Overconfidence Bias, Trading Volume and Returns Volatility: Evidence from Pakistan

Muhammad Fayyaz Sheikh and Khalid Riaz

Department of Banking and Finance, G.C. University, Faisalabad, Pakistan
Department of Management Sciences, COMSATS Institute of Information Technology, Islamabad, Pakistan

Abstract: Investor overconfidence has been proposed to explain various anomalous findings in security markets. The theory of investor overconfidence provides testable implications assuming investor overestimation of their abilities and private information and biased self-attributions. High (low) trading activity following market gains (losses) and excessive volatility are the two testable implications among others. We test these implications in Karachi Stock Exchange, Pakistan using multivariate time series analysis. Consistent with overconfidence hypothesis predictions, we find significant positive relationship between current trading activity and past returns after controlling for returns dispersion and returns volatility. However, we do not find significant positive contribution of overconfidence related trading to conditional returns volatility.


Key words: Karachi stock exchange . overconfidence . self attribution . trading volume . conditional volatility . vector autoregression . EGARCH

INTRODUCTION

The stylized facts in security markets such as high trading volume and excessive returns volatility have captured financial economists’ interest since long. Many researchers have developed theoretical models assuming investor overconfidence to justify these stylized facts.

The theoretical models of investor overconfidence provide various testable implications. This paper focuses on the two empirically testable propositions. First, overconfident traders show high (low) trading activity following high (low) observed market returns i.e. Past returns are positively correlated with current turnover. Second, aggressive trading in securities by overconfident traders contributes to observed excessive returns volatility. In theory, these testable implications are obtained under two general assumptions. First, investors are overconfident about the precision and accuracy of their private information. Second, the level of investor overconfidence varies with realized market returns due to biased self-attribution.

Previous studies [13, 27, 30] contend that the people overestimating their trading and investment skills may be more likely to choose their career as traders or they may trade actively on their own. Moreover, these overconfident traders can survive and dominate the markets in the longer horizon [1, 7, 13, 18, 23, 35]. Therefore, if most investors suffer from overconfidence and if overconfidence is a systematic cognitive bias, it is possible to trace investor overconfidence by analyzing the market level trading behavior (investors’ aggregated trading behavior).

Many studies [2, 3, 16, 34], using market level data, have found a positive relationship between current trading level and past returns that is consistent with overconfidence theory. These studies test the implication of investor overconfidence related to trading volume within the framework of vector autoregression (VAR). There are other studies [5, 14, 15, 29] which analyze the predictions of overconfidence theory by focusing on trading activity of individual investors. These studies find positive link of trading activity with past returns using unique datasets consist of individual investors’ accounts. Since data on individual investor accounts is not available in Pakistan we stick only to market level data to test the overconfidence predictions.

The researchers [5, 15] analyzing individual investors’ portfolios posit that only high portfolio returns can lead investors to buy high risky stocks, therefore, dynamic changes in investor overconfidence can only trigger from their
past portfolio returns rather than from prior market returns. However, models of overconfident investors [13, 27] tell that that overconfident investors trade aggressively following market gains especially in bull market. A recent study [3] tested the predictions of overconfidence models and finds that both individual and institutional investors trade more aggressively following market gains. The findings of the study [3] also indicate the investor tendency to trade more in riskier securities following market gains.

The models of overconfidence [13, 34] argue that overconfidence is a market-wide phenomenon and can be traced at market level, while other studies [5, 15] argue that level of overconfidence varies with the individual portfolio returns rather than market returns. Therefore, implications of investor overconfidence should be tested at both levels i.e., at market level and at individual portfolio level. However, data on individual investor trading accounts is not available in Pakistan; therefore we test the implications of overconfidence at market level only using market level data.

The positive relationship between trading and past returns is also consistent with the disposition effect (The disposition effect: propensity for investors to sell profitable securities too early and hold losing securities too long) [28, 33]. However, researchers [34] argue that overconfidence works as a driving force for disposition effect as overconfidence persuades investors to trade asymmetrically between winners and losers. They [34] differentiate overconfidence from the disposition effect in the following two ways. First, the disposition effect describes an investor’s attitude towards a specific stock [8, 28], while, overconfidence influences the securities market in general. Second, the disposition effect accounts for the incentive for one side of a transaction only, whereas, overconfidence can support both sides of a given transaction.

We consider Karachi Stock Exchange, Pakistan (KSE) to test overconfidence hypothesis due to its distinctive characteristics. KSE is believed to be a small, opaque and highly volatile emerging market of Asia. It is considerably more active than the markets of the same size. KSE was declared as the “Best Performing Stock Market of the World” by Business Week in 2002. It has provided high returns with relatively greater variation during the last decade [19]. KSE has very high turnover ratio; in fact, it was ranked first and third in the world with respect to turnover in 2003 and 2006 respectively (Source: Standard & Poor’s, Global Stock Markets Factbook (2004, 2007)). However, KSE is a highly concentrated market and has less ability to mobilize the new investment [19].

