variables. The models are then re-estimated and new forecasts are generated after each daily observation is sequentially added to the estimation sample for the remaining 37 years of data. The forecast errors are formed by the differences between observed returns and the four forecasted returns.

Table 3 summarizes the performance of the four forecasting models where trading intensity is represented either by the value-weighted turnover (VWVOL) in the left panels or the equal-weighted turnover (EWVOL) in the right panels. For forecast evaluations, we first consider root mean squared forecast errors (RMSFE) of the four competing models. Panel A shows that model U that includes trading volume VWVOL and the model R that does not include the variable generate essentially the same forecast errors (1.063% after rounding) for the full sample period 1963-2010. However, both models underperform the simple historical average forecasts which have a RMSFE of 1.060%. The more complicated model W that also includes market volatility carries even larger forecast errors of 1.064%. In the middle of Panel A, we further test if the forecasts from the four models are statistically different from each other by employing the ENC-NEW encompassing test. The null hypothesis is that the forecasts from models in the first column (H_0) encompass those from models in the first row (H_A). Note that model W nests U, which in turn nests R. The C model is nested by all other three. We are interested in three null hypotheses: C encompasses U, R encompasses U, and U encompasses W. The null hypotheses that C encompasses U is strongly rejected, which contradicts the RMSFE measures. The hypothesis that R encompasses U is also rejected at the 5% level, meaning that lagged trading volume contains additional useful information about the next day's excess market return relative to the pure autoregressive model R. In line with the simple RMSFE measure, the test for model U encompassing W also concludes that market volatility has no significant predictive power beyond what is captured by lagged returns and the trading volume. Finally, we report in the bottom of Panel A realized utility levels associated with the four forecasting models. Model R has an annualized utility of 10.322%, which is slightly higher than that of the U model (10.187%).[19] Therefore, the rankings of models R and U by monetary gains are in line with those based on the RMSFE metric. Similarly, the realized utility of model W (9.835%) is lower than those of both models R and U, which is also consistent with the ranking by RMSFE and the encompassing test. However, the benchmark historical average, while delivering the smallest average forecast errors, attains the lowest level of utility.

Panel B presents quite a different picture of model rankings for the equal-weighted portfolio. The unrestricted model U generates the smallest average forecast errors followed by models W, R, and C. The hypotheses that model R and C encompass U are rejected at any conventional level, further confirming that trading volume helps predict returns to the equal-weighted portfolio. Realized utility based on model U's forecasts is 0.73% per annum higher than that of the R model. Both U and R models beat the historical average forecasts in economic gains by large margins. There is also some gain in utility (0.28%) by model W which further

[19] Following the literature, we set the risk aversion parameter r at 3. We also constrain the equity share in the optimal portfolio to the closed interval [0, 1], excluding short sales. Varying these parameters changes the magnitudes of the computed utility estimates but generally does not alter the models' rankings.

adds information on market volatility for forecasting. Nevertheless, in support of the RMSFE measure, the test of model U encompassing W has a statistic of 0.243 which is insignificant.

In panels C through E we compare the forecast performance of the four competing models for the two sub-sample periods. For the value-weighted portfolio, Panel C shows no significant evidence that model U generates better forecasts than model R during the period 1973-1999, although both appear to perform somewhat better than the historical average. Market volatility shows no additional predictive power in addition to past returns and the trading volume by all three evaluation statistics. As in Panel B for the equal-weighted portfolio, Panel D provides strong evidence that forecasts from model U are more accurate than those from model R. Although model W has the same RMSFE as model U, both the realized utility measure and the encompassing test result suggest that market volatility does contain some useful information for future stock return not captured by trading volume during the first sub-sample period.

In striking contrast to the results in panels C and D, panels E and F clearly show that, during the more recent 2000-2010 period, the simple forecasts based on the historical averages are more accurate for returns to both types of portfolios than those generated from the other three competitors.[20] This result based on the RMSFE measure is supported by the other two forecast evaluation methods for value-weighted portfolios. It however contradicts the rankings by the latter two for equal-weighted portfolios. The realized utility of model U is 24.561%, which is about 1% higher than that of the R model and more than 10% higher than that of the historical average. Furthermore, neither model R nor C encompasses model U.

[20] Recall that forecasts are recursively generated. As for the full sample analysis, the initial estimation sample for the first sub-sample analysis includes first ten years of data (1963-1973). And the initial estimation sample for the second sub-sample forecasting exercises spans a longer period of 1963-1999. This is why the null hypothesis of model R encompassing model U is rejected in Panel C only at the 10% level with a statistic of 1.488, while it is rejected in Panel E at the 5% level with a smaller statistic of 1.219.

Replacing the direct measures of trading volume VWVOL and EWVOL with the high-volume premiums VWHVP and EWHVP, we re-estimate and generate forecasts from all four models examined in Table 3. Their performances are summarized in Table 4. The main findings from Table 3 all hold true in Table 4. In particular, the additional predictive power of the high volume return premium is significant only for the equal-weighted portfolio. More specifically, the U model performs better than the other three models by all three measures reported in Panel B for the full sample and in Panel D for the first sub-sample. Note, however, that the annualized utility gains relative to the R model are small in both cases (0.4% and 0.5%, respectively). Finally, as in Table 3, there is no consistent evidence of a positive relationship between EWHVP and future market returns during the 2000-2010 forecasting period. Model U with EWHVP has a larger RMSFE than model C (1.215% vs. 1.195%), although it still contains additional information useful for forecasting according to the encompassing test. The utility gain of model U over model R is merely 0.12%, essentially nonexistent.

Robustness

Studies focusing on out-of-sample forecast errors rather than within sample pricing errors have found that empirical results from dynamic models may be sensitive to the choice of predictive variables, assets, and in-sample window lengths (e.g., Cooper and Gulen, 2006). To guard against possible data snooping bias, we re-examine the performance of the four forecasting models, doubling the length of the initial estimation sample from 10 years (as used in tables 3 and 4) to 20 years (July 1963-July 1983). The new results are summarized in Appendix tables 1 and 2 for turnover and the high volume return premium, respectively. Note that because forecasts are generated recursively, the change in the initial in-sample length has no impact on the second forecasting period (January 2000-December 2010). Comparing the results in the two appendix tables to their counterparts in tables 3 and 4, we can easily see that changing the in-sample window length does not affect any of our earlier findings. In short, turnover and the volume premium display additional predictive power for future returns only in equal-weighted portfolios by both statistical and monetary measures. Utility gains for the unrestricted model U are in the range of 0.65%~0.78% when using equal-weighted turnover as the extra predictor, and 0.47%~0.71% when using the volume premium. These estimates are similar in magnitudes to the initial estimation using the first ten years of data.

We have so far found from tables 3 and 4 that trading volume consistently helps predict future returns only for the equal-weighted portfolio. This suggests that the lead/lag Granger-causal relationship may exist more prominently in small stocks.[21] To shed further light on this issue, we also study the predictive power of aggregate trading volume for returns to portfolios of various sizes. Specifically, we sort stocks into small, medium, and large portfolios based on the breakpoints for the low 30%, medium 40%, and high 30% of the ranked values of market capitalization. The value- and equal-weighted returns are computed for each of these three size portfolios. To save space, Appendix Table 3 only presents the performance of the four models in forecasting the daily returns to the small and large portfolios. The results for small stocks are tabulated in panels A, C, and E in the left half of the table, and those for large stocks in panels B, D, and F in the right half. By all three forecast evaluation criteria, model U with the extra variable EWVOL performs better than model R and the historical average for small stocks in the full and the first sub-samples. The forecasting ability of the equal-weighted volume premium is also significant in the recent sub-sample period by the two statistical measures. However, incorporating this variable in the forecasting information set decreases an investor's utility by

[21] One possible explanation from microstructure theory is that small stocks are generally less liquid than their large counterparts. A sell or buy order of the same magnitude can have much larger price impacts on small stocks.

about 0.4% compared to model R. There is also no economic gain in forecasting small stock returns by using past market volatility in either the full sample or the two sub-samples. Consider the three panels in the right half of the table, we find that model U does not outperform model R for large stocks by the average forecast errors and the realized utility measures, although the encompassing test results continue to show that volume information relevant for forecasting future returns is not fully captured by historical returns. Overall, the predictive power of trading volume recorded in Table 3 is indeed mainly driven by small stocks.

Both Campbell, Grossman, and Wang (1993) and Llorente et al. (2002) predict a nonlinear relation between trading volume and stock returns. Hiemstra and Jones (1994) also present empirical evidence of nonlinear Granger causality between the two variables. However, model (1) does not allow for such nonlinearity in the test. To address the potential misspecification issue, we adopt the robust method of quantile regressions (Koenker, 2005). The lag structures for the quantile regressions are the same as those used in the linear models in Tables 3 and 4. In deriving forecasts from the quantile regressions, we estimate a total of 99 regressions for quantiles 0.01, 0.02 .., 0.99. We then generate one-step-ahead return forecasts for each quantile and form the mean forecast by taking a simple average over these 99 estimates. We find that the quantile models perform slightly better than the linear ones for the equal-weighted portfolios during the 1973-1999 sub-sample period when the predictive power of trading volume is most noticeable. The Granger causality tests based on the restricted and unrestricted quantile regressions provide strong evidence for the predictive power of trading volume as in Tables 3 and 4 using the linear models.

In sum, the out-of-sample forecasting results provide partial support for the findings of the in-sample regressions. The additional predictive power of information on lagged trade activity is corroborated by the out-of-sample tests for future stock returns to equal-weighted portfolios but not for returns to value-weighted ones.

6. International Evidence

In this section we extend our analysis on the U.S. equity market to other developed and emerging markets with two goals. As pointed out earlier, Kaniel, Ozoguz, and Starks (2012) document significant high volume premiums in 15 out of 20 developed markets, although the volume effect is not as persistent in the emerging economies as it is in the developed markets.[22] The sample they use ends in 2001. Our first goal is therefore to examine whether the high volume return premium continues to exist in the international markets after we include data from the past decade. As shown in the last section, the out-of-sample predictive power of trading volume for stock returns is vanishing in the U.S. market and the volume premium is smaller in the 2000-2010 sample period. We are therefore interested in whether the burst of the technology bubble and the economic recession around the new millennium, and the latest financial crisis and the ensuing economic recession have had significant impact on the volume premium and market behavior in the international markets. Our second goal is to examine whether the in-sample and out-of-sample predictive power of trading volume on stock returns found in the U.S. market is also present in international markets. To develop perspective on whether the trading volume effect is integrated across regions, we further examine whether turnover and the volume premium in the U.S. market have predictive power for stock returns in the other markets. This part of the analysis is motivated by such empirical evidence as the spillover effect documented in the literature of idiosyncratic volatility (e.g., Guo and Savicks, 2008) and the leading role for the

[22] There are a few other studies on the volume-return relation in and across international markets, notably, Gagnon and Karolyi (2009). An early example is Saatcioglu and Starks (1998) focusing exclusively on Latin American markets.

U.S. with respect to monthly international excess return predictability reported in Rapach, Strauss, and Zhou (forthcoming).

