Investor Overconfidence and Trading Volume

Meir Statman, Steven Thorley and Keith Vorkink *

January 2004

* Meir Statman is the Glenn Klimek Professor of Finance at Santa Clara University's Leavey School of Business. Steven Thorley is the H. Taylor Peery Professor of Finance at Brigham Young University's Marriott School of Management. Keith Vorkink is an Assistant Professor of Finance at Brigham Young University's Marriott School of Management. For comments, we thank Grant McQueen, Terrance Odean, Michael Pinegar, James Brau, and participants at the University of Washington and Arizona State University seminars, and the PACAP/FMA and EFA conferences. Meir Statman acknowledges support from the Dean Witter Foundation. We thank Greg Adams for database support.

Send correspondence to Steven Thorley, Marriott School at BYU, 666 TNRB, Provo, UT 84602.
Phone: (801) 422-6065 Fax: (801) 422-0108 Email: steven_thorley@byu.edu

Investor Overconfidence and Trading Volume

Abstract

The proposition that investors are overconfident about their valuation and trading skills can explain high observed trading volume. In addition, biased self-attribution may cause the level of investor overconfidence to vary with past market returns. We test the market trading volume prediction of formal overconfidence models and find that turnover is positively related to lagged returns for months and perhaps years, even after controlling for turnover trend and contemporaneous volume-volatility relationships. Vector autoregressions and associated impulse response functions indicate that individual security turnover is positively related to lagged market returns, as well as lagged returns to the security. These results are consistent with disposition effect trading in conjunction with the trading volume prediction of investor overconfidence. We show that both overconfidence and disposition effect trading are more pronounced in small-cap stocks and in earlier periods where individual investors hold a greater proportion of shares. We also test the return autocorrelation predictions of formal overconfidence models and find some confirmatory results.

Investor Overconfidence and Trading Volume

Financial market economists have long puzzled over investor's enthusiasm for active trading in presumably efficient stock markets. While stock prices appear to be consistent with rational models, the investors themselves remain irrationally confident in their ability to exploit mispriced securities. Psychological theory and empirical evidence in non-market activities suggest that overconfidence is a pervasive behavioral norm in most competitive settings. Investor's overconfidence in security valuation and trading skills has recently become a formalized hypothesis among many financial economists. Theorists have focused on the impact of investor overconfidence in otherwise rational models of financial market pricing and trading. Daniel, Hirshleifer, and Subrahmanyam (1998) and Odean (1998) develop general equilibrium results that incorporate the assumption that some investors overestimate the precision of their private information. Researchers develop theory and testable implications under two general assumptions; first, that investors are overly confident about the precision of their private information, and second, that biased self-attribution causes the degree of overconfidence to vary with realized market outcomes. Gervias and Odean (2001) further develop the theory that overconfidence is enhanced in investors that experience high returns, even when those returns are simultaneously enjoyed by the entire market.

Odean (1998) and Gervais and Odean (2001) suggest that changes in trading volume is the primary testable implication of overconfidence theory. Alternatively, Daniel, Hirshleifer, and Subrahmanyam (1998) focus on security return implications and explain a number of previously documented pricing anomalies. Surprisingly little new empirical research has been directly motivated by the theoretical treatments of investor overconfidence. The paucity of overconfidence related empirical work may be due in part to the lack of well-defined and testable implications. In this study we focus on the new and fairly distinct implications associated with changes in trading volume over time for both the market and individual securities. We also explore some of the return implications of Daniel, Hirshleifer, and Subrahmanyam (1998) and others, although the large amount of past empirical research on return prediction makes it difficult to distinguish new implications.

The overconfidence model of Gervais and Odean (2001) and Odean (1998) predicts that high total market returns make some investors overconfident about the precision of their information. Although the price increases are market wide, investors mistakenly attribute gains in wealth to their ability to pick stocks. Overconfident investors trade more frequently in subsequent periods because of inappropriately tight error bounds around return forecasts. Alternatively, market losses reduce investor overconfidence and trading, although perhaps not in a symmetric fashion. Glaser and Weber (2003) distinguish between the "miscalibration" version of overconfidence and the idea that most investors simply believe their investment skills are better than average. We find little difference in the trading patterns implications between these two versions of investor overconfidence and the tests we conduct do not distinguish between them.

We differentiate investor overconfidence with the disposition effect of Shefrin and Statman (1985). The disposition effect describes a desire for investors to realize gains by selling stocks that have appreciated, but to delay the realization of losses. Odean (1998b) and other empirical researchers confirm that investors tend to sell the stocks in their portfolio that have paper gains, but are reluctant to acknowledge losses. We interpret the more recently developed overconfidence hypothesis as a separate theory of trading activity that relates to investors beliefs about trading in general, rather than an attitude about individual stocks they currently hold. Some recent empirical research on behavioral assertions, including overconfidence and the disposition effect, is based on proprietary trading records and portfolio holdings of specific individual investors. While such studies can help us understand the motives of some investors, they do not address the larger issue of whether these motivations are pervasive enough to impact the structure of the market in terms of realized trading volume and price discovery.