The relationship between trading volume and lagged returns is found to be stronger for developing, opaque, corrupted and volatile economies than for developed and transparent economies [16]. Since Pakistani bourse satisfies these conditions and in addition, investors in Pakistani securities face additional risk due to uncertainties related to political instability and war against terrorism. Therefore, it is interesting to analyze the investor behavior in such a market. This study fills this gap by investigating the implications of overconfidence hypothesis related to trading volume and returns volatility in Karachi Stock Exchange (KSE), the main equity market in Pakistan, for period November, 1999 to October, 2010.

We employ vector autoregression (VAR) technique to examine the link between current turnover and past returns. To account for other possible explanations to lead-lag association between return and turnover, such as portfolio rebalancing and heterogeneous interpretation of informational events, we introduce returns volatility and cross sectional dispersion in our VAR analysis.

For implication related to returns volatility, we first decompose the turnover into two components: one associated with overconfidence and other not associated with it. Then both the components are incorporated in EGARCH and TARCH models to see any effects on conditional volatility.

The study proceeds as follows. Section 2 discusses hypothesis development and the empirical framework. Data and variables are described in section 3. Section 4 reports empirical results and followed by the conclusions in section 5.

HYPOTHESES AND EMPIRICAL FRAMEWORK

**Investor overconfidence and trading volume:** Overconfident investors mistakenly attribute past market returns to their trading and valuation skills. They overestimate precision and accuracy of their information. Consequently, they trade more aggressively subsequent to high market returns to maximize their utility [1, 13, 27]. Investor overconfidence decreases subsequent to market losses but it takes relatively long time due to slow learning process of overconfident traders [7, 13]. This indicates that increased trading volume due to investor overconfidence will persist for a longer period of time. However, theory does not provide any time frame for the presence of overconfidence related trading activity.
Based on prediction of theoretical models of overconfidence, we can state our first hypothesis as follows:

**Hypothesis 1:** Investors are overconfident, therefore, the current trading activity is positively related to past market returns and this relationship is persistent for quite a longer period of time.

We use Vector autoregression (VAR) and associated impulse response functions to test this hypothesis. The feedback contemporaneous relationship between trading volume and returns in Pakistan [21, 25] and ability of VAR to capture evolution of and the interdependencies between multiple time series led us to choose VAR for study of overconfidence hypothesis following many studies [3, 16, 34].

The lead-lag relationship between trading volume and realized returns may be conditioned by other factors. For example, portfolio rebalancing trades can occur if the spread between the individual stock returns widens, reminding investors to rebalance their portfolios. We included cross sectional returns dispersion in VAR system to capture the effect of any potential trading activity associated with portfolio rebalancing or diversification. Cross sectional returns dispersion is standard deviation of returns for all stocks in market portfolio at one point in time. It is associated with degree of realized idiosyncratic risk of securities (Idiosyncratic risk: Also known as unsystematic risk; the risk of change in price that is unique to an individual asset or security, derived from asset’s specific characteristics).

Likewise, heterogeneous or diverse explanations to informational events can lead to association between returns, volatility and trading volume. A large amount of research [4, 6, 12, 17, 20, 24, 25, 32] empirically establishes the concurrent and lead-lag relationships between returns, volatility and trading volume. To account for such relationships, we introduced returns volatility as another control variable in VAR system following other studies [16, 34]. Without accounting for volatility of returns, the effect of returns on turnover can be offset by effect of volatility, leading to a blurry impact of returns on turnover.

The standard form of VAR is:

\[
Y_t = \alpha + \sum_{i=1}^{P} A_i Y_{t-i} + \sum_{j=0}^{S} B_j X_{t-j} + \epsilon_t
\]

where \(Y_t\) is an \(n\times1\) vector of the endogenous variables at time \(t\), \(X_t\) is a vector of exogenous variables and \(\epsilon_t\) is an \(n\times1\) vector of residuals. The coefficient matrices \(A_i\) and \(B_j\) estimate the time-series associations between the endogenous and exogenous variables in the system. \(P\) is the number of lags included for endogenous variables and \(S\) is number lags included for exogenous variables.

We also estimated impulse response functions IRF associated with VAR. The IRF traces out the responses of dependent variables in the VAR system to the innovations in residual vector \(\epsilon_t\). The IRFs can trace out the impact of such innovations for several periods in the future, providing a simple picture of the relationships.

Formal overconfidence models do not precisely advise the time frame for lead-lag association between turnover and return for empirical studies [14]. Therefore, following other studies [16, 34] we based the lag selection in VAR on Akaike Information Criterion (AIC) and Schwartz Information Criterion (SIC) (AIC and SIC take the following forms: Where \(m\) is the number of parameters estimated, \(T\) is the number of observations and \(O\) is the log likelihood function value using the \(m\) estimated parameters). Our VAR model (eq. 1) becomes of the following form:

\[
\begin{bmatrix}
\text{TURNN}_t \\
\text{RETNN}_t \\
\end{bmatrix} = \begin{bmatrix}
\text{eNN}_t \\
\text{eRETNN}_t \\
\end{bmatrix} + \sum_{i=1}^{2} A_i \begin{bmatrix}
\text{TURNN}_{t-i} \\
\text{RETNN}_{t-i} \\
\end{bmatrix} + \sum_{j=0}^{5} B_j \begin{bmatrix}
\text{EGVOLNN}_{t-j} \\
\text{DISPN}_{t-j} \\
\end{bmatrix} + \begin{bmatrix}
\text{eNN}_t \\
\text{eRETNN}_t \\
\end{bmatrix}
\]

(2)

where \(\text{TURNN}\) (proxy for trading activity) and \(\text{RETNN}\) (market return) are endogenous variables and \(\text{EGVOLNN}\) (returns volatility) and \(\text{DISPN}\) (cross sectional dispersion of returns) are exogenous variables.