Data

For international analysis, we first consider stock markets in six non-U.S. Group of Seven (G-7) countries: Canada, France, Germany, Italy, Japan, and the United Kingdom (U.K.). We also study another 12 countries whose primary exchanges make up the top 20 worldwide major stock exchanges by market capitalization. These 12 countries/regions, including both developed and developing economies, are Australia, Brazil, China, Hong Kong, India, Korea, Russia, Singapore, South Africa, Spain, Switzerland, and Taiwan. Our international data on firm-level daily returns, trading volume and monthly market capitalization, both for currently trading and defunct securities, are obtained from Thomson-Reuter Datastream.[23] Given the potential data errors or outliers in the Datastream as identified by previous research (e.g., Ince and Porter, 2006), we implement two sets of filtering rules on the raw data in addition to those sampling requirements applied to the U.S. data in forming volume-based portfolios. We first follow Kaniel, Ozoguz, and Starks (2012) and remove stocks whose local currency prices fall below the lowest five percentile of stock prices in the country's sample for that year. We then follow Guo and Savickas (2008) and set the daily return on a stock to a missing value if the recorded return is greater than 300% on that day. If the price of the stock falls by more than 90% in a day *and* it has increased by more than 200% within the previous 20 trading days, we set all daily returns between the two dates to missing values. Similarly, if the price of a stock increases by more than 100% in a day *and* it has decreased by more than 200% within the previous 20 trading days, we also set all daily returns between the two dates to missing values. The price we pay for these

[23] Market portfolio returns are measured as the percentage change of total return indexes. The empirical results reported later are very similar if they are measured as the percentage of prices only.

more stringent data-filtering rules is that, compared to Kaniel, Ozoguz, and Starks (2012), we work with a smaller number of stocks and shorter samples for some countries during the overlapping periods.

Based on the filtered data, we form the estimates of daily value- and equal-weighted aggregate turnover for each of the18 countries. We also follow the same strategies used in previous sections for the U.S. data in forming the volume portfolios and estimating the high volume return premium. Because the limited availability of the volume data, the number of trading stocks meeting all the selection criteria are small for many countries, particularly in the earlier years of the samples. Therefore, for the international markets, we sort all stocks into quintiles rather than centiles as we did for the U.S. market. The value- (equal-)weighted high volume return premium is defined as the difference between the value- (equal-)weighted portfolio returns on the top volume quintile and the returns on the bottom volume quintile. Conceivably, HVP would be higher were it based on centiles rather than on quintiles. Different stock exchanges within the same country may have different trade-volume dynamics due to differences in the institutional details. To reduce this type of heterogeneity and its possible impact on the estimation, we follow the literature and only study the primary exchanges in each country.[24]

Table 5 reports the start date and the average number of firms that the samples comprise for the G-7 countries (excluding the U.S.) in Panel A and for the other 12 countries in Panel B. Although the end date is December 2010 for all 18 countries, the effective sample start date ranges from January 1977 to September 2005. The average number of firms considered also

[24] The list of countries that have multiple primary exchanges includes Canada, China, Germany, India, Korea, Russia, and Spain.

varies considerably from 61 of Brazil to 1712 of Japan, with an average of 486 and a median of 410.

Table 6 provides mean statistics for the monthly market portfolio returns and the two types of high volume return premiums for the non-U.S. G-7 countries in Panel A and for the 12 other countries in Panel B. The third and the fourth columns of the table present our estimates of value- and equal-weighted market portfolio returns.[25] The value-weighted estimates are close to Datastream's total market estimates for the respective countries in Column 2. The equal-weighted market portfolio returns are higher than the value-weighted ones with the exceptions of Germany, South Africa, Spain, and Switzerland. Column 5 reports the average value-weighted high volume premium. The premium is positive for all non-U.S. G-7 countries and statistically different from zero for France, Japan, and UK at the 5% level. The equal-weighted premium is both positive and statistically significant in all six countries. These results are similar to those reported in Kaneil, Ozoguz, and Starks (2012) who find that significant volume effects exist in all G-7 countries but Italy. The magnitudes of our estimates are also similar to their counterparts in Kaneil, Ozoguz, and Starks (2012) despite using quite different sample periods. For the remaining 12 countries, the value-weighed premium is positive in all but three markets, China, Korea, and Spain. Nevertheless, the premium is statistically significant in Hong Kong only. Similar to the U.S. market, the equal-weighted premium is estimated more precisely and larger compared to the value-weighted one. It is significant in eight countries, although the estimate is again negative for the Korean market. Russia, which is not included in Kaneil, Ozoguz, and

[25] Market portfolio returns are given in the form of total returns instead of real or excess returns so that they are comparable in magnitudes across countries since the inflation rate and the risk-free rate vary significantly from country to country, especially during the early sample period. And different estimates of risk free rates are sometimes used in the literature.

Starks (2012), also has a large point estimate of 1.57% per month. Nevertheless, the estimate is noisy, likely because a small number of stocks and a short sample period are used (see Table 5).

In-sample Granger causality test

Table 7 summarizes in-sample Granger causality test results for the G-7 countries excluding the U.S.[26] Panel A presents the results when the value-weighted return and turnover series are used in the test. In striking contrast to the positive relation found in the U.S. market (Panel E in Table 2), high volume today predicts lower return tomorrow for all six other G-7 countries. Although this negative effect is smaller in magnitude than the positive effect in the U.S., it is statistically significant for the Canadian and the UK markets by the HC-*t* values. Panel B shows that the association between equal-weighted turnover and future returns can be positive or negative. But it is insignificant in all six countries, meaning that trading volume contains no predictive power for future returns on the equal-weighted market portfolios after we control for past returns and market volatility. Like in the U.S. market, the higher value-weighted volume premium does predict higher returns tomorrow shown in Panel C, although the effect is imprecisely estimated and there also appears to be a significant reversal in returns the day after tomorrow for Germany and Japan. The exception is the U.K. where the dynamic volume-return relation is negative. Panel D of Table 7 presents evidence of the predictive power of the equal-weighted volume premium, which is more in line with that from the U.S. shown in Panel J of Table 2. All γ estimates are positive, and the associated HC-*t* values are no less than 1 for all six countries and larger than 2 for the Canadian and German markets.

Table 8 includes the in-sample test results for the 12 other countries where aggregate turnover is used as one of the predictive variables. Similar to the non-U.S. G-7 countries, the

[26] In deriving excess market returns, we use, when available, International Financial Statistics' (IFS) government treasury bill rates as risk free rates. If such series are either missing during the sample period or simply unavailable, we use money market rates or short-term deposit rates as substitutes

value-weighted turnover predicts lower returns for all but three countries (Panel A). The predictive power is significant in half of the cases. In Panel B, we can see that the sign of γ associated with one-day lagged turnover is positive in seven countries and negative in the remaining five. The positive coefficients are significant for China, Hong Kong, and Taiwan at the 10% or better levels, while none of the negative coefficients are significant.

The empirical results for the in-sample Granger causality test using the two high volume return premiums are presented in Table 9. According to Panel A, the value-weighted premium VWHVP has significant predictive power for future returns in the Indian and Taiwanese markets; and the associated γ estimate is positive in both cases. The premium also appears to have some marginal power in predicting a lower future return for the Spanish stocks. Among the 12 countries, a higher equal-weighted premium in panel B on average is followed by higher future returns in 9 countries. However, the predictive power is significant only in Australia and Switzerland, and marginally significant in India.

Out-of-sample Granger causality test

We conduct out-of-sample Granger causality tests for the 18 international markets based on the models specified in Tables 7, 8, and 9. As before, the initial samples for most countries include first ten years of data. However, the sample start dates of Germany, Brazil, and Russia are 1999 or later. Therefore, the in-samples only contain the first five years of data for Germany and Brazil, and the first three years of data for Russia. Although we consider the same set of forecasting models for each of the 18 markets as we did for the U.S. market, to save space, we only report the results for the following three models: model C with a constant only, model RZ with past returns and market volatility, and model W with all three sets of predictive variables (past returns, measures of trading intensity, and market volatility). Obviously, model W nests RZ, which in turn nests C. To test if past trading volume and the high volume premium contain useful information in forecasting current returns, we consider the null hypothesis that forecasts from the restricted model RZ encompass those from the unrestricted model W. If this hypothesis is rejected we further test if the forecasts of the simple historical averages encompass those from model W.

We first present the test results for the six G-7 countries in Table 10. As shown in Panel A, model W that uses value-weighted turnover as a predictor performs no better than model RZ that does not include the variable in terms of the root mean squared forecast errors (RMSFE) of five markets. It achieves relatively high realized utility only in the French market (2.4% per annum vs. 0.74% by model C and −0.75% by RZ). The encompassing test results are consistent with the rankings by the RMSFE measure. The U.K. appears to be an exception where turnover used in model W has additional predictive power by all three measures in comparison to the RZ model. We also reject the null hypothesis that the historical average forecasts contain all of the useful information that model W does. The gain in utility by including volume information is about 1% relative to the RZ model, and 0.7% higher than forecasts from model C. In Panel B we use equal-weighted turnover as a proxy for trade intensity. The evidence is more consistent in the sense that turnover does not have additional predictive power for future returns for any market by all three evaluation methods after past returns and market volatility are controlled for in the regressions.

The results tabulated in Panel C suggest that the unrestricted model W only performs better than both C and RZ models by the two statistical measures in the German market. It however obtains lower utility than the restricted RZ model by 0.9%. The evidence is similar in the U.K. market. If proxied by the equal-weighted volume premium in forecasting equations

(Panel D), trading volume shows consistent evidence of forecasting ability for returns in the Canadian market by all three evaluation measures. The volume premium also appears to have predictive power for German stock returns according to the utility gain and the encompassing test result. However, the W model with the volume premium produces relatively larger average forecast errors than the historical average forecasts.

Tables 11 and 12 summarize the results of the out-of-sample causality test for the remaining 12 countries using turnover and the high volume premium, respectively. Given that the sample sizes are generally small for the international markets (in terms of the period spanned and/or the number of stocks included), inferences drawn from the statistical measures may be more likely to diverge from those drawn from the economic measure. For this reason and for brevity as well, from now on we define the two trading activity proxies as having predictive power for stock returns if (1) the unrestricted model W that uses this information achieves higher utility than both model RZ and the simple historical average forecasts (model C); and (2) the latter two models have larger forecast errors on average and do not encompass the unrestricted model W at the 5% significance level. Applying these rules we can see from Panel A of Table 11 that model W with the variable of value-weighted turnover performs better than the C and RZ models in three markets, India, Russia, and Switzerland. Panel B presents results for the aggregate measure of turnover when it is constructed on an equal-weight basis. Although the equal-weighted turnover EWVOL helps forecast returns by the encompassing test in Australia, Hong Kong, and Taiwan, there is no economic gain in doing so since the realized utility of the unrestricted model is slightly lower than the restricted model RZ in all three cases.

As shown in Panel A of Table 12, the only market in which the value-weighted volume premium VWHVP helps forecast future returns by all three evaluation methods is India. Still, the economic benefit is admittedly small, an increases of less than 0.1% per annum relative to the model without the volume information. And there is no consistent evidence that the cross-sectionally constructed volume premium helps forecast returns in any of the other 11 countries. Like the value-weighted measure, the equal-weighted high volume premium does not have predictive power for all but one country (Panel B). The exception is the Chinese stock market where the unrestricted model W performs better than both the historical average and the restricted RZ model by both statistical and economic measures. Economically, an investor would be better off with additional annual returns of 1.1% if volume information is used in rebalancing her/his portfolio. Based on the two statistical measures (RMSFE and ENC-NEW), the equal-weighted HVP also helps forecast returns in the Australian and the Swiss markets. However, the evidence is negative by the economic measure of realized utility.

Finally, we briefly discuss the empirical results on whether trading volume from the U.S. market contains additional information for forecasting returns in the 18 other markets controlling for volume and volatility information from the domestic markets. We consider one-step ahead forecasts from four models. As before, model C includes an intercept only. Model W contains past returns, volume, and volatility of a domestic market. The other two models, WR and WRV, augment model W with U.S. market information. Specifically, WR adds one lag of U.S. market returns, and WRV adds both one lag of U.S. market returns and one lag of trading volume. Based on the four sets of return forecasts, we compute annualized utility levels according to equations (3) and (4). Appendix Table 4 presents the estimated economic gains associated with model C, W, WR, and WRV for each market when trading activity is approximated by two aggregate turnover series. To facilitate the presentation, we again define U.S. trading volume information as having predictive power for stock returns on another market if model WRV that includes this information (1) achieves higher utility than the more restricted models C, W, and WR models, and (2) the latter three models have larger forecast errors and do not encompass model WRV at the 5% significance level. Although past U.S. market returns contain substantial information for predicting current returns to the other markets, the value-weighted turnover of the U.S. market only provides additional information useful for predicting the Indian market. The added economic gain is 2.8% per annum. The equal-weighted turnover shows predictive power in three more markets (Canada, Japan, and Korea). Nevertheless, the gains in utility are smaller, ranging from 0.12% to 0.74%.