Our empirical research is time-series oriented and based on monthly observations of turnover for the NYSE/AMEX market and individual securities. We use vector autoregressions and impulse response functions to test specific implications of how trading activity relates to lagged returns. Our results can be summarized into four key findings. First, we find a statistically and economically significant positive relationship between market-wide turnover and lagged market returns, consistent with the prediction of the overconfidence hypothesis. The test was motivated by the overconfidence models developed in recent theory and to our knowledge is a newly documented empirical regularity regarding market volume. Second, we find that individual security turnover is positively related to both lagged security returns and lagged market returns. The positive security turnover response to own lagged return is consistent with the disposition effect, while we interpret the positive turnover response to lagged market returns as evidence of investor overconfidence. The relatively pronounced dependence of security turnover on lagged market returns in a regression that also includes lagged security turnover and returns is striking.

Our third empirical finding is that the lead-lag relationship between security returns and turnover is stronger in small-capitalization stocks and in earlier time periods. We hypothesize that this finding relates to the relatively larger roll of individual investor volume versus institutional and arbitrage based trading volume in small stocks and earlier time periods. Our fourth and final result relates to the predictability of security returns based on past trading volume. While the individual time-series methodologies we employ have little statistical power in explaining the realized cross-sectional variation in stock returns, the results are not inconsistent with the return sign predictions in Daniel, Hirshleifer, and Subrahmanyam (1998). We find a tendency for high volume stocks to have positive one month returns followed by several months of negative returns.

We are unaware of any other behavioral or rational expectations models that predict the positive lead-lag relationship between returns and volume we document. Portfolio rebalancing in the wake of large price movements can induce trading activity, but the implication would be for higher volume after large price increases and decreases (i.e., following large positive and negative returns). Further, one might expect rebalancing trades to quickly follow large price changes, not to be delayed by many months. In any event, we use concurrent and lagged observations of return volatility and return dispersion to control for alternative trading motivations.

The study proceeds as follows. In Section 1 we review overconfidence theory and past empirical work on trading volume. In Section 2 we introduce our data and empirical methodologies. Specifically, section 2 introduces the vector autoregression and impulse-response function methodology we employ, and describes a number of related econometric issues. Sections 3 and 4 report our empirical results on market and security time series, respectively. The market results in Section 3 are discussed in more detail in order to introduce the econometric framework. Our finding that market-wide turnover is responsive to past shocks in market return is consistent with the overconfidence hypothesis, the disposition effect, or some combination of both theories. This ambiguity motivates the analysis of individual security volume and returns in Section 4. We estimate vector autoregression models on many (hundreds) individual security multivariate time series and provide average results on impulse-response functions. Section 4 also discusses sub-period and firm-size based findings, and the predictability of security returns. We state our conclusions in Section 5.

1. Overconfidence Theory and Volume Research

Black (1986) first argued that noise traders offer an exit from the no-trading equilibrium of perfectly rational models of securities markets. Recently, Odean (1998), and then Gervais and Odean (2001) developed a multi-period model where the overconfidence of noise traders increases as they attribute high returns in bull markets to their trading skills. These models do not specify an exact time frame for the lead-lag relationship between returns and trading activity, only that high (low) market returns lead to high (low) subsequent volume. Models of investor overconfidence and biased self-attribution are also developed by and Daniel, Hirshleifer, and Subrahmanyam (1998). Daniel, Hirshleifer, and Subrahmanyam (1998) make an important distinction between public and private informational events, and develop predictions about stock price over and under reaction. Under certain circumstances, security returns are positively auto correlated in the short run but negatively autocorrelated over longer horizons.

The disposition effect was first proposed by Shefrin and Statman (1985) who combined prospect theory from Kahneman and Tversky (1979) with the emotions of pride and regret. Investors in the Shefrin-Statman model think about stocks within mental accounts, one for each stock. Pride accompanies the realization of paper gains, and regret accompanies the realization of paper losses. Ironically, the disposition effect motivates trading that may be in direct opposition to tax-efficient portfolio management. Empirical evidence for the disposition effect is provided by Lakonishok and Smidt (1986), Ferris, Haugen and Makhija (1988), Odean (1998b), and Heath, Huddart and Lang (1999).

The empirical implications of the more recently developed theory of investor overconfidence differ from the disposition effect in two important ways. First, the disposition effect explains the motivation for only one side of a trade. The desire to sell a specific stock with a paper gain is expressed by trading with other investors without this bias, and thus may affect the pricing equilibrium for that stock. If disposition related selling (for gains) or a resistance to sell (for loses) comprises a material part of total volume, prices will be slow to react to new information. For example, Grinblatt and Han (2002) find that disposition motivated trading is the root cause of the Jegadeesh and Titman (1993) momentum anomaly; positive autocorrelations in returns lasting several months. Goetzmann and Massa (2003) refine the notion that the one-sided nature of disposition motivated trades can affect prices. In contrast, overconfidence in stock picking expertise among investors can explain both sides of a given transaction. Differences of opinion together with inappropriately tight error bounds by two investors will result in a trade that need not include other liquidity traders or rational information traders.