In the impulse response function associated with above VAR models (Eq. 2), the innovation in one residual, say \(e_{\text{RETNN}}\), would immediately modify the existing value of \(\text{RETNN}\) but would also influence future values of \(\text{RETNN}\) and \(\text{TURNN}\) because past values of \(\text{RETNN}\) exist in both equations through the coefficient matrix \(A_i\). So, to examine the prediction of the overconfidence hypothesis, one sample standard deviation shock of return residual \(e_{\text{RETNN}}\) can be applied to track how turnover responds to innovation in \(e_{\text{RETNN}}\) over time. Hence, despite the dense parameter structure of VAR, we can have a simple picture of relationship between endogenous variables through the associated impulse response function.
Excessive trading of overconfident investors and returns volatility: Overconfident investors trade excessively they reveal a large amount of their private information to public, hence, increase informativeness and price volatility [1]. Many studies [7, 27, 30] have developed arguments of increased volatility due to overconfident traders. Based on theory of investor overconfidence, we state our hypothesis as follows:

Hypothesis 2: Excessive trading of overconfident traders in stocks positively contributes to the observed returns volatility.

Various studies [1, 2] categorize the entire trading volume as respective contribution of liquidity traders, market makers, rational traders and overconfident traders. So, we test our hypothesis of increased volatility due to overconfident investors by separating the contribution of overconfident traders from the entire trading volume. Initially, we decomposed the trading volume into two components using following regression:

$$Turn_t - a + \sum_{j=1}^{p} \beta_j R_{t-j} + \varepsilon_t - \left[ \sum_{j=1}^{p} \beta_j R_{t-j} \right] + \left[ a + \varepsilon_t \right] - OVER_t + NONOVER_t$$  \hspace{1cm} (3)

The variable \(Turn\) is detrended turnover and \(R_{t-j}\) is past return. The sum of residuals terms and constant \((NONOVER)_t\) is that component of turnover which is irrelevant to investors’ overconfidence. The difference between trading volume and \(NONOVER\) is the part of trading volume related to investors’ overconfidence \((OVER)_t\) due to past market returns.

Both variables, \(NONOVER\) and \(OVER\), were then incorporated in the variance equations of EGARCH and TARCH models. Both models account for the asymmetric response of volatility to the sign of return innovations. Specifically, they allow for the possibility that returns volatility would rise more in response to a negative return shock of a given magnitude than to positive return shock of the same magnitude.

Our EGARCH and TARCH models take the following forms:

ARMA (1, 1)-EGARCH (1, 1) model:

$$R_t = \alpha + \beta R_{t-1} + \gamma \varepsilon_{t-1} + \varepsilon_t \zeta_t (\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots, R_{t-1}, R_{t-2}, \ldots) \sim GED(0, h_t)$$  \hspace{1cm} (4)

$$\ln(h_t) = \omega + \theta \left[ \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \right] + \alpha_1 \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} + \beta \ln(h_{t-1}) + f_1 NONOVER_t + f_2 OVER_t$$

and ARMA (1, 1)-TARCH (1, 1) model:

$$R_t = \alpha + \beta R_{t-1} + \gamma \varepsilon_{t-1} + \varepsilon_t \zeta_t (\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots, R_{t-1}, R_{t-2}, \ldots) \sim GED(0, h_t)$$  \hspace{1cm} (5)

$$h_t = \omega + \theta \left[ \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \right] + \alpha_1 \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} + \beta \ln(h_{t-1}) + f_1 NONOVER_t + f_2 OVER_t$$

In EGARCH model parameter \(k\) captures the asymmetric response of volatility to market return. if \(k < 0\), then there is asymmetric relationship between volatility of return and return. In TARCH model, \(d_{t-1}\) is a dummy that takes the value “one” if \(e_{t-1} < 0\) and takes the value “zero” if \(e_{t-1} \geq 0\). Hence, if \(0 > \theta > 0\), the negative \(e_{t-1}\) innovations would have greater effect on volatility than positive \(e_{t-1}\) innovations. The distribution for conditional errors is assumed to be a Generalized Error Distribution (GED) [26].

The empirical frame work suggested in Eqs. (4) and (5) along with Eq. (3) differentiates the trading activity because of investor overconfidence from other factors that contribute to volatility of market returns. The \(f_2\) parameter catches the effect of overconfidence on returns volatility, whereas, the \(f_1\) parameter represents the effect of other possible explanations of volatility of returns. If the overconfidence led trading activity adds more to the conditional volatility of returns then condition \(f_2 > f_1 > 0\) would hold.