Judged on all three statistical and economic criteria, we also find little evidence that U.S. market trading activity carries additional predictive power for international markets when it is represented by the high volume return premium. We see from Appendix Table 5 that when the value-weighted volume premium is included in the models, investors' welfare improves in the French and Russian markets. For the equal-weighted portfolio investment, the improvement is found in one market only (Hong Kong). And in all three cases, the economic gains are small (0.16~0.41%).

7. Concluding Remarks

We provide a comprehensive reexamination of the lead-lag relationship between trading volume and stock returns within the popular framework of Granger causality. While also presenting the standard in-sample evidence of the predictive power of trading activity for stock returns, our contribution to the literature rests importantly in the paper's emphasis being on detailing out-of-sample evidence, thereby increasing the reliability of our empirical findings. In the U.S. market, higher trading volume, whether measured by aggregate time series of turnover or by the cross-sectionally constructed high volume return premium, is indeed followed by higher stock returns. However, such predictive power of trading volume should be interpreted with caution. As emphasized in the introduction, Granger causality may or may not reflect underlying causality. Even if the causal relation is fundamental in this case and investors can increase returns by simply trading more, the associated economic gain is quantitatively small for the market as a whole. Furthermore, the predictive power of trading volume becomes insignificant even statistically in the more recent period featuring high-profile high-speed trading.

The international evidence of the predictive power at best is mixed based on in-sample regressions. More importantly, with only a few exceptions, the Granger causal effect of trading volume on stock returns fail to pass the rigorous statistical and economic tests in out-of-sample regressions.

References

Andersen, T.G. 1996. Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility. Journal of Finance 51, 169-204.

Ashley, R., C.W.J. Granger, and R.I. Schmalensee. 1980. Advertising and Aggregate Consumption: An Analysis of Causality. Econometrica 48, 1149-1168.

Barber, B.M., Y-T. Lee, Y-J. Liu, and T. Odean. 2009. Just How Much Do Individual Investors Lose by Trading? Review of Financial Studies 22, 609-630.

Blume, L., D. Easley, and M. O'Hara. 1994. Market Statistics and Technical Analysis: The Role of Volume. Journal of Finance 49, 153-81.

Brennan, M., T. Chordia, and A. Subrahmanyam. 1998. Alternative Factor Specifications, Security Characteristics, and the Cross-Section of Expected Stock Returns. Journal of Financial Economics 49, 345-373.

Butler, A.W., G. Grullon, and J.P. Weston. 2005. Can Managers Forecast Aggregate Market Returns? Journal of Finance 60, 963-986.

Campbell, J.Y., S.J. Grossman, and J. Wang. 1993. Trading Volume and Serial Correlation in Stock Returns. Quarterly Journal of Economics 108, 905-939.

Campbell, J.Y., and S.B. Thompson. 2008. Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average? Review of Financial Studies 21, 1509-1531.

Chen, J., H. Hong, and J. Stein. 2001. Forecasting Crashes: Trading Volume, Past Returns and Conditional Skewness in Stock Prices. Journal of Financial Economics 61, 345-381.

Chordia, T., S-W. Huh, and A. Subrahmanyam. 2007. The Cross-Section of Expected Trading Activity. Review of Financial Studies 20, 709-740.

Chordia, T., R. Roll, and A. Subrahmanyam. 2011. Recent Trends in Trading Activity and Market Quality. Journal of Financial Economics 101, 243-263.

Chordia, T., and B. Swaminathan. 2000. Trading Volume and Cross-Autocorrelations in Stock Returns. Journal of Finance 55, 913-935.

Clark, T.E., and M.W. McCracken. 2001. Tests of Equal Forecast Accuracy and Encompassing for Nested Models. Journal of Econometrics 105, 85-110.

Cochrane, J.H. 2013. Finance: Function Matters, not Size. Journal of Economic Perspectives 27, 29-50.

Conrad, J., A. Hameed, and C. Niden. 1994. Volume and Autocovariances in Short-Horizon Security Returns. Journal of Finance 49, 1305-1329.

Conrad, J., and G. Kaul. 1988. Time-Variation in Expected Returns. Journal of Business 61, 409-425.

Cooper, M. 1999. Filter Rules Based on Price and Volume in Individual Security Overreaction. Review of Financial Studies 12, 901-935.

Cooper, M., and H. Gulen. 2006. Is Time-Series-Based Predictability Evident in Real Time? Journal of Business 79, 1263-1292.

Corradi, V., and N.R. Swanson. 2006. Predictive Density Evaluation. In Elliott, G., C.W.J. Granger, and A. Timmermann (eds.) *Handbook of Economic Forecasting*, Vol. 1, Elsevier, Amsterdam, pp. 197-284.

Easley, D., M. O'Hara, and P.S. Srinivas. 1998. Option Volume and Stock Prices: Evidence on Where Informed Traders Trade. Journal of Finance 53, 431-465.

Fama, E.F. 1991. Efficient Capital Markets: II. Journal of Finance 46, 1575-1617.

Fama, E., and K. French. 1998. Value versus Growth: The International Evidence. Journal of Finance 53, 1975-1999.

French, K. 2008. The Cost of Active Investing. Journal of Finance 63, 1537-1573.

Gagnon, L., and G.A. Karolyi. 2009. Information, Trading Volume, and International Stock Return Comovements: Evidence from Cross-Listed Stocks. Journal of Financial and Quantitative Analysis 44, 953-986.

Gallant, A.R., P.E. Rossi, and G. Tauchen. 1992. Stock Prices and Volume. Review of Financial Studies 5, 199-242.

Gervais, S., R. Kaniel, and D. Mingelgrin. 2001. The High Volume Return Premium. Journal of Finance 56, 877-919.

Ghysels, E., P. Santa-Clara, and R. Valkanov. 2005. There Is a Risk-Return Tradeoff after All. Journal of Financial Economics 76, 509-548.

Gomes, F.J. 2005. Portfolio Choice and Trading Volume with Loss-Averse Investors**.** Journal of Business 78, 675-706.

Granger, C.W.J. 1969. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. Econometrica 36, 424-438.

Granger, C.W.J. 1988. Some Recent Developments in a Concept of Causality. Journal of Econometrics 39, 199-211.

Greenwood, R., and D.S. Scharfstein. 2013. The Growth of Finance. Journal of Economic Perspectives 27, 3-28.

Griffin, J.M. 2002. Are the Fama and French Factors Global or Country-Specific? Review of Financial Studies 15, 783-803.

Griffin, J.M., J.H. Harris, and S. Topaloglu. 2003. The Dynamics of Institutional and Individual Trading. Journal of Finance 58, 2285-2320.

Griffin, J.M., F. Nardari, and R.M. Stulz. 2007. Do Investors Trade More When Stocks Have Performed Well? Evidence from 46 Countries. Review of Financial Studies 23, 905-951.

Guo, H., and R. Savickas. 2008. Average Idiosyncratic Volatility in G7 Countries. Review of Financial Studies 21, 1259-1296.

He, X., and R. Velu. Forthcoming. Volume and Volatility in a Common Factor Mixture of Distributions Model. Journal of Financial and Quantitative Analysis.

Hiemstra, C., and J.D. Jones. 1994. Testing for Linear and Nonlinear Granger Causality in the Stock Price-Volume Relation. Journal of Finance 49, 1639-1664.

Ince, O.S., and R.B. Porter. 2006. Individual Equity Return Data from Thomson Datastream: Handle with Care! Journal of Financial Research 29, 463-479.

Inoue, A., and L. Kilian. 2005. In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use? Econometric Reviews 23, 371-402.

Kandel, E., and N.D. Pearson. 1995. Differential Interpretation of Public Signals and Trade in Speculative Markets. Journal of Political Economy 103,831-872.

Kaniel, R., A. Ozoguz, and L. Starks. 2012. The High Volume Return Premium: Cross Country Evidence. Journal of Financial Economics 103, 255-279.

Karpoff, J. 1987. The Relation between Price Changes and Trading Volume: A Survey. Journal of Financial and Quantitative Analysis 22, 109-126.

Koenker, R. 2005. Quantile Regression. Cambridge University Press, Cambridge.

Lamoureux, C.G., and W.D. Lastrapes. 1990. Heteroskedasticity in Stock Return Data: Volume versus GARCH Effects. Journal of Finance 45, 221-229.

Li, J., and C. Wu. 2006. Daily Return Volatility, Bid-Ask Spreads, and Information Flow: Analyzing the Information Content of Volume. Journal of Business 79, 2697-2739.

Llorente, G., R. Michaely, G. Saar, and J. Wang. 2002. Dynamic Volume-Return Relation of Individual Stocks. Review of Financial Studies 15, 1005-1047.

Lo, A.W., and C. MacKinlay. 1988. Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test. Review of Financial Studies 1, 41-66.

Lo, A.W., and J. Wang. 2000. Trading Volume: Definitions, Data Analysis, and Implications of Portfolio Theory. Review of Financial Studies 13, 257-300.

Lo, A.W., and J. Wang. 2006. Trading Volume: Implications of an Intertemporal Capital Asset Pricing Model. Journal of Finance 61, 2805-2840.

Malcolm, B., and J.C. Stein. 2004. Market Liquidity as a Sentiment Indicator. Journal of Financial Markets 7, 271-299.

Marquering, W., and M. Verbeek. 2004. The Economic Value of Predicting Stock Index Returns and Volatility. Journal of Financial and Quantitative Analysis 39, 407-429.

Mayfield, E.S. 2004. Estimating the Market Risk Premium. Journal of Financial Economics 73, 465-496.

Merton, R. 1973. An Intertemporal Capital Asset Pricing Model. Econometrica 41, 867-887.

Odean, T. 1999. Do Investors Trade Too Much? American Economic Review 89, 1279-1298.

Rapach, D.E., J.K. Strauss, and G. Zhou. 2010. Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy. Review of Financial Studies 23, 821-862.

Rapach, D.E., J.K. Strauss, and G. Zhou. Forthcoming. International Stock Return Predictability: What is the Role of the United States? Journal of Finance, DOI: 10.1111/jofi.12041.

Rouwenhorst, K.G. 1999. Local Return Factors and Turnover in Emerging Stock Markets. Journal of Finance 54, 1439-1464.

Saatcioglu, K., and L.T. Starks. 1998. The Stock Price-Volume Relationship in Emerging Stock Markets: The Case of Latin America. International Journal of Forecasting 14, 215-225.

Schneider, J. 2009. A Rational Expectations Equilibrium with Informative Trading Volume. Journal of Finance 64, 2783-2805.

Schwert, W. 1989. Why Does Stock Market Volatility Change Over Time? Journal of Finance 44, 1115-1155.

Statman, M., S. Thorley, and K. Vorkink. 2006. Investor Overconfidence and Trading Volume. Review of Financial Studies 19, 1531-1565.

Thornton, D. L., and G. Valente. Forthcoming. Out-of-Sample Predictions of Bond Excess Returns and Forward Rates: An Asset Allocation Perspective. Review of Financial Studies, doi:10.1093/rfs/hhs069.

Wang, J. 1994. A Model of Competitive Stock Trading Volume. Journal of Political Economy 102, 127-68.

Welch, I., and A. Goyal. 2008. A Comprehensive Look at the Empirical Performance of Equity Premium Prediction. Review of Financial Studies 21, 1455-1508.