Second, the disposition effect is generally understood to refer to an investor's attitude towards specific stocks in his or her portfolio. Consequently, most recent empirical research on the disposition effect (e.g., Odean (1998b), Ranguelova (2001), and Dhar and Zhu (2002)) employs data on individual investor transactions in specific stocks at various prices. In contrast, the theory of intertemporal changes in investor overconfidence suggests an attitude about the stock market in general. If investors overestimate their ability to increase wealth by active trading, they are likely to maintain this belief about stocks in general rather than the specific securities they currently hold. Time-series observations on a specific investor's portfolio return and trading activity are valuable in that the investor's own portfolio return would likely influence his or her level of overconfidence more than the general market return. However, to be useful, the individual investor portfolio return data would have to span longer periods than investor-specific databases currently allow. As a result, the best chance of finding evidence for the overconfidence hypothesis may be in the long-term trading volume data available from the Center for Research in Security Price (CRSP). In addition, variations in realized trading volume as opposed to the trades of a specific set of investors may be the more important variable for financial market economists to explain.

Considerable empirical research relates contemporaneous volume and returns including Karpoff (1987), Stoll and Whaley (1987), Bessembinder and Seguin (1993), Bessembinder, Chan and Seguin (1996), Chordia, Roll, and Subrahmanyam (2000), and Lo and Wang (2000). However, little prior empirical research relates *current* volume to *lagged* returns. Gallant, Rossi and Tauchen (1992) document several regularities, but the non-parametric methodology of their research does not yield interpretations relevant to the overconfidence hypothesis. Chordia and Swaminathan (2000) examine volume and return cross-autocorrelations at very short (i.e., daily) horizons to explore the speed at which information is priced. Recent research by Cooper (1998), Lee and Swaminathan (2000), Llorente, Michaely, Saar, and Wang (2002) and Gervais, Kaniel and Mingelgrin (2002) examine the ability of volume to predict returns, not the ability of returns to predict volume. Another large branch of theoretical and empirical research relates volume to concurrent return volatility, as well and lead-lag relationships between volatility and volume, for example Harris and Raviv (1993) and Shalen (1993). We include contemporaneous and lagged observations of return volatility and cross-sectional security return dispersion in our vector autoregression models.

2. Data and Methodology

2.1 Data description

Our database consists of monthly observations on all NYSE/AMEX common stocks, excluding closed-end funds, REITs and ADRs, from August 1962 to December 2002, the span of the daily CRSP files.[1] We exclude NASDAQ stocks because the dealer market has volume

[1] We use monthly observations for turnover and returns, but our estimate of volatility is constrained by the availability of daily returns which start in July 1962. Our start date is delayed to August 1962 to obtain uniform availability of the prior month-end shares outstanding data.

measurement conventions that differ from the exchange traded securities (see Atkins and Dyl (1997), among others). In separate unreported regressions, we verify that NASDAQ stocks exhibit results that are similar to the smallest three size quintiles of the NYSE/AMEX database.

An examination of long-term U.S. trading activity indicates that the number of outstanding shares for the typical stock has increased markedly over the last four decades. The median number of outstanding shares per security has increased almost twenty-fold from 1.8 million to 34 million, making raw share volume a poor measure of trading activity.[2] Following Lo and Wang (2000) we measure trading activity with turnover (shares traded divided by outstanding shares) and aggregate security turnover into market turnover on a value-weighted rather than equal-weighted basis. Value-weighted market turnover is algebraically equivalent to total dollar turnover; the dollar value of all shares traded divided by the total dollar value of the market.[3]

Figure 1 presents NYSE/AMEX turnover from August 1962 to December 2002. Turnover generally stayed below 2 percent per month until the mid-1970's, equivalent to a complete (i.e., 100 percent) turnover time of more than 4 years. Turnover gradually increased to a level of about 6 percent per month by the late 1980's, reducing the complete turnover time to 18 months. Trading dropped off for a few years after the spike associated with the October 1987 market crash, but then picked up through the 1990's and the turn of the century, ending 2002 at almost 10 percent per month; a complete turnover time of 10 months. One possible explanation for the long-term trends in trading activity are changes in investor overconfidence induced by realized portfolio returns, the focus of this study. Alternative explanations include the deregulation of brokerage fees (May Day 1975) and other changes in transaction costs, trading

[2] Only a portion of the dramatic increase in the average number of shares per stock is attributable to increases in real company size. Rather, institutional considerations such as price discreteness and listing requirements, expressed though stock splits, have kept the median share price between \$15 and \$20 over the last four decades, despite a six-fold increase in the Consumer Price Index (Angel (1997)).