DATA

The data set consists of weekly and monthly observations on Karachi Stock Exchange (KSE) from November 1999 to October 2010. Studies [13, 16, 27, 34] argue that change in investor overconfidence can occurs on weekly,
monthly or annual basis; therefore, we analyze weekly and monthlydata. Daily data for market as a whole and 43 selected listed companies is collected first and then transformed into weekly (Wednesday close to Wednesday close) and monthly (end of month to end of month) frequencies (If there were holidays in a week or in a month, the next working day is treated as the next day). The data is compiled from official website of Karachi stock exchange (www.kse.com.pk), www.businessrecorder.com, www.khistocks.com and monthly reports of State Bank of Pakistan.

Following other studies [2, 3, 24, 34], we scaled the total trading value by total market capitalization to form market turnover as proxy to trading activity in KSE. Specifically, we divided daily market trading value by daily contemporaneous total market capitalization and cumulated to form monthly and weekly turnover respectively. We log transform the turnover. It appears to be trend stationary using ADF unit root test (ADF test specification with time trend:

\[
\Delta \ln \frac{t}{t-1} = \alpha + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \varepsilon_t
\]

where \( \varepsilon_t \) is series in question and the null hypothesis is that the series contains unit root (i.e., \( \gamma = 0 \)). We then detrended the log turnover series using following selected model for both weekly and monthly observations. The model selection was based on econometric theory [9].

\[
\operatorname{Lnturn}_t = \alpha + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \varepsilon_t
\]

(6)

where \( \varepsilon_t \) is log turnover series. The residuals from this model represent detrended log turnover series (TURN).

Table 1 shows the monthly and Weekly descriptive statistics in Panels A and B respectively. Note that due to log transformation detrended log turnover represents the percentage change in the turnover with respect to the trend in turnover.

The returns on KSE 100 index are used as proxy for market returns (RET). KSE 100 index is a value weighted index of 100 companies. It captures over 90\% of the total market capitalization of the companies listed on the Karachi Stock Exchange. KSE-100 is a total return index that adjusts dividends, bonus issues and right issues. We calculated the index returns as difference of natural log of ending value of index on monthly and weekly basis.

\[
R_t = \ln(P_t) - \ln(P_{t-1}) = \ln(P_t/P_{t-1})
\]

(7)

were \( R_t \) is market return for period \( t \), \( P_t \) is current period closing value of index and \( P_{t-1} \) is previous period closing value of index.

Following large body of literature returns volatility is estimated using an EGARCH model. Non-synchronous trading of securities causes daily portfolio returns to be auto-correlated, especially at lag one [10, 31]. Therefore we used AR(1) specification for meanequation in EGARCH model. Specifically, We fitted the following AR(1)-EGARCH (1,1) model to daily KSE 100 index returns to estimate monthly and weekly volatilities.

\[
R_t = \alpha + \beta R_{t-1} + \varepsilon_t
\]

(8)

\[
\varepsilon_t | (R_{t-1}, R_{t-2}, \ldots, R_{t-k}, \ldots) \sim t(\alpha, h_t)
\]

\[
\ln(h_t) = \omega + \gamma \frac{|\varepsilon_{t-1}|}{\sqrt{h_{t-1}}} + \theta \ln(h_{t-1})
\]

(9)

where \( R_t \) represents current market return and \( R_{t-1} \) is previous day market return. Note that left hand side of variance equation is natural log of the variance \( (h_t) \). Therefore, estimates of the conditional variances are guaranteed to be non-negative. Estimated daily variances from AR(1)-EGARCH (1,1) model are cumulated to form monthly and weekly variances. Square roots of these variances represent the volatility of returns (EGVOL).

Cross sectional dispersion (DISP) is estimated using observations on 43 selected companies. These 43 companies are selected from KSE100 Index lists as on October 01, 2009 and April 01, 2010 (old and new sectors respectively) on the basis of following criteria. First, we selected the largest market capitalization company from

\[
\gamma = 0)). We then detrend the log turnover series using following selected model for both weekly and monthly observations. The model selection was based on econometric theory [9].

\[
\operatorname{Lnturn}_t = \alpha + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \varepsilon_t
\]

(6)

were \( \operatorname{Lnturn}_t \) is log turnover series. The residuals from this model represent detrended log turnover series (TURN).

Table 1 shows the monthly and Weekly descriptive statistics in Panels A and B respectively. Note that due to log transformation detrended log turnover represents the percentage change in the turnover with respect to the trend in turnover.

The returns on KSE 100 index are used as proxy for market returns (RET). KSE 100 index is a value weighted index of 100 companies. It captures over 90\% of the total market capitalization of the companies listed on the Karachi Stock Exchange. KSE-100 is a total return index that adjusts dividends, bonus issues and right issues. We calculated the index returns as difference of natural log of ending value of index on monthly and weekly basis.

\[
R_t = \ln(P_t) - \ln(P_{t-1}) = \ln(P_t/P_{t-1})
\]

(7)

were \( R_t \) is market return for period \( t \), \( P_t \) is current period closing value of index and \( P_{t-1} \) is previous period closing value of index.