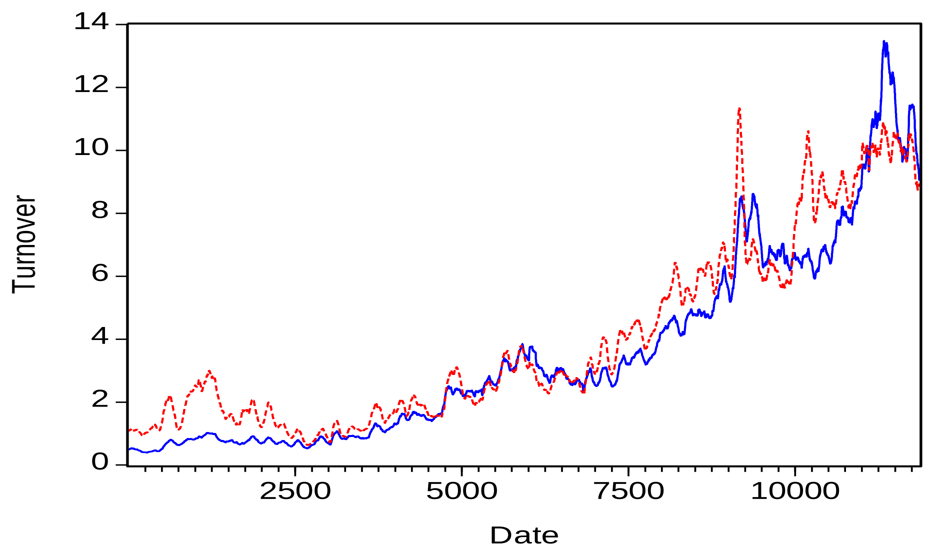
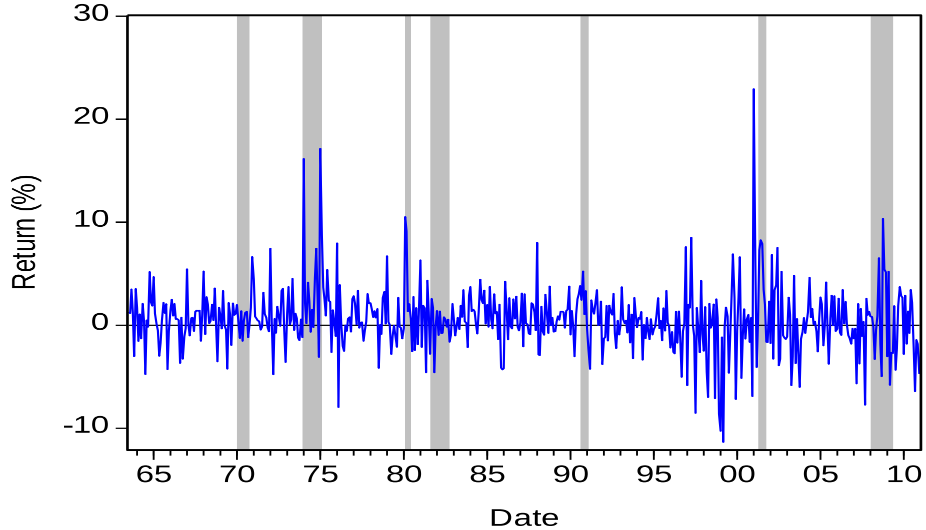


Figure 1. Daily value-weighted and equally-weighted turnover in the U.S. market (raw measures)

The solid and broken lines are 100-day moving averages of value- and equal-weighted turnover ratios, respectively. They are estimated using daily data from July 1, 1963 to December 31, 2010. The first observation on the horizontal axis corresponds to November 20, 1963 and the last observation 11860 corresponds to December 31, 2010.

Panel A. Value-weighted HVP



Panel B. Equal-weighted HVP

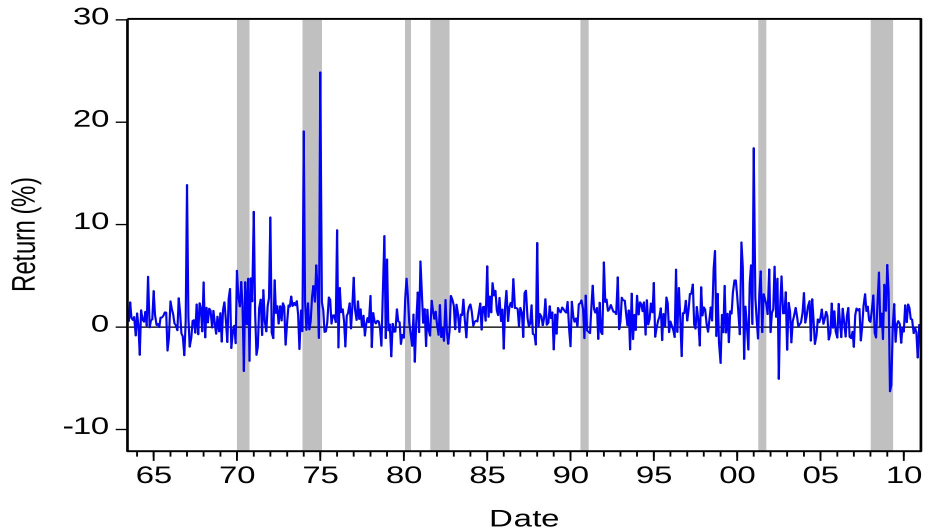


Figure 2. The monthly high volume return premium (HVP) in the U.S. market

The high volume return premium, defined as the return on a zero-cost portfolio that is long on stocks experiencing unusually high trading volume on the last day of each month and short on low-volume stocks, is estimated using daily and monthly U.S. data from July 1963 to December 2010. Shaded areas indicate U.S. economic recessions as designated by the National Bureau of Economic Research (NBER).

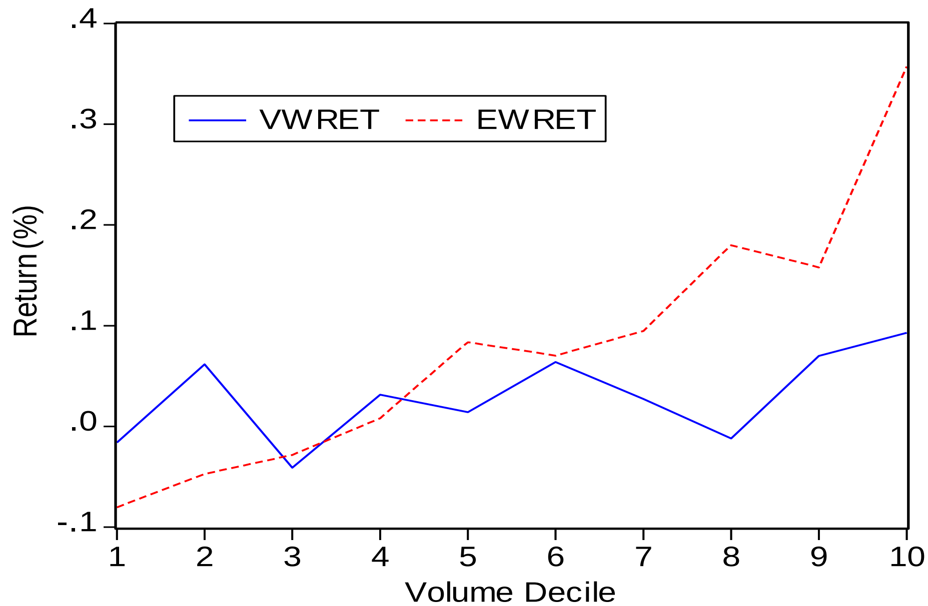


Figure 3. Average daily market returns when sorted on the lagged volume

This figure reports on subset daily U.S. market returns from July 1963 to December 2010, which are sorted into ten decile groups based on the one-day lagged trading volume. Group 1 is the lowest volume decile and Group 10 the highest decile.

Table 1. Descriptive statistics of volume and returns in the U.S. market

VWRET and EWRET are value- and equal-weighted market portfolio returns, and VWVOL and EWVOL are the corresponding turnover. VWHVP and EWHVP are value- and equal-weighted high volume return premiums. They are estimated using U.S. market daily data from July 1, 1963 through December 31, 2010, of which the first 100 observations are used for detrending the trading volume series (therefore, the effective sample for the first four rows starts at November 21, 1963). VWMKT and EWMKT are CRSP market portfolio returns in excess of the risk free rate.

Note that our VWRET and EWRET are not comparable to those of the CRSP counterparts. Because the volume data of NASDAQ stocks are not included in the database until November 1982, these stocks are excluded from the estimation of the two portfolio returns during this period. Differences in other stock selection criteria for the whole sample period also contribute to the differences between CRSP's estimates and ours.

The mean and standard deviation are both in percentage forms. AR(1) is the first-order autocorrelation coefficient. Q(1) is the Ljung-Box Q-statistic for the null hypothesis that there is no autocorrelation up to order 1, which follows the χ^2 distribution with one degree-of-freedom. Symbols *, **, and *** indicate significant at the 10%, 5%, and 1% levels, respectively.

Panel A. Daily volume and returns (July 01, 1963-December 30, 2010)

Variable	Mean	Std. dev.	Skewness	Kurtosis	AR(1)	Q(1)
VWVOL	1.197**	0.217	–0.252***	2.959***	0.604	>100***
EWVOL	0.879	0.237	0.048**	1.614***	0.734	>100***
VWMKT	0.010	0.991	–0.522***	17.474***	0.064	48.482***
EWMKT	0.061***	0.921	–0.529***	12.092***	0.234	>100***
VWHVP	0.027***	0.727	–0.135***	18.428***	0.059	41.395***
EWHVP	0.061***	0.452	0.779***	8.525***	0.088	92.205***
VWRET	0.021**	0.982	−0.533***	17.118***	0.069	57.233***
EWRET	0.060***	0.831	−0.591***	14.564***	0.237	>100***

Panel B. Monthly volume and returns

	July 1963-Dec. 2010		July 1963-Dec. 1999		Jan. 2000- Dec. 2010	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
VWHVP	0.574***	3.134	0.618***	2.865	0.430	1.197
EWHVP	1.294***	2.568	1.343***	2.541	1.132***	4.483
VWHML	0.538**	4.713	0.451*	4.451	0.827	0.030
EWHML	1.150***	4.255	1.106***	3.822	1.294**	2.588

Table 2. In-sample Granger causality test results for the U.S. market

VWMKT and EWMKT are CRSP value- and equal-weighted market portfolio returns, and VWVOL and EWVOL are the corresponding turnover. VWRET and EWRET are our estimates of value- and equal-weighted market portfolio returns; and VWHVP and EWHVP are value- and equal-weighted high volume return premiums. *h* is the conditional market variance estimated from a GARCH model. Market volatility is approximated by the realized market variance (σ^2), which in turn is the sum of the squared daily returns in the past three months. These variables are all estimated/constructed using U.S. daily data from July 1963 to December 2010. L_y is the number of lags on stocks returns in Equation (1) in the text. L_x and L_z are the numbers of lags on one of the two trading volume measures (turnover and the volume premium) and market volatility, respectively. They are selected by minimizing Schwarz's Bayesian information criterion (BIC). The null hypothesis is that trading volume does not predict stock returns, implying the joint zero restrictions on parameters associated with volume terms γ_n. Coefficient γ_1 in Panels A-F is multiplied by 100 for ease of presentation.
The symbols *, **, and *** indicate significant at the 10%, 5%, and 1% levels, respectively.