[3] Let V_i be trading volume and S_i be outstanding shares for security i, so that turnover is $T_i = V_i/S_i$. Market-wide turnover is defined as the value-weighted average of individual security turnover, $T_M =$? $w_i T_i$ where the weights, w_i, are based on the capitalization of the security divided by the sum of capitalization for all securities in the market. Capitalization is the beginning of period price per share, P_i, times shares outstanding, S_i, for each security. Some algebra indicates that T_M is equal to the sum of the dollar volume for each security, P_i*S_i, divided by the sum of the capitalizations.

associated with financial derivatives arbitrage, and the growing influence of institutional rather than individual traders (see Smidt (1990)).

The secular trends in Figure 1 indicate that the turnover time-series is non-stationary, leading to bias in the coefficient standard errors of vector autoregressions (VAR) we employ in this study. Turnover is definitionally restricted to non-negative values and the natural log transformation helps, eliminating the visual correlation in Figure 1 between the level of the trend and volatility around the trend. However, logged turnover still exhibits non-linear secular trends over time, and we fail to reject to null-hypothesis of a unit root using the Phillips and Perron (1988) test.[4] We employ the Hodrick-Prescott (1997) (henceforth HP) algorithm for detrending turnover, although the qualitative results we present are invariant to alternative detrending methodologies.[5] Simpler (e.g., linear time-trend) detrending methodologies appear inadequate for the market turnover in Figure 1, and are not flexible enough for the many turnover patterns in individual security turnover time-series we also examine. For reference purpose, Figure 1 includes the exponentiated HP trend (dotted line) computed from log turnover. The detrended time-series used in this study is the monthly difference between log turnover and its trend.

Removing the trend from the market turnover time-series presents an important bias against finding long-term trading activity responses to market returns. Gervais (1998) and Gervais and Odean (2001) hypothesize that realized returns impact trading activity for some unspecified period of time, leading to possible long-term trends in turnover. For example, the long-term drop in activity after the 1987 crash could be indicative of a drop in investors' confidence in the value of active trading. Similarly, the long-term increases in trading starting in

[4] The Phillips and Perron (1988) test-statistic is -1.8 for log market turnover and -16.0 for detrended log market turnover. The five-percent critical value for the test statistics is -2.9.

[5] The Hodrick-Prescott (1997) algorithm creates trend series, s, by minimizing the variance of the raw series, y, around the trend, subject to a penalty on the second difference of the trend. Specifically, the HP filter chooses s_t to minimize: $\sum_{t=1}^{T}(y_t - s_t)^2 + h\sum_{t=2}^{T-1}[(s_{t+1}-s_t)-(s_t-s_{t-1})]^2$ where h is the penalty parameter (the trend becomes more smooth as h is increased). We follow the common practice of setting h=14,400 for monthly observations. The HP filter is two-sided, meaning that data before and after time t is used in the smoothing process. Our motivation for detrending is to extract a stationary time-series, not to predict the trend.

the late 1970's and early 1990's are coincident with long-term bull markets. While consistent with the overconfidence hypothesis, as well as the disposition effect, these long-term observations are anecdotal; too infrequent for statistical analysis and hypothesis testing. We therefore present detrended log turnover as our base-case in this study and comment on how the results differ for non-detrended log turnover.

Table 1 presents summary statistics on monthly market turnover and market return as well as two market-wide based control variables; market volatility and dispersion. Table 1 has descriptive data on the full sample (August 1962 to December 2002) as well as four non-overlapping 120 month (decade long) sub-periods that are used later in robustness checks. The sub-period means and standard deviations for raw turnover again suggest a non-stationary series. For example, the average turnover in the last sub-period, 7.07 percent, is several standard deviations above the average value of 1.48 percent in the first sub-period. The sub-period descriptive statistics for detrended monthly log turnover, *mturn*, are indicative of a stationary time-series.

Market return, *mret,* in Table 1 is the monthly return with dividends on a value-weighted portfolio of all NYSE/AMEX non-fund common stocks. For example, the mean monthly return over the full sample is 0.94 percent; a simple annualized average return of 12*0.94 = 11.28 percent. Unreported statistics on the higher moments of the returns data in this study are consistent with established distributional characteristics of monthly market returns; significant leptokurtosis and slight negative skewness. The sub-sample means and standard deviations for *mret* are relatively stable, indicative of stationary time-series.