Following large body of literature returns volatility is estimated using an EGARCH model. Non-synchronous trading of securities causes daily portfolio returns to be auto-correlated, especially at lag one [10, 31]. Therefore we used AR(1) specification for meanequation in EGARCH model. Specifically, We fitted the following AR(1)-EGARCH (1,1) model to daily KSE 100 index returns to estimate monthly and weekly volatilities.

\[
R_t = \alpha + \beta R_{t-1} + \varepsilon_t
\]

(8)

\[
\varepsilon_t | (R_{t-1}, R_{t-2}, \ldots, R_{t-k}, \ldots) \sim t(\alpha, h_t)
\]

\[
\ln(h_t) = \omega + \gamma \frac{|\varepsilon_{t-1}|}{\sqrt{h_{t-1}}} + \theta \ln(h_{t-1})
\]

(9)

where \( R_t \) represents current market return and \( R_{t-1} \) is previous day market return. Note that left hand side of variance equation is natural log of the variance \( (h_t) \). Therefore, estimates of the conditional variances are guaranteed to be non-negative. Estimated daily variances from AR(1)-EGARCH (1,1) model are cumulated to form monthly and weekly variances. Square roots of these variances represent the volatility of returns (EGVOL).

Cross sectional dispersion (DISP) is estimated using observations on 43 selected companies. These 43 companies are selected from KSE100 Index lists as on October 01, 2009 and April 01, 2010 (old and new sectors respectively) on the basis of following criteria. First, we selected the largest market capitalization company from
Table 1: Descriptive statistics

<table>
<thead>
<tr>
<th></th>
<th>Turnover</th>
<th>TURN</th>
<th>RET</th>
<th>EGVOL</th>
<th>DISP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Monthly descriptive statistics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>21.34</td>
<td>0.00</td>
<td>1.66</td>
<td>6.68</td>
<td>9.71</td>
</tr>
<tr>
<td>Median</td>
<td>17.86</td>
<td>14.96</td>
<td>1.93</td>
<td>6.17</td>
<td>9.03</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>15.53</td>
<td>106.76</td>
<td>9.21</td>
<td>2.09</td>
<td>4.30</td>
</tr>
<tr>
<td>Maximum</td>
<td>75.57</td>
<td>140.66</td>
<td>24.11</td>
<td>12.58</td>
<td>30.36</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.00</td>
<td>-806.53</td>
<td>-44.88</td>
<td>2.53</td>
<td>2.45</td>
</tr>
<tr>
<td>Observations</td>
<td>132.00</td>
<td>132.00</td>
<td>132.00</td>
<td>132.00</td>
<td>132.00</td>
</tr>
<tr>
<td><strong>Panel B: Weekly descriptive statistics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>4.91</td>
<td>0.00</td>
<td>0.37</td>
<td>3.15</td>
<td>4.35</td>
</tr>
<tr>
<td>Median</td>
<td>4.01</td>
<td>15.14</td>
<td>0.63</td>
<td>2.88</td>
<td>3.85</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>3.91</td>
<td>122.70</td>
<td>3.87</td>
<td>1.16</td>
<td>2.19</td>
</tr>
<tr>
<td>Maximum</td>
<td>22.75</td>
<td>178.51</td>
<td>15.57</td>
<td>8.13</td>
<td>20.47</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.00</td>
<td>-944.29</td>
<td>-19.64</td>
<td>0.80</td>
<td>0.00</td>
</tr>
<tr>
<td>Observations</td>
<td>573.00</td>
<td>573.00</td>
<td>573.00</td>
<td>573.00</td>
<td>573.00</td>
</tr>
</tbody>
</table>

All the variables are reported in percentage (%) points

Table 2: ADF unit root test

<table>
<thead>
<tr>
<th></th>
<th>ADF statistic</th>
<th>a2</th>
<th>PP statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Tests for monthly observations</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Turnover</td>
<td>-4.0997</td>
<td>-2.6061</td>
<td>-4.0325</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.0081)</td>
<td>(0.0103)</td>
<td>(0.0099)</td>
</tr>
<tr>
<td>TURN</td>
<td>-6.6310</td>
<td>------</td>
<td>-4.3837</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.0000)</td>
<td></td>
<td>(0.0000)</td>
</tr>
<tr>
<td>RET</td>
<td>-9.7801</td>
<td>------</td>
<td>-9.7803</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.0000)</td>
<td></td>
<td>(0.0000)</td>
</tr>
<tr>
<td>EGVOL</td>
<td>-7.3388</td>
<td>------</td>
<td>-7.0813</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.0000)</td>
<td></td>
<td>(0.0000)</td>
</tr>
<tr>
<td>DISP</td>
<td>-5.1620</td>
<td>------</td>
<td>-8.1024</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.0000)</td>
<td></td>
<td>(0.0000)</td>
</tr>
<tr>
<td><strong>Panel B: Tests for weekly observations</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Turnover</td>
<td>-4.5070</td>
<td>-2.4596</td>
<td>-4.9008</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.0016)</td>
<td>(0.0142)</td>
<td>(0.0003)</td>
</tr>
<tr>
<td>TURN</td>
<td>-4.9533</td>
<td>------</td>
<td>-5.4368</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.0000)</td>
<td></td>
<td>(0.0000)</td>
</tr>
<tr>
<td>RET</td>
<td>-13.8862</td>
<td>------</td>
<td>-21.8203</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.0000)</td>
<td></td>
<td>(0.0000)</td>
</tr>
<tr>
<td>EGVOL</td>
<td>-8.8225</td>
<td>------</td>
<td>-8.8476</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.0000)</td>
<td></td>
<td>(0.0000)</td>
</tr>
<tr>
<td>DISP</td>
<td>-8.8981</td>
<td>------</td>
<td>-19.6974</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.0000)</td>
<td></td>
<td>(0.0000)</td>
</tr>
</tbody>
</table>