L_y	L_x	L_z	γ_1	*t*-stat.	*HC*- *t*-stat.	Adj-R^2 (%)	BIC
Panel A $VWMKT_t = \alpha + \sum_{m=1}^{L_y} \beta_m VWMKT_{t-m} + \sum_{n=1}^{L_x} \gamma_n VWVOL_{t-n} + \varepsilon_t$							
2	1		0.119	2.841***	1.952*	0.651	–9.233
Panel B $EWMKT_t = \alpha + \sum_{m=1}^{L_y} \beta_m EWMKT_{t-m} + \sum_{n=1}^{L_x} \gamma_n EWVOL_{t-n} + \varepsilon_t$							
3	1		0.362	10.315***	7.049***	6.993	–9.444
Panel C $VWMKT_t / h_t = \alpha + \sum_{m=1}^{L_y} \beta_m VWMKT_{t-m} / h_{t-m} + \sum_{n=1}^{L_x} \gamma_n VWVOL_{t-n} + \varepsilon_t / h_t$							
1	1		0.099	2.214**		1.816	0.097
Panel D $EWMKT_t / h_t = \alpha + \sum_{m=1}^{L_y} \beta_m EWMKT_{t-m} / h_{t-m} + \sum_{n=1}^{L_x} \gamma_n EWVOL_{t-n} + \varepsilon_t / h_t$							
5	1		0.382	9.249***		11.394	0.038
Panel E $VWMKT_t = \alpha + \sum_{m=1}^{L_y} \beta_m VWMKT_{t-m} + \sum_{n=1}^{L_x} \gamma_n VWVOL_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma_{t-r}^2 + \varepsilon_t$							
2	1	1	0.121	2.880***	1.956*	0.648	–9.232
Panel F $EWMKT_t = \alpha + \sum_{m=1}^{L_y} \beta_m EWMKT_{t-m} + \sum_{n=1}^{L_x} \gamma_n EWVOL_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma_{t-r}^2 + \varepsilon_t$							
3	1	1	0.371	10.532***	7.083***	7.077	–9.444
Panel G $VWRET_t = \alpha + \sum_{m=1}^{L_y} \beta_m VWRET_{t-m} + \sum_{n=1}^{L_x} \gamma_n VWHVP_{t-n} + \varepsilon_t$							
2	1		0.034	2.758***	1.579	0.660	–9.248

Panel H $EWRET_t = \alpha + \sum_{m=1}^{L_y} \beta_m EWRET_{t-m} + \sum_{n=1}^{L_x} \gamma_n EWHVP_{t-n} + \varepsilon_t$							
5	1		0.065	3.984***	2.286**	6.666	–9.643
Panel I $VWRET_t = \alpha + \sum_{m=1}^{L_y} \beta_m VWRET_{t-m} + \sum_{n=1}^{L_x} \gamma_n VWHVP_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma_{t-r}^2 + \varepsilon_t$							
2	1	1	0.034	2.749***	1.579	0.652	–9.243
Panel J $EWRET_t = \alpha + \sum_{m=1}^{L_y} \beta_m EWRET_{t-m} + \sum_{n=1}^{L_x} \gamma_n EWHVP_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma_{t-r}^2 + \varepsilon_t$							
5	1	3	0.064	3.937***	2.290**	6.856	–9.638

Table 3. Out-of-sample Granger causality test results for the U.S. market (trading volume proxied by turnover)

Model C includes a constant only. Model R includes only lagged returns. Model U includes both lags of CRSP value-weighted (equal-weighted) market portfolio returns VWMKT (EWMKT) and lags of the corresponding market turnover VWVOL (EWVOL). Model W includes lags of stock returns, turnover, and market volatility. The number of lags of each variable (L_y, L_x, and L_z) is the same as in Table 2.
Root mean squared forecast errors (RMSFE) and the annualized realized utility have been multiplied by 100 for ease of presentation. All the statistics in the table are based on one-step-ahead recursive forecast errors from the above models for the daily excess returns to the U.S. stock market portfolio.
ENC-NEW is the encompassing test statistic of Clark and McCracken (2001), in which the associated null hypothesis is that forecasts from the model in the first column (H_o) encompass those from the model in the second row (H_A). Thus, the H_o model is preferred to H_A. The critical values of the ENC-NEW test are linearly interpolated from the unpublished Appendix of Clark and McCracken (2001). The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Forecasting models				Forecasting models			
	C	R	U	W	C	R	U	W
L_y	0	2	2	2	0	3	3	3
L_x	0	0	1	1	0	0	1	1
L_z	0	0	0	1	0	0	0	1
	Panel A. VWVOL & VWMKT, 1973-10				Panel B. EWVOL & EWMKT, 1973-10			
	RMSFE				RMSFE			
	1.060	1.063	1.063	1.064	0.956	0.939	0.935	0.936
	Encompassing test statistic				Encompassing test statistic			
C			149.933***				839.127***	
R			3.112**				60.535***	
U				−8.551				0.243
	Realized Utility				Realized Utility			
	5.074	10.322	10.187	9.835	13.733	32.722	33.448	33.730
	Panel C. VWVOL & VWMKT, 1973-99				Panel D. EWVOL & EWMKT, 1973-99			
	RMSFE				RMSFE			
	0.901	0.898	0.898	0.899	0.744	0.704	0.700	0.700

	Encompassing test statistic				Encompassing test statistic			
C			221.754***				1099.131***	
R			1.488*				64.656***	
U				−3.830				5.005***
	Realized Utility				Realized Utility			
	6.800	16.792	16.693	16.208	15.203	36.562	37.188	37.627
	Panel E. VWVOL & VWMKT, 2000-10				Panel F. EWVOL & EWMKT, 2000-10			
	RMSFE				RMSFE			
	1.366	1.380	1.379	1.381	1.329	1.341	1.337	1.339
	Encompassing test statistic				Encompassing test statistic			
C			−4.842				108.919***	
R			1.219**				11.874***	
U				−3.459				−1.252
	Realized Utility				Realized Utility			
	0.971	−5.042	−5.263	−5.298	10.238	23.598	24.561	24.470

Table 4. Out-of-sample Granger causality test results for the U.S. market (trading volume proxied by the high volume premium)

Model C includes a constant only. Model R includes only lagged returns. Model U includes both lags of value-weighted (equal-weighted) market portfolio returns VWRET (EWRET) and lags of the corresponding high volume return premiums VWHVP (EWHVP). Model W includes lags of stock returns, volume premium, and market volatility. The number of lags of each variable (L_y, L_x, and L_z) is the same as in Table 2.

Root mean squared forecast errors (RMSFE) and the annualized realized utility have been multiplied by 100 for ease of presentation. All the statistics in the table are based on one-step-ahead recursive forecast errors from the above models for the daily excess returns to the U.S. stock market portfolio.

ENC-NEW is the encompassing test statistic of Clark and McCracken (2001), in which the associated null hypothesis is that forecasts from the model in the first column (H_o) encompass those from the model in the second row (H_A). Thus, the H_o model is preferred to H_A. The critical values for linear models of the ENC-NEW test are linearly interpolated from the unpublished Appendix of Clark and McCracken (2001). The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Forecasting models				Forecasting models			
	C	R	U	W	C	R	U	W
L_y	0	2	2	2	0	5	5	5
L_x	0	0	1	1	0	0	1	1
L_z	0	0	0	1	0	0	0	3
	Panel A. VWHVP & VWRET, 1973-10				Panel B. EWHVP & EWRET, 1973-10			
	RMSFE				RMSFE			
	1.057	1.059	1.059	1.060	0.857	0.842	0.841	0.847
	Encompassing test statistic				Encompassing test statistic			
C			166.687***				851.692***	
R			2.395*				10.720***	
U				−10.216				−20.798
	Realized Utility				Realized Utility			
	4.277	13.023	12.895	12.817	13.439	31.283	31.676	31.409
	Panel C. VWHVP & VWRET, 1973-99				Panel D. EWHVP & EWRET, 1973-99			
	RMSFE				RMSFE			
	0.886	0.879	0.879	0.880	0.666	0.622	0.621	0.627
	Encompassing test statistic				Encompassing test statistic			
C			261.705***				1219.171***	
R			0.058				11.402***	
U				−3.772				−7.251

	Realized Utility				Realized Utility			
	5.801	19.535	19.269	19.218	15.062	35.618	36.124	35.457
	Panel E. VWHVP & VWRETD, 2000-10				Panel F. EWHVP & EWRETD, 2000-10			
	RMSFE				RMSFE			
	1.381	1.398	1.398	1.400	1.195	1.216	1.215	1.221
	Encompassing test statistic				Encompassing test statistic			
C			−7.901				89.519***	
R			1.351**				2.157***	
U				−4.369				−8.089
	Realized Utility				Realized Utility			
	0.633	−2.524	−2.322	−2.466	9.558	20.929	21.050	21.739

Table 5. Basic information on international samples of stock returns and trading volume

This table reports the sample start dates and the average numbers of included stocks for international data which are obtained from Datastream. All samples end in December 2010. The numbers of stocks used to estimate aggregate turnover are generally higher than those used in estimating the cross-sectional high volume return premiums as reported in this table.

Country	Start date	Average number of stocks
	Panel A G-7 countries excluding the U.S.	
Canada	January 1977	453
France	July 1991	452
Germany	January 1999	384
Italy	April 1994	194
Japan	December 1990	1712
UK	January 1991	702
	Pane B Twelve other countries	
Australia	January 1984	436
Brazil	August 2003	61
China	March 1996	930
Hong Kong	June 1988	384
India	January 1995	1118
Korea	September 1987	811
Russia	September 2005	72
Singapore	January 1984	176
South Africa	January 1996	185
Spain	January 1991	105
Switzerland	May 1990	128
Taiwan	April 1991	448

Table 6. Descriptive statistics for international markets

TOTMKT is Datastream total market returns for the country. VWRET and EWRET are our estimates of value- and equal-weighted market portfolio returns, and VWHVP and EWHVP are value- and equal-weighted high volume return premiums. All three market portfolio returns are in percentage. The symbols *, **, and *** denote that the entry (i.e., the mean statistic) is significant at the 10%, 5%, and 1% levels, respectively, based on heteroskedasticity-and-autocorrelation consistent (HAC) errors.
The numbers in parentheses are pair-wise correlations between the volume premiums of the 18 countries and that of the U.S. market.

Countries	TOTMKT	VWRET	EWRET	VWHVP	EWHVP
Panel A G-7 countries excluding the U.S.					
Canada	1.049***	1.073***	1.602***	0.189	1.041***
France	0.843**	0.796**	0.879**	0.548**	1.266***
Germany	0.444	0.504	0.304	0.656	0.841**
Italy	0.629	0.624	0.578	0.297	0.647***
Japan	-0.04	−0.050	0.206	0.479**	0.980***
UK	0.826***	0.833***	0.885*	0.641**	1.494***
Pane B Twelve other countries					
Australia	1.104***	1.052***	1.215***	0.273	1.603***
Brazil	1.970**	2.370***	2.728**	0.579	1.604***
China	1.267	1.458*	1.960**	−0.519	−0.630
Hong Kong	1.393***	1.388***	1.534**	0.858***	1.500***
India	1.547**	1.623**	2.435**	0.585	0.187
Korea	1.199**	1.074*	1.687**	−0.469	−0.894***
Russia	1.280	1.492	2.745	0.535	1.566
Singapore	0.826*	0.952**	1.204*	0.385	0.625**

South Africa	1.421***	1.334***	1.243***	0.432	0.688***
Spain	0.945**	0.934**	0.859*	−0.250	0.655**
Switzerland	0.824**	0.827**	0.814*	0.285	0.593**
Taiwan	0.872	0.700	0.907	0.461	0.381

Table 7. In-sample Granger causality test results for G-7 countries excluding the U.S.