Our first control variable, market volatility, *misg*, is the monthly temporal volatility of market returns for the value-weighted composite of all NYSE/AMEX non-fund common shares, measured in percentage points. The volatility control is based on Karpoff's (1987) survey of research on the contemporaneous volume-volatility relationship, as is similar to the mean absolute deviation (MAD) measure in the trading volume study of Bessembinder, Chan, and Seguin (1996). The monthly realized volatility estimates are based on daily market returns within the month, correcting for realized autocorrelation, as specified in French, Schwert, and Stambaugh (1987).[6]

The second control variable, dispersion, *disp*, is the monthly cross-sectional standard deviation of returns for the previously defined list of NYSE/AMEX non-fund common stocks, measured in percentage points. Consistent with calculation of market return as the value-weighted mean security return, the cross-sectional standard deviation of returns is value-weighted (i.e., market-cap weights are used in the averaging of squared deviations from the mean). Return dispersion is related to the level of realized idiosyncratic risk in stocks and has been relatively stable around the 7.10 percent full-sample mean shown in Table 1, although somewhat higher in recent years as reported by Campbell, Lettau, Malkiel, and Xu (2001). Return dispersion is included as a control variable to account for potential trading activity associated with portfolio rebalancing. For example, large spreads between the individual stock returns might lead to trading activity among investors seeking to maintain fixed (e.g., equal) portfolio weights.

In addition to the market data presented in Table 1, we perform time-series analysis on the returns and trading volume of individual stocks. The full-sample span of 485 months on 2,000 to 2,500 NYSE/AMEX securities per month results in over one million individual stock observations. These observations are organized as multivariate time-series for each stock with heterogeneous start and stop dates associated with the listing of that stock. The VAR models we estimate require a continuous time-series with a sample size that allows for the estimation of relatively large number of parameters. We establish a requirement of at least ten years (120 months) of contiguous volume, return, and standard deviation data in the CRSP monthly and daily files.[7] The data sufficiency requirement yields 1,878 securities with start and stop months

---

[6] Specifically, we calculate month t's volatility as $msig_t^2 = \sum_{t=1}^{T} r_t^2 + 2\sum_{t=1}^{T} r_t r_{t-1}$, where r_t is day t's return, and ? is the number of trading days in month t.

[7] The frequency of missing volume observations in the CRSP database is high enough that excluding an entire time-series for one missing month may be an overly strict criterion that introduces unintended biases associated with the exclusion of certain types of stocks (i.e., small-cap). We replace occasional (not more than one percent of a given time-series) isolated (single occurrence) missing turnover

at least ten years apart, but within the August 1962 to December 2002 full-sample period. The time-series that meet the data sufficiency requirement contain a little over one-half a million monthly observations, as shown in Table 2.

Table 2 presents summary statistics for three security variables used in the individual stock VAR estimations: turnover, return, and volatility. Turnover is the volume of shares traded in month *t* divided by the prior month-end shares, adjusting for stock splits that occur during the month. The security turnover averages in Table 2 are slightly higher than the value-weighted market composite in Table 1 because smaller capitalization stocks tend to have higher turnover rates. The standard deviation of turnover rates across securities and time in Table 2 is large, with rare minimum values of zero turnover for some securities in some months, as well as extraordinary maximum turnover rates of several hundred percent. An examination of the turnover time-series charts for individual securities, similar to Figure 1, indicate long-term trends and non-stationarities similar to, but often more complex than, the market composite. As with the market, we log and detrend individual security turnover using the HP filter, producing the detrended monthly log turnover variable, *turn*.

Individual security return, *ret*, is the monthly return with dividends measured in percentage points. As with turnover, the variation in returns across stocks and months shown in Table 2 is substantial, with extreme minimum and maximum values common to a large return database. Individual security volatility, *sig*, is the return volatility for a given firm and month, measured in percentage points. As with the market, we use daily security return data and the French, Schwert, and Stambaugh (1987) method to calculate realized volatility for each stock each month. A comparison of the mean security volatility estimates, *sig*, the standard deviation of security returns, *ret*, and the mean cross-sectional dispersion on security returns from Table 1, *disp*, provides a confirmatory reasonableness check on the observation calculations.

observations with the prior month turnover before a security time series is measured against the ten-year contiguous data criterion.

2.2 Empirical Methodology

Vector autoregressions (VAR) and associated impulse response functions are used to study the interaction of turnover and return time-series for both the market and individual securities. The general form of the VAR model is

$$\mathbf{Y}_t = \mathbf{a} + \sum_{k=1}^{K} \mathbf{A}_k \mathbf{Y}_{t-k} + \sum_{l=0}^{L} \mathbf{B}_l \mathbf{X}_{t-l} + \mathbf{e}_t \qquad (1)$$

where $\mathbf{Y}_t$ is a nx1 vector of period t observations of endogenous variables, for example turnover and return, $\mathbf{X}_t$ is a vector of period t observations of the exogenous (i.e., control) variables, and $\mathbf{e}_t$ is a nx1 residual vector. The regression coefficients, $\mathbf{A}_k$, and $\mathbf{B}_l$, estimate the time-series relationships between the endogenous and exogenous variables, where K is the number of lagged endogenous observations, and L is the number of lagged exogenous observations. The VAR methodology allows for a covariance structure to exist in the residual vector, $\mathbf{e}_t$, that captures the contemporaneous correlation between endogenous variables.