each sector of KSE then looked for availability of data for the whole investigation period. If data was not available, the company was dropped and the next largest company from the same sector was given a chance to be selected on data availability criterion.
Table 3: VAR estimation

<table>
<thead>
<tr>
<th>Panel A: Monthly VAR estimation</th>
<th>TURNt Coefficient</th>
<th>RETt Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>TURNt-1</td>
<td>1.0817***</td>
<td>0.0016</td>
</tr>
<tr>
<td>TURNt-2</td>
<td>-0.5556***</td>
<td>0.0344**</td>
</tr>
<tr>
<td>TURNt-3</td>
<td>0.1447</td>
<td>-0.0190</td>
</tr>
<tr>
<td>RETt-1</td>
<td>0.1010</td>
<td>0.0967</td>
</tr>
<tr>
<td>RETt-2</td>
<td>-0.0329</td>
<td>-0.0311</td>
</tr>
<tr>
<td>RETt-3</td>
<td>1.1081*</td>
<td>-0.0744</td>
</tr>
<tr>
<td>EGVOLt</td>
<td>9.1336***</td>
<td>-2.1469***</td>
</tr>
<tr>
<td>EGVOLt-1</td>
<td>-9.8610**</td>
<td>1.1566**</td>
</tr>
<tr>
<td>EGVOLt-2</td>
<td>-2.4332</td>
<td>-0.5268</td>
</tr>
<tr>
<td>DISPt</td>
<td>0.8647</td>
<td>0.1060</td>
</tr>
<tr>
<td>DISPt-1</td>
<td>-2.1418</td>
<td>0.5160**</td>
</tr>
<tr>
<td>DISPt-2</td>
<td>4.6580***</td>
<td>-0.3295</td>
</tr>
<tr>
<td>Constant</td>
<td>-13.1874</td>
<td>8.6370**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Weekly VAR estimation</th>
<th>TURNt Coefficient</th>
<th>RETt Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>TURNt-1</td>
<td>0.8961***</td>
<td>0.0030</td>
</tr>
<tr>
<td>TURNt-2</td>
<td>0.1056*</td>
<td>-0.0080*</td>
</tr>
<tr>
<td>TURNt-3</td>
<td>-0.0715*</td>
<td>0.0080**</td>
</tr>
<tr>
<td>RETt-1</td>
<td>1.1413**</td>
<td>-0.0666</td>
</tr>
<tr>
<td>RETt-2</td>
<td>-2.4355***</td>
<td>0.2047***</td>
</tr>
<tr>
<td>RETt-3</td>
<td>-0.2930</td>
<td>0.0262</td>
</tr>
<tr>
<td>EGVOLt</td>
<td>18.1693***</td>
<td>-2.0498***</td>
</tr>
<tr>
<td>EGVOLt-1</td>
<td>-19.5357***</td>
<td>1.4038***</td>
</tr>
<tr>
<td>DISPt</td>
<td>5.4581***</td>
<td>0.1335*</td>
</tr>
<tr>
<td>DISPt-1</td>
<td>-2.2417**</td>
<td>0.1210</td>
</tr>
<tr>
<td>Constant</td>
<td>-8.9570</td>
<td>1.2342**</td>
</tr>
</tbody>
</table>

Significance: 10% (*), 5% (**) and 1% (***)

The focus of the selection of 43 companies was to have such a portfolio that mimic KSE 100 index in term of returns. Therefore, we select one company from each sector of KSE. A value weighted portfolio is constructed from these 43 stocks. The returns of 43-portfolio were highly correlated with KSE100 Index returns (correlation coefficients: 0.9568 and 0.9492 for monthly and weekly observations respectively). The t-test failed to reject the hypothesis of equal means of KSE100 index returns and 43-portfolio returns over time. Monthly and Weekly cross sectional dispersion is calculated as follows:

\[
disp_t = \sqrt{\sum_{i=1}^{43} w_{it}(R_{it} - R_{mt})^2}
\]

\[
R_{it} = L_{it}(P_{it}) - L_{t-1}(P_{i(t-1)})
\]

where \( \disp_t \) is cross sectional dispersion at time \( t \), \( w_{it} \) is weight of each security in 43-portfolio at time \( t \), \( R_{it} \) is individual security return at time \( t \), \( R_{mt} \) is market return at time \( t \) proxied by KSE100 index returns, \( R_{it} \) is current period closing price of individual security at time \( t \) and \( P_{i(t-1)} \) is lag one closing price of individual security.