VWRET and EWRET are value- and equal-weighted market portfolio returns, and VWVOL and EWVOL are the corresponding turnover. VWHVP and EWHVP are value- and equal-weighted high volume return premiums. Market volatility is represented by the realized market variance (σ^2), which in turn is the sum of the squared daily returns in the past three months. These variables are all estimated/constructed using Datastream daily data for the respective periods tabulated in Table 6. L_y is the number of lags on stocks returns in equation (1) in the text. L_x and L_z are the numbers of lags on one of the two trading volume measures (turnover and the volume premium), and market volatility, respectively. They are selected by minimizing Schwarz's Bayesian information criterion (BIC). The null hypothesis is that trading volume does not predict stock returns, implying the joint zero restrictions on parameters associated with volume terms γ_n. Coefficient γ_1 in panels A and B is multiplied by 100 for ease of presentation.
When two lags are chosen for VWHVP in Panel C (L_x = 2) in cases of Germany and Japan, γ_2 is reported one row below that of γ_1.
The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Market	L_y	L_x	L_z	γ_1	t-stat.	HC- t-stat.	Adj-R^2 (%)	BIC
Panel A $VWRET_t = \alpha + \sum_{m=1}^{L_y} \beta_m VWRET_{t-m} + \sum_{n=1}^{L_x} \gamma_n VWVOL_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma_{t-r}^2 + \varepsilon_t$								
Canada	1	1	6	−0.064	−2.438**	−2.248**	2.127	−9.416
France	1	1	2	−0.094	−1.648*	−1.478	0.528	−8.802
Germany	1	1	2	−0.001	−0.083	−0.064	0.846	−8.608
Italy	5	1	2	−0.084	−1.334	−1.166	1.102	−8.577
Japan	1	1	1	−0.015	−0.251	−0.261	0.224	−8.643
UK	6	1	2	−0.143	−2.586***	−2.360**	1.332	−8.981
Panel B $EWRET_t = \alpha + \sum_{m=1}^{L_y} \beta_m EWRET_{t-m} + \sum_{n=1}^{L_x} \gamma_n EWVOL_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma_{t-r}^2 + \varepsilon_t$								
Canada	4	1	1	−0.000	−0.080	−0.060	8.662	−9.603
France	4	1	2	−0.000	−0.123	−0.140	12.130	−10.242
Germany	4	1	2	0.012	0.300	0.319	10.780	−9.361
Italy	4	1	2	0.023	0.604	0.541	4.570	−9.312
Japan	1	1	2	0.001	0.231	0.196	5.244	−9.059
UK	8	1	2	−0.046	−1.302	−1.124	18.729	−10.242
Panel C $VWRET_t = \alpha + \sum_{m=1}^{L_y} \beta_m VWRET_{t-m} + \sum_{n=1}^{L_x} \gamma_n VWHVP_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma_{t-r}^2 + \varepsilon_t$								
Canada	1	1	6	0.016	1.406	0.944	2.094	−9.399
France	1	1	2	0.009	0.574	0.428	0.455	−8.792
Germany	1	2	2	0.016	0.824	0.467	1.321	−8.539
				−0.068	−3.625***	−2.141**		

Italy	5	1	2	0.006	0.345	0.249	1.014	−8.565
Japan	1	2	1	0.033	1.252	0.908	0.452	−8.636
				−0.092	−3.506***	−2. 685***		
UK	6	1	2	−0.042	−2.636***	−1.940*	1.376	−8.972

Panel D $EWRET_t = \alpha + \sum_{m=1}^{L_y} \beta_m EWRET_{t-m} + \sum_{n=1}^{L_x} \gamma_n EWHVP_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma_{t-r}^2 + \varepsilon_t$

Canada	4	1	1	0.047	4.696***	3.171***	5.873	−9.561
France	4	1	2	0.027	1.556	1.128	9.325	−10.121
Germany	4	1	2	0.069	3.274***	2.003**	4.886	−8.956
Italy	4	1	2	0.019	1.108	1.098	4.092	−9.264
Japan	1	1	2	0.077	2.055**	1.501	4.164	−8.960
UK	8	1	4	0.033	1.759*	1.316	13.297	−10.374

Table 8. In-sample Granger noncausality test results for 12 other countries (trading volume proxied by turnover)

VWRET and EWRET are value- and equal-weighted market portfolio returns, and VWVOL and EWVOL are the corresponding turnover. Market volatility is represented by the realized market variance (σ^2), which in turn is the sum of the squared daily returns in the past three months. These variables are all estimated/constructed using Datastream daily data for the respective periods tabulated in Table 6. L_y is the number of lags on stocks returns in equation (1) in the text. L_x and L_z are the numbers of lags on trading volume (turnover), and market volatility, respectively. They are selected by minimizing Schwarz's Bayesian information criterion (BIC). The null hypothesis is that trading volume does not predict stock returns, implying the joint zero restrictions on parameters associated with volume terms γ_n. Coefficient γ_1 in panels A and B is multiplied by 100 for ease of presentation.

When two lags are chosen for India (L_x = 2), γ_2 is reported one row below that of γ_1.

The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Country	L_y	L_x	L_z	γ_1	t-stat.	HC- t-stat.	Adj-R^2 (%)	BIC
Panel A $VWRET_t = \alpha + \sum_{m=1}^{L_y} \beta_m VWRET_{t-m} + \sum_{n=1}^{L_x} \gamma_n VWVOL_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma^2_{t-r} + \varepsilon_t$								
Australia	1	1	6	−0.068	−1.941*	−1.584	1.175	−9.168
Brazil	1	1	3	−0.016	−0.318	−0.416	1.141	−8.028
China	1	1	1	0.113	1.741*	1.496	0.687	−7.942
Hong Kong	1	1	2	−0.144	−2.184**	−1.707*	0.309	−8.204
India	1	2	1	0.170	1.644	1.354	2.082	−8.163
				−0.443	−4.298***	−3.852***		
Korea	1	1	2	−0.074	−1.118	−1.008	0.522	−7.986
Russia	1	1	1	−0.657	−2.818***	−2.717***	0.445	−7.178
Singapore	1	1	3	−0.056	−1.390	−1.097	1.677	−8.685
South Africa	1	1	2	−0.131	−2.213**	−2.243**	1.569	−8.805
Spain	1	1	1	−0.048	−0.951	−0.848	0.086	−8.753
Switzerland	5	1	2	−0.149	−2.788***	−2.433**	1.083	−9.013
Taiwan	1	1	1	0.141	2.326**	2.062**	0.146	−8.250
Panel B $EWRET_t = \alpha + \sum_{m=1}^{L_y} \beta_m EWRET_{t-m} + \sum_{n=1}^{L_x} \gamma_n EWVOL_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma^2_{t-r} + \varepsilon_t$								
Australia	8	1	3	−0.000	−0.264	−0.183	8.921	−9.248
Brazil	1	1	3	−0.042	−0.948	−0.855	2.300	−8.415
China	1	1	1	0.150	2.138**	1.853*	0.564	−7.794
Hong Kong	3	1	2	0.116	2.339**	1.850*	6.264	−8.386
India	4	2	1	0.166	1.446	1.155	13.095	−8.317
				−0.455	−4.021***	−3.152***		
Korea	1	1	1	−0.054	−0.918	−0.791	2.989	−8.117
Russia	1	1	1	0.028	0.242	0.188	3.126	−7.827
Singapore	4	1	2	0.000	0.064	0.049	4.964	−8.493
South Africa	5	1	1	−0.062	−1.292	−1.349	5.805	−9.448
Spain	4	1	2	0.032	1.143	1.140	2.738	−9.450
Switzerland	13	1	3	−0.038	−1.103	−0.907	8.348	−9.961
Taiwan	1	1	1	0.194	3.403***	3.134***	1.110	−8.321

Table 9. In-sample Granger causality test results for 12 other countries (trading volume proxied by the high volume premium)

VWRET and EWRET are value- and equal-weighted market portfolio returns, and VWHVP and EWHVP are the corresponding high volume return premiums. Market volatility is represented by the realized market variance (σ^2), which in turn is the sum of the squared daily returns in the past three months. These variables are all estimated/constructed using Datastream daily data for the respective periods tabulated in Table 6. L_y is the number of lags on stocks returns in equation (1) in the text. L_x and L_z are the numbers of lags on the volume premium, and market volatility, respectively. They are selected by minimizing Schwarz's Bayesian information criterion (BIC). The null hypothesis is that trading volume does not predict stock returns, implying the joint 0 restrictions on parameters associated with volume terms γ_n.
The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Country	L_y	L_x	L_z	γ_1	t-stat.	HC- t-stat.	Adj-R^2 (%)	BIC
Panel A $VWRET_t = \alpha + \sum_{m=1}^{L_y} \beta_m VWRET_{t-m} + \sum_{n=1}^{L_x} \gamma_n VWHVP_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma^2_{t-r} + \varepsilon_t$								
Australia	1	1	6	−0.010	−0.718	−0.477	1.105	−9.146
Brazil	1	1	3	−0.012	−0.487	−0.373	0.954	−7.987
China	1	1	2	0.031	1.172	0.840	0.236	−7.931
Hong Kong	1	1	2	0.004	0.209	0.147	0.241	−8.190
India	1	1	1	0.089	3.975***	3.103***	1.883	−8.160
Korea	1	1	2	0.041	2.342**	1.642	0.552	−7.975
Russia	1	1	1	−0.000	−0.000	−0.000	-0.254	−7.146
Singapore	1	1	3	0.024	1.897*	1.462	1.676	−8.661
South Africa	1	1	2	−0.004	−0.212	−0.170	1.354	−8.784
Spain	2	1	2	−0.031	−2.419**	−1.837*	0.511	−8.733
Switzerland	5	1	1	0.013	0.921	0.642	0.759	−8.982
Taiwan	1	1	1	0.048	2.477**	1.995**	0.159	−8.249
Panel B $EWRET_t = \alpha + \sum_{m=1}^{L_y} \beta_m EWRET_{t-m} + \sum_{n=1}^{L_x} \gamma_n EWHVP_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma^2_{t-r} + \varepsilon_t$								
Australia	4	1	6	0.052	4.138***	2.507**	5.357	−9.200
Brazil	1	1	3	−0.015	−0.712	−0.622	1.728	−8.451
China	1	1	1	0.061	1.792*	1.309	0.427	−7.780
Hong Kong	3	1	2	0.010	0.449	0.314	4.226	−8.284
India	3	1	1	0.080	2.302**	1.893*	8.720	−8.245
Korea	1	1	1	0.012	0.447	0.257	2.706	−8.077
Russia	1	1	1	0.000	0.000	0.000	0.777	−7.871
Singapore	4	1	2	0.001	0.041	0.025	3.384	−8.416
South Africa	3	1	2	0.018	1.503	1.476	4.531	−9.479
Spain	1	1	2	−0.008	−0.611	−0.420	1.964	−9.343
Switzerland	4	1	2	0.040	2.629***	1.969**	4.992	−9.791
Taiwan	1	1	1	0.037	1.520	1.176	0.820	−8.312

Table 10. Out-of-sample Granger causality test results for G-7 countries excluding the U.S.

Model C includes a constant only. Model RZ includes both lags of value-weighted (equal-weighted) market portfolio returns VWRET (EWRET) and lags of market volatility (σ^2), where market volatility is represented by the realized market variance, which in turn is the sum of the squared daily returns in the past three months. Model W includes lags of market portfolio returns, the corresponding aggregate turnover (VWVOL and EWVOL) or the high volume return premium (VWHVP and EWHVP), and market volatility. The number of lags of each variable is the same as in Table 7.

Root mean squared forecast errors (RMSFE) and the annualized realized utility have been multiplied by 100 for ease of presentation. All the statistics in the table are based on one-step-ahead recursive forecast errors from the above models.

ENC-NEW is the encompassing test statistic of Clark and McCracken (2001), in which the associated null hypothesis is that forecasts from model H_o encompass those from H_A. Thus, the H_o model is preferred to H_A. The critical values of the ENC-NEW test are linearly interpolated from the unpublished Appendix of Clark and McCracken (2001). The symbols ** and *** denote significance at the 5% and 1% levels, respectively.