Formal overconfidence theories do not specify a time frame for the relationship between returns and turnover, so we let the data determine the number of monthly lags to include. Specifically, we set K=10 based on the Schwartz Information Criteria (SIC).[8] The SIC model selection criterion is based on the log likelihood function adjusted by a penalty for the number of parameters, and performs well for model selection with large sample sizes (see Lutkepohl (1991)). The model selection criterion also leads to L=2, so that contemporaneous and two lagged monthly observations of the exogenous variables are used to explain and predict the endogenous variables. The optimal lag lengths of K=10 and L=2 are established based on the market-wide time series, and then applied uniformly to the individual VAR estimations for each security. While the SIC suggests some variation in the optimal lag lengths for individual security models, we fix the lag lengths at K=10 and L=2 to facilitate the cross-sectional comparison of the coefficient estimates.

[8] Specifically, the Schwartz Information Criteria takes the form $SIC = \frac{-2\Omega}{T} - p\left(\frac{\log(T)}{T}\right)$, where p is the number of estimated parameters, T is the number of observations, and Ω is the value of the log likelihood function using the p estimated parameters.

Based on the VAR model, we employ impulse response functions to aggregate over coefficient estimates and illustrate how the endogenous variables relate to each other over time (see Hamilton (1994)). Impulse response functions trace the effect of a one standard deviation shock (measured within sample) in one residual to current and future values of the endogenous variables through the dynamic structure of the VAR. For example, the market version of the general model in Equation (1) contains two endogenous variables, market turnover and market return, and two exogenous variables, market volatility and dispersion:

$$\begin{bmatrix} mturn_t \\ mret_t \end{bmatrix} = \begin{bmatrix} a_{mturn} \\ a_{mret} \end{bmatrix} + \sum_{k=1}^{10} \mathbf{A}_k \begin{bmatrix} mturn_{t-k} \\ mret_{t-k} \end{bmatrix} + \sum_{l=0}^{2} \mathbf{B}_l \begin{bmatrix} msig_{t-l} \\ disp_{t-l} \end{bmatrix} + \begin{bmatrix} e_{mturn,t} \\ e_{mret,t} \end{bmatrix} \tag{2}$$

Changes in one residual, say $e_{mturn,t}$, will immediately change the value of current value of *mturn*, but will also affect future values of *mturn* and *mret* since lagged values of *mturn* appear in both equations through the coefficient matrix $\mathbf{A}_k$.[9] For example, to test the market prediction of the overconfidence hypothesis, we shock the market return residual, $e_{mret,t}$ by one sample standard deviation. Using the estimated parameters and the dynamic structure of the VAR, we track how market turnover responds over time to the $e_{mret,t}$ shock. Thus, while the VAR model is dense in parameters, the impulse response function provides a simple picture of how the endogenous variables are related over time.

The VAR model for individual securities is trivariate, with endogenous variables security turnover, security return, and market return, and has a single exogenous variable, security volatility:

$$\begin{bmatrix} turn_t \\ ret_t \\ mret_t \end{bmatrix} = \begin{bmatrix} a_{turn} \\ a_{ret} \\ a_{mret} \end{bmatrix} + \sum_{k=1}^{10} \mathbf{A}_k \begin{bmatrix} turn_{t-k} \\ ret_{t-k} \\ mret_{t-k} \end{bmatrix} + \sum_{l=0}^{2} \mathbf{B}_l sig_{t-l} + \begin{bmatrix} e_{turn,t} \\ e_{ret,t} \\ e_{mret,t} \end{bmatrix}. \tag{3}$$

As with the market-wide model, we are primarily concerned with how security turnover responds to shocks in return, both security return and market-wide return. We also examine the response of security returns to shocks in security turnover to test some of the overconfidence implications in Daniel, Hirshleifer, and Subrahmanyam (1998). Market returns are not predictable based on a

[9] The residuals often have a common component which cannot be associated with a specific variable. For the purpose of impulse-response functions, the residuals are sometimes orthogonalized (for example, by the Cholesky decomposition) resulting in a diagonal covariance matrix. Unreported orthogonalized impulse-response functions are qualitatively similar to those we report.

single security in any plausible theory, so the third endogenous variable, *mret*, could have also been introduced as an exogenous variable in Equation (3). However, the Shefrin and Statman (1985) disposition effect motivates a direct comparison of past individual and market return impacts on security turnover, so we introduce the market return, *mret*, in the same format as the individual security return, *ret*.

We estimate Equation (3) on the 1,878 individual NYSE/AMEX securities that meet the minimum data sufficiency requirement. Each of these VAR models has a full set of coefficient estimates and standard errors, producing extensive output that can not be reported on a security by security basis. For brevity, we report the average coefficient across all securities for each the individual elements of the endogenous coefficient matrix, $\mathbf{A}_k$. While the VAR coefficient standard error estimates are valid for each security, there are material dependencies between the coefficient estimates across securities, making the statistical significance of the mean coefficient we report difficult to ascertain. For example, simply estimating the standard error of a coefficient mean by the cross-sectional standard deviation divided by the square root of the sample size grossly underestimates the standard error, leading to spurious power in associated t-tests.