The four time series, detrended turnover, return, volatility and dispersion appeared to be stationary when formally tested using ADF test of unit root. Table 2 presents the results of ADF tests.
Table 4: Relationship between the conditional volatility and turnover

\begin{align*}
R_t &= \alpha + \beta R_{t-1} + \gamma \ln(x_{t-1}) + \epsilon_t \\
\ln(x_t) &= \omega + \eta \ln(x_{t-1}) + k \frac{R_{t-1}}{\sqrt{\sigma_{t-1}}} + \delta \ln(x_{t-1}) + f_1 \text{NONOVER}_t + f_2 \text{OVER}_t
\end{align*}

<table>
<thead>
<tr>
<th>Weekly results</th>
<th>Monthly results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>p-value</td>
</tr>
<tr>
<td>α</td>
<td>0.5876</td>
</tr>
<tr>
<td>β</td>
<td>0.8022</td>
</tr>
<tr>
<td>γ</td>
<td>-0.6220</td>
</tr>
<tr>
<td>Variance Equation</td>
<td></td>
</tr>
<tr>
<td>ω</td>
<td>0.2486</td>
</tr>
<tr>
<td>η</td>
<td>0.5120</td>
</tr>
<tr>
<td>k</td>
<td>-0.1340</td>
</tr>
<tr>
<td>δ</td>
<td>0.7435</td>
</tr>
<tr>
<td>f1</td>
<td>0.0001</td>
</tr>
<tr>
<td>f2</td>
<td>0.0020</td>
</tr>
</tbody>
</table>

Panel B: ARMA (1,1)-TARCH (1,1) Model

\begin{align*}
R_t &= \alpha + \beta R_{t-1} + \gamma \ln(x_{t-1}) + \epsilon_t \\
\ln(x_t) &= \omega + \eta \ln(x_{t-1}) + \theta \ln(x_{t-1}) + \delta \ln(x_{t-1}) + f_1 \text{NONOVER}_t + f_2 \text{OVER}_t
\end{align*}

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>p-value</th>
<th>Coefficient</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>0.3619</td>
<td>0.0325</td>
<td>1.8550</td>
</tr>
<tr>
<td>β</td>
<td>0.7730</td>
<td>-0.2536</td>
<td>0.2450</td>
</tr>
<tr>
<td>γ</td>
<td>-0.5506</td>
<td>0.0000</td>
<td>0.4646</td>
</tr>
<tr>
<td>Variance Equation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω</td>
<td>4.6323</td>
<td>0.0000</td>
<td>6.5194</td>
</tr>
<tr>
<td>η</td>
<td>0.3583</td>
<td>0.0355</td>
<td>-0.0221</td>
</tr>
<tr>
<td>θ</td>
<td>0.1234</td>
<td>0.6809</td>
<td>-0.0625</td>
</tr>
<tr>
<td>δ</td>
<td>0.3216</td>
<td>0.0014</td>
<td>1.0022</td>
</tr>
<tr>
<td>f1</td>
<td>0.0050</td>
<td>0.0001</td>
<td>-0.0328</td>
</tr>
<tr>
<td>f2</td>
<td>-0.0573</td>
<td>0.4171</td>
<td>-0.2075</td>
</tr>
</tbody>
</table>

**EMPIRICAL RESULTS**

**VAR estimation and test results:** Table 3 summarizes the results VAR for detrended turnover (TURN) and market returns (RET), for monthly and weekly observations. Panels A and B report monthly and weekly results respectively. The table is organized by columns for VAR equations associated with the dependent variables and by rows for lagged dependent variable and exogenous variable. The selection criteria (AIC and SIC) for deciding upon number lags to include in VAR system (Eq. 2) led us to select P = 3 and S = 2 for monthly data. While, for weekly data P = 3 and S = 1.

**Overconfidence and turnover: Evidence from generalized impulse response functions:** We used generalized impulse response (GIR) analysis [22] to have a simple picture of lead-lag relationship of turnover and returns. The generalized impulse response (GIR) analysis lets the data decide the correlation structure for shocks across variables and makes the inferences independent of the order of placement of variables in VAR. Variable order is particularly important in our study, as an orthogonalized IRF with turnover ordered before returns could miss intra period relationships between past returns and current turnover (A separate sensitivity analysis (unreported), we estimated orthogonalized IRFs and found similar results to those we report).
Figure 1 and 2 respectively show the monthly and weekly responses of turnover (TURN) and return (RET) to a one standard deviation shock in TURN and RET along with two-standard error bands. Due to log transformation, turnover responses represent the percentage increase in turnover relative to non-shock turnover level (i.e., trend). While return responses represent the difference in market return relative to the mean of market return.

**Monthly perspective:** Consistent with our Hypothesis 1, Panel A of Fig. 1 shows a large and persistent positive response of turnover to a market return shock over more than 6 months. This is the key finding of the study. The accumulated response of turnover to market return shock for first six months is about 94% which is economically and statistically significant phenomenon when compare to average levels. This result provides evidencethat market returns impact investor confidence and subsequent trading activity in Karachi Stock Exchange (The similar impulse response functions using raw log turnover (nondetrended) declines but remains persistently positive over more than 24 months (unreported)).