Country	RMSFE			ENC-NEW		Realized Utility		
	C	RZ	W	RZ vs. W	C vs. W	C	RZ	W
Panel A $VWMKT_t = \alpha + \sum_{m=1}^{L_y} \beta_m VWMKT_{t-m} + \sum_{n=1}^{L_x} \gamma_n VWVOL_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma_{t-r}^2 + \varepsilon_t$								
Canada	0.941	0.965	0.966	0.789		3.604	14.814	14.166
France	1.356	1.357	1.357	0.940		0.741	−0.753	2.387
Germany	1.307	1.308	1.308	−0.830		2.809	6.556	5.915
Italy	1.381	1.373	1.373	0.226		−2.475	5.825	5.250
Japan	1.431	1.440	1.440	−0.292		−0.673	1.359	1.657
UK	1.294	1.295	1.294	2.506**	19.667***	3.276	2.940	3.946
Panel B $EWMKT_t = \alpha + \sum_{m=1}^{L_y} \beta_m EWMKT_{t-m} + \sum_{n=1}^{L_x} \gamma_n EWVOL_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma_{t-r}^2 + \varepsilon_t$								
Canada	0.889	0.862	0.862	−2.233		28.244	46.720	46.783
France	0.712	0.677	0.677	−0.189		23.053	40.029	39.940
Germany	0.963	0.923	0.923	−0.395		63.138	74.020	73.833
Italy	1.017	0.998	0.998	−0.198		−3.199	20.589	20.110
Japan	1.181	1.187	1.188	−0.830		3.675	24.242	23.263
UK	0.736	0.673	0.673	0.296		6.545	36.676	36.619
Panel C $VWRET_t = \alpha + \sum_{m=1}^{L_y} \beta_m VWRET_{t-m} + \sum_{n=1}^{L_x} \gamma_n VWHVP_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma_{t-r}^2 + \varepsilon_t$								
Canada	0.945	0.969	0.969	−0.160		3.846	15.117	14.785
France	1.360	1.361	1.361	−0.538		0.949	−0.428	0.232
Germany	1.346	1.346	1.344	6.071***	25.868***	2.050	6.355	5.471
Italy	1.369	1.361	1.362	−0.410		−1.362	5.990	5.828
Japan	1.431	1.440	1.438	4.309***	-0.999	−0.620	1.508	3.009
UK	1.315	1.317	1.316	2.455**	19.111***	2.334	1.846	1.668

Panel D $EWRET_t = \alpha + \sum_{m=1}^{L_y} \beta_m EWRET_{t-m} + \sum_{n=1}^{L_x} \gamma_n EWHVP_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma_{t-r}^2 + \varepsilon_t$

Canada	0.873	0.862	0.861	12.615***	290.92***	23.724	37.032	38.323
France	0.742	0.717	0.717	0.078		7.146	27.503	27.651
Germany	1.148	1.162	1.163	4.040***	59.470***	9.190	28.439	30.185
Italy	1.019	1.001	1.001	0.433		−1.295	18.543	18.452
Japan	1.208	1.211	1.212	−0.873		1.461	20.294	18.300
UK	0.829	0.788	0.788	0.164		3.725	30.915	31.052

Table 11. Out-of-sample Granger causality test for 12 other countries (trading volume proxied by turnover)

Model C includes a constant only. Model RZ includes both lags of value-weighted (equal-weighted) market portfolio returns VWRET (EWRET) and lags of market volatility (σ^2), where market volatility is represented by the realized market variance, which in turn is the sum of the squared daily returns in the past three months. Model W includes lags of market portfolio returns, the corresponding aggregate turnover (VWVOL and EWVOL), and market volatility. The number of lags of each variable is the same as in Table 8.

Root mean squared forecast errors (RMSFE) and the annualized realized utility have been multiplied by 100 for ease of presentation. All the statistics in the table are based on one-step-ahead recursive forecast errors from the above models.

ENC-NEW is the encompassing test statistic of Clark and McCracken (2001), in which the associated null hypothesis is that forecasts from model H_o encompass those from H_A. Thus, the H_o model is preferred to H_A. The critical values of the ENC-NEW test are linearly interpolated from the unpublished Appendix of Clark and McCracken (2001). The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Country	RMSFE			ENC-NEW		Realized Utility		
	C	RZ	W	RZ vs. W	C vs. W	C	RZ	W
Panel A $VWRET_t = \alpha + \sum_{m=1}^{L_y} \beta_m VWRET_{t-m} + \sum_{n=1}^{L_x} \gamma_n VWVOL_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma_{t-r}^2 + \varepsilon_t$								
Australia	0.965	0.968	0.968	1.063		3.134	4.267	4.742
Brazil	2.138	2.173	2.173	−0.278		−7.076	−8.195	−7.966
China	2.340	2.342	2.344	−0.515		3.679	5.672	2.732
Hong Kong	1.533	1.532	1.532	1.664*	8.041***	4.363	11.312	12.011
India	1.811	1.802	1.795	6.832***	25.705***	10.384	24.699	26.854
Korea	1.942	1.939	1.938	0.300		4.823	15.203	16.690
Russia	1.850	1.853	1.849	4.633***	4.117***	12.436	7.577	24.703
Singapore	1.262	1.265	1.265	0.441		1.110	13.124	13.314
South Africa	1.372	1.374	1.373	0.780*	10.109***	5.154	15.786	15.693
Spain	1.350	1.356	1.356	-0.100		1.874	−4.921	−3.051
Switzerland	1.223	1.222	1.221	2.869**	18.038***	−1.313	3.247	3.752
Taiwan	1.435	1.435	1.436	−0.233		2.646	5.296	6.561
Panel B $EWRET_t = \alpha + \sum_{m=1}^{L_y} \beta_m EWRET_{t-m} + \sum_{n=1}^{L_x} \gamma_n EWVOL_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma_{t-r}^2 + \varepsilon_t$								
Australia	1.020	0.975	0.975	3.800***	4.575*	36.358	55.501	55.430
Brazil	1.626	1.632	1.632	0.145		4.531	23.010	22.070
China	2.580	2.574	2.575	0.004		13.185	38.649	34.278
Hong Kong	1.463	1.410	1.410	2.958**	216.31***	20.396	60.671	60.091
India	1.654	1.532	1.531	0.626		8.173	68.212	66.970
Korea	1.798	1.778	1.779	−0.534		8.933	47.012	47.271
Russia	1.470	1.451	1.451	−0.092		82.960	85.736	85.558

Singapore	1.460	1.424	1.424	−0.709		8.379	43.366	42.714
South Africa	0.810	0.794	0.795	−0.206		27.910	38.183	38.556
Spain	0.900	0.895	0.895	0.235		6.249	18.365	18.574
Switzerland	0.791	0.765	0.765	0.008		8.442	28.649	28.499
Taiwan	1.427	1.417	1.417	3.857***	27.054***	4.857	22.081	21.996

Table 12. Out-of-sample Granger causality test results for other 12 countries (trading volume proxied by the high volume premium)

Model C includes a constant only. Model RZ includes both lags of value-weighted (equal-weighted) market portfolio returns VWRET (EWRET) and lags of market volatility (σ^2), where market volatility is represented by the realized market variance, which in turn is the sum of the squared daily returns in the past three months. Model W includes lags of market portfolio returns, the corresponding high volume return premium (VWHVP and EWHVP), and market volatility. The number of lags of each variable is the same as in Table 9.

Root mean squared forecast errors (RMSFE) and the annualized realized utility have been multiplied by 100 for ease of presentation. All the statistics in the table are based on one-step-ahead recursive forecast errors from the above models.

ENC-NEW is the encompassing test statistic of Clark and McCracken (2001), in which the associated null hypothesis is that forecasts from model H_o encompass those from H_A. Thus, the H_o model is preferred to H_A. The critical values of the ENC-NEW test are linearly interpolated from the unpublished Appendix of Clark and McCracken (2001). The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Country	RMSFE			ENC-NEW		Realized Utility		
	C	RZ	W	RZ vs. W	C vs. W	C	RZ	W
Panel A $VWRET_t = \alpha + \sum_{m=1}^{L_y} \beta_m VWRET_{t-m} + \sum_{n=1}^{L_x} \gamma_n VWHVP_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma^2_{t-r} + \varepsilon_t$								
Australia	0.969	0.972	0.972	−1.072		2.675	3.743	3.540
Brazil	2.161	2.196	2.198	−0.681		−2.923	-5.341	−5.363
China	2.322	2.323	2.323	−0.001		7.326	6.150	7.773
Hong Kong	1.549	1.548	1.548	−0.721		3.666	10.721	11.099
India	1.802	1.792	1.790	3.196***	22.461***	11.904	26.373	26.455
Korea	1.968	1.964	1.963	7.630***	19.816***	5.141	15.289	15.083
Russia	1.750	1.753	1.753	−0.092		11.409	10.356	9.535
Singapore	1.273	1.276	1.276	1.498*	41.085***	1.053	12.900	12.161
South Africa	1.374	1.377	1.377	−0.275		5.231	14.517	14.348
Spain	1.363	1.369	1.368	3.002***	−5.407	0.442	−0.963	−0.894
Switzerland	1.225	1.224	1.224	0.097		−1.513	2.932	3.055
Taiwan	1.453	1.453	1.454	0.425		2.847	5.559	3.340
Panel B $EWRET_t = \alpha + \sum_{m=1}^{L_y} \beta_m EWRET_{t-m} + \sum_{n=1}^{L_x} \gamma_n EWHVP_{t-n} + \sum_{r=1}^{L_z} \lambda_r \sigma^2_{t-r} + \varepsilon_t$								
Australia	1.021	1.001	0.999	10.014***	176.27***	9.018	32.342	31.472
Brazil	1.673	1.691	1.692	−0.188		8.332	18.146	17.433
China	2.563	2.560	2.558	0.763*	4.300***	13.650	30.964	32.082
Hong Kong	1.507	1.474	1.475	−0.681		8.913	46.052	45.831
India	1.700	1.588	1.590	−1.331		7.882	67.834	67.756
Korea	1.862	1.838	1.839	−1.299		6.754	43.844	43.314
Russia	1.315	1.312	1.312	−0.079		45.353	51.751	51.600
Singapore	1.495	1.471	1.472	−1.669		0.570	33.649	32.957
South Africa	0.835	0.826	0.827	−0.517		10.279	24.833	24.854

Spain	0.938	0.933	0.934	−1.000		2.796	13.850	12.645
Switzerland	0.837	0.815	0.814	3.797***	140.73***	3.058	24.840	24.213
Taiwan	1.445	1.436	1.436	0.447		2.847	5.559	3.340

Appendix Table 1. Out-of-sample Granger causality test results for the U.S. market (trading volume proxied by turnover, in-sample length 20 years)

Model C includes a constant only. Model R includes only lagged returns. Model U includes both lags of CRSP value-weighted (equal-weighted) market portfolio returns VWMKT (EWMKT) and lags of the corresponding market turnover VWVOL (EWVOL). Model W includes lags of stock returns, turnover, and market volatility (σ^2). Market volatility is represented by the realized market variance, which in turn is the sum of the squared daily returns in the past three months. The number of lags of each variable (L_y, L_x, and L_z) is the same as in Table 2.

Root mean squared forecast errors (RMSFE) and the annualized realized utility have been multiplied by 100 for ease of presentation. All the statistics in the table are based on one-step-ahead recursive forecast errors from the above models for the daily excess returns to the U.S. stock market portfolio.

ENC-NEW is the encompassing test statistic of Clark and McCracken (2001), in which the associated null hypothesis is that forecasts from the model in the first column (H_o) encompass those from the model in the second row (H_A). Thus, the H_o model is preferred to H_A. The critical values of the ENC-NEW test are linearly interpolated from the unpublished Appendix of Clark and McCracken (2001). The symbols * and *** denote significance at the 10% and 1% levels, respectively.