To address this problem, we introduce a boot-strap procedure that provides standard errors for the cross-sectional mean of each VAR coefficient, as well as the security impulse response function coefficients. The time-series nature of the VAR model and the size of the security database lead to an econometrically complicated and computationally intensive bootstrap procedure, which we outline briefly below. Our choices generally follow the VAR-bootstrap specifications in Runkle (1987). (Hamilton (1994) also includes a review of VAR bootstrapping.) The bootstrap program creates 5,000 artificial databases with time-series structure and dependencies similar to the single live database. The observed variation of the coefficient means within each of these databases provides an accurate estimate of the standard error of the live database means.

A key aspect of the VAR bootstrap program is that each artificial database has one random sequencing of months that is applied consistently across the 1,878 securities in order preserve the cross-sectional dependencies in the original data. We first estimate Equation (3) for each security and store the estimated coefficients and fitted residuals. Then, for each security we create a Tx3 matrix of random variables, u, by sampling with replacement from the appropriate set of fitted residuals. We then use the estimated coefficients and the structure of Equation (3) to construct an artificial time series for that security. We repeat this process for all 1,878 securities, creating one database of simulated values of security turnover, security return, and market return.[10] By repeating this process 5,000 times, we obtain sufficiently precise estimates of the standard error for each mean coefficient in the endogenous coefficient matrix, $\mathbf{A}_k$ (i.e., the standard error for each of the sixty coefficient means reported in Table 4 is different). Our bootstrap procedure also provides standard errors for the mean coefficients of the security impulse response functions. The results indicate that a naive employment of the sample standard deviation of coefficient values in the single live database understates the actual standard error by as much as a degree of magnitude, confirming the need for the bootstrapping program.

3. Market VAR Estimation and Test Results

3.1 Market Vector Autoregression

Table 3 presents the results of the full-sample bivariate VAR of detrended log market turnover, *mturn*, and market return, *mret*. The table is organized by rows for each dependent variable (*mturn* and *mret*) and by columns for lagged dependent variable and exogenous variables coefficients. For each coefficient we report the estimated value, standard error, and the p-value for a two-sided t-test that the true coefficient value is zero. Significance levels (in parentheses) are rounded to two digits so that a reported p-value of (0.00) indicates significance at the 0.005 level or less. In our review of test results, we generally refer to coefficients with p-values of (0.05) or less as significant, and coefficients with p-values of (0.01) or less as highly significant.

[10] Our bootstrap method results in 1,878 unique time-series of the market return for each database, while the live database naturally has only one. As an alternative, we constructed one time-series of the market return for each database by using the average security coefficients and average residuals. Preliminary results of this alternative are not materially different from the adopted bootstrap method.

Table 3 indicates that market turnover is autocorrelated, with highly significant first lagged coefficient of 0.284 (standard error of 0.047) and coefficients of generally declining magnitude on the second and higher lags. Notably, market turnover is also dependent on lagged market returns even in a regression that includes lagged observations of turnover and other control variables. The first lagged market return coefficient in the *mturn* dependent variable regression is 0.819 (standard error of 0.133) and highly significant, as is the second lagged market return. Indeed, the magnitude of the first lagged market return coefficient estimate is about three times greater than lagged turnover (compare 0.819 to 0.284), although with a correspondingly larger standard error. The positive and highly significant association between market turnover and lagged market returns represents the first key empirical finding of this study. We will discuss this key result in more detail in the context of impulse response functions.

Consistent with previous trading volume studies (i.e. Karpoff (1987)), the association of turnover with contemporaneous volatility is large and positive, with a coefficient estimate of 1.712. The first and second lagged *misg* coefficients indicate that the volatility-volume relationship reverses sign in subsequent months. However, individual coefficients on lagged *misg* must be interpreted with caution because of well-known autocorrelation in volatility. Market returns exhibit significant second moment (i.e., GARCH) temporal dependencies, causing material multicollinearity among the lagged *misg* observations in the VAR model. We find that the lagged *misg* coefficient estimates vary greatly with the number of lags included in model, as well as the inclusion of the other exogenous control variable, *disp*.

Dispersion also has a highly positive contemporaneous association with market turnover as shown by the concurrent *disp* coefficient of 5.024. Like volatility, the sign of the first and second lagged *disp* coefficients is negative. However, univariate analysis indicates that dispersion exhibits even more autocorrelation than volatility, and dispersion and volatility are contemporaneously correlated. We include the exogenous variables *misg* and *disp* to control for alternative explanations of trading activity, although their inclusion and specified lag structure does not materially change our key findings.[11] Presumably, the strong positive contemporaneous association between trading activity and both inter-temporal market price changes (*misg*) and cross-sectional variation in security price changes (*disp*) is due to the impact of information events in the market.