The persistent positive response of turnover to market return shock over a longer period supports the overconfidence hypothesis. The realized returns impact trading activity for several periods of time, leading to possible long-term trends in turnover.

In Panel B, the impulse response of market return to turnover shock at first month indicates that a one standard deviation shock to market turnover results in about 3% increase in the next month's return relative to mean returns. The response is statistically significant for up to three months, indicating turnover could predict future market returns at KSE-a contradiction to efficient market hypothesis.

**Weekly perspective:** Panel A of Fig. 2 shows a large and persistent positive response of turnover to a market returns shock over more than 10 weeks, which is consistent with the trading volume implications of formal overconfidence hypothesis. The accumulated response of turnover to market return shock for first four weeks is 40% which
Fig. 2: Weekly impulse response functions with two standard error band

In Panel B, we can see response of return to turnover statistically different from zero at various weeks. This finding is contrast to the weak-form market efficiency which holds that past events cannot predict future returns. So, KSE has an anomaly of turnover predicting future returns in weekly data as well

**Overconfidence and conditional volatility: results from EGARCH and TARCH models:** Our second hypothesis is that trading related to overconfidence increases returns volatility. We decompose the turnover into two components, one related to overconfidence and the other not related to it (To decompose turnover we set P = 2 for weekly observations and P = 4 for monthly observations in Eq. (3) based on standard t-test along with F-test, Wald test and LM test to decide upon any omitted or redundant regressor). These two components were incorporated into ARMA(1,1)-EGARCH(1,1) model (Eq. 4) and ARMA(1,1)-TARCH model (Eq. 5) to investigate the relationship between overconfidence led trading and conditional volatility of market returns.

Table 4 summarizes the monthly and weekly estimates of relationship between conditional volatility and turnover. Panel A presents results of EGARCH model while Panel B presents results of TARCH model.

In both models, there is no support for hypothesis of increased volatility due to trading related to investor overconfidence as the coefficient is either insignificant or it is significant with negative sign. Therefore, our hypothesis of increased returns volatility due to trading related to investor overconfidence is rejected.

Concerning asymmetric relationship between returns and volatility of returns (leverage effect), there is no supporting evidence as the coefficients k and θ are consistently insignificant in all the cases.

**ROBUST CHECKS**

To test the robustness of our results we first replicate our study using raw trading volume (shares traded) instead to turnover. Second, we replicate the study using returns volatilities estimated from different competing methods

1746
such as various ARCH specifications and following other studies [11, 34]. But, our results were qualitatively similar to what we report.

Regarding our EGARCH and TARCH models to test the impact of overconfidence related turnover on returns volatility, we also employ other specifications with different lags and with various distributional assumptions but could not find any significant deviations in the results. Therefore we stick to more parsimonious models. Both models present no remaining heteroscedasticity. Moreover, we employ various other symmetric models of conditional volatility from ARCH family, however, results remained unaltered qualitatively.

CONCLUSIONS

Theories of overconfidence explain number of stylized facts in financial markets. This paper investigates the testable implications of overconfidence theory. One implication of theory is that overconfident investors trade more aggressively. Assuming that past returns lead investors to become overconfident, we tested the hypothesis that turnover was positively related to past returns. Another implication of theory is that trading by overconfident investors contributes to the returns volatility. Accordingly, we tested the hypothesis that returns volatility is positively associated with overconfidence related turnover. The two hypotheses were tested in a multivariate time series framework using market level data from Karachi Stock Exchange for the period November 1999 to October 2010.

For the first hypothesis, the Vector Autoregression (VAR) and associated impulse response functions (IRF) are employed. We found significant positive response of turnover to market return shock after controlling for concurrent and lagged return dispersion and returns volatility. This response was persistent for quite a long time for both monthly and weekly IRF. Thus, results confirm the presence of investor overconfidence at KSE. This is the key finding of this study.

Consistent with previous studies, we also found significant contemporaneous positive relationship between turnover and returns volatility in our VAR analysis. Regarding portfolio rebalancing, we found that investors took two months to respond to cross sectional variations in security prices to rebalance their portfolios for eliminating unsystematic risk. Moreover, returns predictability based on past turnover in the VAR and associated impulse response function analysis was found. This violates the strict market efficiency hypothesis—a finding that has been documented by many financial economists.

For the second hypothesis, the turnover was decomposed into two parts: one associated with past market return (overconfidence) and the other unassociated with past market returns. The interaction between the conditional volatility of market returns and these two parts of turnover was investigated using EGARCH and TARCH models. Overall there was no evidence of positive contribution of overconfidence related trading to conditional volatility of returns for both weekly and monthly data.

The presence of predictable positive returns persisting for several periods, following a positive turnover shock, is indicative of some degree of market inefficiency. Future research should address the nature of and the reasons for, this inefficiency. In particular, it should address the question, why trading does not lead to elimination of these excess returns sooner.

REFERENCES