	Forecasting models				Forecasting models			
	C	R	U	W	C	R	U	W
	Panel A. VWVOL & VWMKT, 1983-10				Panel B. EWVOL & EWMKT, 1983-10			
	RMSFE				RMSFE			
	1.113	1.121	1.121	1.123	1.000	0.996	0.992	0.994
	Encompassing test statistic				Encompassing test statistic			
			38.608***				437.42***	
			1.724*				40.698***	
				−7.512				−2.361
	Realized Utility				Realized Utility			
	3.894	6.908	6.305	6.046	14.955	29.502	30.277	30.249
	Panel C. VWVOL & VWMKT, 1983-99				Panel D. EWVOL & EWMKT, 1983-99			
	RMSFE				RMSFE			
	0.899	0.903	0.903	0.905	0.691	0.665	0.661	0.661
	Encompassing test statistic				Encompassing test statistic			
C			70.747***				540.91***	
R			−0.187				43.198***	
U				-3.507				−0.201
	Realized Utility				Realized Utility			
	5.890	15.079	14.214	13.801	18.178	33.535	34.182	34.197

Appendix Table 2. Out-of-sample Granger causality test results for the U.S. market (trading volume proxied by the volume premium, in-sample length 20 years)

Model C includes a constant only. Model R includes only lagged returns. Model U includes both lags of value-weighted (equal-weighted) market portfolio returns VWMKT (EWMKT) and lags of the corresponding high volume return premium VWHVP (EWHVP). Model W includes lags of stock returns, the volume premium, and market volatility (σ^2). Market volatility is represented by the realized market variance, which in turn is the sum of the squared daily returns in the past three months. The number of lags of each variable (L_y, L_x, and L_z) is the same as in Table 2. Root mean squared forecast errors (RMSFE) and the annualized realized utility have been multiplied by 100 for ease of presentation. All the statistics in the table are based on one-step-ahead recursive forecast errors from the above models for the daily excess returns to the U.S. stock market portfolio.
ENC-NEW is the encompassing test statistic of Clark and McCracken (2001), in which the associated null hypothesis is that forecasts from the model in the first column (H_o) encompass those from the model in the second row (H_A). Thus, the H_o model is preferred to H_A. The critical values of the ENC-NEW test are linearly interpolated from the unpublished Appendix of Clark and McCracken (2001). The symbols ** and *** denote significance at the 5% and 1% levels, respectively.

	Forecasting models				Forecasting models			
	C	R	U	W	C	R	U	W
	Panel A. VWHVP & VWRET, 1983-10				Panel B. EWHVP & EWRET, 1983-10			
	RMSFE				RMSFE			
	1.110	1.119	1.119	1.121	0.903	0.903	0.903	0.910
	Encompassing test statistic				Encompassing test statistic			
C			44.677***				418.73***	
R			2.017**				5.306***	
U				−8.963				−21.609
	Realized Utility				Realized Utility			
	4.579	9.220	9.162	8.975	14.058	27.427	27.898	27.589
	Panel C. VWHVP & VWRET, 1983-99				Panel D. EWHVP & EWRET, 1983-99			
	RMSFE				RMSFE			
	0.880	0.881	0.881	0.882	0.633	0.607	0.607	0.616
	Encompassing test statistic				Encompassing test statistic			
C			91.911***				568.59***	
R			−0.149				3.104***	
U				−3.455				−15.380
	Realized Utility				Realized Utility			
	7.251	17.178	16.944	16.729	17.104	31.828	32.536	31.550

Appendix Table 3. Out-of-sample Granger causality test results for the U.S. market, by firm size (trading volume proxied by turnover)

We sort CRSP stocks into small, medium and large portfolios based on the breakpoints for the low 30%, medium 40%, and high 30% of the ranked values of market capitalization. For each portfolio, we calculate equal-weighted portfolio returns (EWRET) and the corresponding aggregate measure of turnover (EWVOL) for the sample period July 1963 to December 2010. Model C includes a constant only. Model R includes only lagged portfolio returns. Model U includes both lags of portfolio returns and the corresponding turnover. Model W includes lags of portfolio returns, turnover, and market volatility (σ^2), where market volatility is approximated by the realized market variance, which in turn is the sum of the squared daily returns in the past three months. L_y, L_x, and L_z are the numbers of lags on stocks returns, turnover, and market volatility, respectively, in equation (1) in the text. They are selected by minimizing Schwarz's Bayesian information criterion (BIC).

Root mean squared forecast errors (RMSFE) and the annualized realized utility have been multiplied by 100 for ease of presentation. All the statistics in the table are based on one-step-ahead recursive forecast errors from the above models.

ENC-NEW is the encompassing test statistic of Clark and McCracken (2001), in which the associated null hypothesis is that forecasts from the model in the first column (H_o) encompass those from the model in the second row (H_A). Thus, the H_o model is preferred to H_A. The critical values of the ENC-NEW test are linearly interpolated from the unpublished Appendix of Clark and McCracken (2001). The symbol *** denotes significance at the 1% level.

	Forecasting models				Forecasting models			
	C	R	U	W	C	R	U	W
L_y	0	3	3	3	0	3	3	3
L_x	0	0	1	1	0	0	1	1
L_z	0	0	0	3	0	0	0	1
	Panel A. Small stocks, 1963-10				Panel B. Large stocks, 1963-10			
	RMSFE				RMSFE			
	0.838	0.763	0.756	0.756	1.082	1.079	1.079	1.080
	Encompassing test statistic				Encompassing test statistic			
C			2430.5***				378.63***	
R			163.57***				8.817***	
U				45.942***				−9.342
	Realized Utility				Realized Utility			
	43.881	60.167	60.711	60.168	4.743	19.838	19.643	20.106
	Panel C. Small stocks, 1963-99				Panel D. Large stocks, 1963-99			
	RMSFE				RMSFE			
	0.718	0.648	0.640	0.642	0.845	0.822	0.822	0.823

	Encompassing test statistic				Encompassing test statistic			
C			1942.4***				583.20***	
R			147.82***				5.552***	
U				14.119***				-3.610
	Realized Utility				Realized Utility			
	46.655	59.380	60.324	59.793	5.812	25.524	25.674	25.739
	Panel E. Small stocks, 2000-10				Panel F. Large stocks, 2000-10			
	RMSFE				RMSFE			
	1.070	0.984	0.978	0.975	1.502	1.525	1.524	1.527
	Encompassing test statistic				Encompassing test statistic			
C			620.44***				20.018***	
R			34.425***				2.801***	
U				21.493***				−3.628
	Realized Utility				Realized Utility			
	37.288	62.038	61.629	61.060	2.201	6.335	5.320	6.726

Appendix Table 4. The impact of U.S. market turnover on the international markets

Model C includes a constant only. Model W includes lags of value-weighted (equal-weighted) market portfolio returns VWRET (EWRET)), the corresponding aggregate turnover VWVOL (EWVOL), and market volatility (σ^2). Market volatility is approximated by the realized market variance, which in turn is the sum of the squared daily returns in the past three months. The number of lags of each variable is the same as in Table 11. Model WR augments model W with one lag of the U.S. market returns. Model WRV augments model W with one lag of the U.S. market returns and aggregate turnover.
Each entry is the annualized utility level a risk-averse investor can attain by following the above models' stock return forecasts in allocating her/his investment daily between stocks and risk-free bills. These realized utility estimates are computed according to equations (3) and (4) in the text and have been multiplied by 100 for ease of presentation. Numbers in bold indicate that (1) the unrestricted model WRV achieves higher utility than models C, W, and WR; and (2) the latter three models have larger forecast errors and do not encompass model WRV at the 5% significance level.

	Value-weighted portfolios				Equal-weighted portfolios			
	C	W	WR	WRV	C	W	WR	WRV
Canada	3.60	14.17	18.10	17.76	28.24	46.78	49.02	**49.62**
France	0.74	2.39	36.00	35.99	23.05	39.94	47.17	47.20
Germany	2.81	5.92	18.44	18.73	63.14	73.83	78.93	78.53
Italy	−2.48	5.25	19.38	18.21	-3.20	20.11	17.98	20.04
Japan	−0.67	1.66	53.34	53.61	3.68	23.26	53.44	**53.85**
UK	3.28	3.95	36.31	36.26	6.55	36.62	43.74	43.75
Australia	3.13	4.74	46.49	46.49	36.36	55.43	74.39	74.41
Brazil	−7.08	−7.97	5.54	4.33	4.53	22.07	18.94	19.47
China	3.68	2.73	35.88	37.68	13.19	34.28	51.97	52.72
Hong Kong	4.36	12.01	64.39	64.63	20.40	60.09	77.74	**78.48**
India	10.38	26.85	53.43	**56.21**	8.17	66.97	74.79	**74.91**
Korea	4.82	16.69	63.21	62.65	8.93	47.27	70.14	69.30
Russia	12.44	24.70	54.33	50.69	82.96	85.56	96.74	96.36
Singapore	1.11	13.31	41.32	41.50	8.38	42.71	54.44	54.39
South Africa	5.15	15.69	47.14	47.01	27.91	38.56	55.42	55.35
Spain	1.87	−3.05	22.02	21.90	6.25	18.57	28.98	29.03
Switzerland	−1.31	3.75	31.80	30.98	8.44	28.50	39.61	39.49
Taiwan	2.65	6.56	47.35	47.07	4.86	22.00	42.90	43.01

Appendix Table 5. The impact of U.S. market volume premium on the international markets

Model C includes a constant only. Model W includes lags of value-weighted (equal-weighted) market portfolio returns VWRET (EWRET)), the corresponding volume premium VWHVP (EWHVP), and market volatility (σ^2). Market volatility is approximated by the realized market variance, which in turn is the sum of the squared daily returns in the past three months. The number of lags of each variable is the same as in Table 12. Model WR augments model W with one lag of the U.S. market returns. Model WRV augments model W with one lag of the U.S. market returns and volume premium.

Each entry is the annualized utility level a risk-averse investor can attain by following the above models' stock return forecasts in allocating her/his investment daily between stocks and risk-free bills. These realized utility estimates are computed according to equations (3) and (4) in the text and have been multiplied by 100 for ease of presentation. Numbers in bold indicate that (1) the unrestricted model WRV achieves higher utility than models C, W, and WR; and (2) the latter three models have larger forecast errors and do not encompass model WRV at the 5% significance level.

	Value-weighted portfolios				Equal-weighted portfolios			
	C	W	WR	WRV	C	W	WR	WRV
Canada	3.85	14.79	18.41	18.55	23.72	38.32	41.33	41.32
France	0.95	0.23	35.68	**35.84**	7.15	27.65	36.00	36.05
Germany	2.05	5.47	16.68	19.94	9.19	30.19	40.90	40.88
Italy	−1.36	5.83	21.41	22.31	−1.30	18.45	28.90	28.89
Japan	−0.62	3.01	51.36	51.13	1.46	18.30	50.33	49.98
UK	2.33	1.67	34.62	34.66	3.73	31.05	41.76	41.74
Australia	2.68	3.54	46.20	46.63	9.02	31.47	55.01	55.16
Brazil	−2.92	−5.36	9.00	7.71	8.33	17.43	19.40	20.46
China	7.33	7.77	35.81	34.36	13.65	32.08	53.64	52.20
Hong Kong	3.67	11.10	64.61	65.06	8.91	45.83	66.77	**66.95**
India	11.90	26.46	48.11	46.86	7.88	67.76	77.23	76.95
Korea	5.14	15.08	62.80	62.47	6.75	43.31	64.41	64.69
Russia	11.41	9.54	57.63	**58.04**	45.35	51.60	74.14	72.95
Singapore	1.05	12.16	41.32	40.92	0.57	32.96	47.18	46.40
South Africa	5.23	14.35	45.90	45.40	10.28	24.85	44.22	44.11
Spain	0.44	−0.89	23.80	23.69	2.80	12.65	27.01	27.32
Switzerland	−1.51	3.06	28.75	28.65	3.06	24.21	38.57	38.72
Taiwan	2.85	3.34	47.34	47.18	4.30	22.03	44.37	44.06