We also included calendar month (i.e., January through December) dummies in the market and security VARs and found that trading activity is the highest in December and January, possibly due to tax related trading, and drops off slightly in the summer months. The high turnover rates around the turn of the year for the typical stock are less evident in value-weighted market turnover, suggesting that the seasonality in trading activity is more pronounced in small market capitalization stocks. The mean individual security return is higher in January than any other month, consistent with the small-firm effect first identified by Banz (1981) and Reinganum (1983). The inclusion of calendar month dummies confirms well-established empirical regularities in the U.S. equity market, but does not impact any of our key findings. We do not include dummies in our reported model to reduce the already high number of parameter estimates.

Generally consistent with weak-form market efficiency, the *mret* dependent variable regression in Table 3 shows few statistically significant coefficients. For example, the first lagged *mturn* coefficient is 0.019 (standard error of 0.018). Cooper (1999), Lee and Swaminathan (2000), and Gervais, Kaniel, and Mingelgrin (2001), among others, find that past volume can be used predict returns, however these studies are for individual securities. The only market return prediction anomaly in Table 3 is a positive and significant autocorrelation at the 5th lagged *mret* (coefficient estimate of 0.125), offset by an almost significant negative autocorrelation coefficient at the 6th lagged *mret*. While statistically significant, this result is likely due to a few coincidental large market return observations that are separated by five months. For example, the October 1987 market crash happened to be followed by another negative return in March 1988.

---

[11] We examined alternative specifications for the exogenous variables *disp* and *msig*, including exclusion of one or both, different lag structures (variations in the value of *L* in Equation (2)) and log transformations. None of these perturbations to the base-case specification exhibited materially different results in terms of the impulse-response functions of the endogenous variables.

While not predictive in nature, Table 3 shows that market return is negatively associated with contemporaneous market volatility, *misg*, and positively associated with security return dispersion, *disp*. Some of the lagged observations of the exogenous variables are significant, but as with the turnover dependent variable regression, the lagged exogenous coefficient estimates are subject to material multicollinearity and sensitive to the exact model specification.

3.2 Market Impulse Response Functions

As explained in the methodology section, individual VAR coefficient estimates do not capture the full impact of an exogenous variable observation. Impulse response functions use all the VAR coefficient estimates to trace the full impact of a residual shock that is one sample standard deviation from zero. Figure 2 contains all four possible impulse-response function graphs using the bivariate VAR estimation shown in Table 3. Panels (a) and (b) plot the response of *mturn* to a one standard deviation shock in *mturn* and *mret*, respectively, along with confidence bands spaced out at two standard errors. Because of the log transformation, the vertical axis in panels (a) and (b) measures the percentage increase in turnover relative to the non-shock turnover level. For example, in Panel (a) the first period impulse-response indicates that a one standard deviation shock to market turnover results in approximately a 3.0 percent increase in the next month's turnover. Panel (a) verifies the serial dependence of turnover in that the positive impact of a turnover shock persists for about five months. Note the impulse response function is forced to zero over time because turnover is detrended. Impulse response functions on raw log turnover (unreported) show declining but generally positive values for all 24 months.

Figure 2 panel (b) indicates a large and persistent positive response in turnover to a market return shock; the first key finding of this study. We interpret this result as evidence that market returns impact investor confidence and subsequent trading activity. The figure shows that a one standard deviation market return shock results in an 8.6 percent increase in the next month's turnover, and a 7.3 percent increase in the following month's turnover. The accumulated response over the first six months is a 30 percent increase in turnover compared to average levels, an economically as well as statistically significant event. As a measure of economic significance, consider the descriptive statistics in Table 1 which give the average monthly market return as about 1.0 percent, with a standard deviation of about 4.0 percent. Compare a relatively high market return in one month of say 7.0 percent (1.5 standard deviations above the mean) with a relatively low market return of -5.0 percent (1.5 standard deviations below the mean). The impulse response function suggests that a high return of 7.0 compared to minus 5.0 percent adds the equivalent of an extra month of trading volume (30 percent times 3 standard deviations) over the next six months.

Interestingly, the shape of the true impulse response function may not be monotonically declining. In unreported sensitively analysis, we replicate the market portion of this study using weekly rather than monthly observations, following the trading volume research of Lo and Wang (2000). VAR estimations using weekly market turnover and return observations over the full sample period suggest that the impulse response to a market return shock increases for a couple of weeks before beginning a gradual decline that lasts for about twenty weeks. While the weekly finding for market-wide data is interesting, we focus on monthly observations under the perspective that changes in investor overconfidence should be evident over long time periods.

As in panel (a), the impulse response function in panel (b) using raw (non-detrended) data declines but remains persistently positive over all 24 months. As suggested in the methodology section, an interpretation of this fact favorable to the overconfidence hypothesis is that market returns impact trading activity over a long enough time frame that the impact is identified as a trend and removed by the HP algorithm. Strictly predictive but less sophisticated detrending methodologies, for example a moving average of the prior 60 months of turnover, yield results that are similar to HP detrended data we report. The process used to detrend turnover does not resolve the question of whether the long-term trends are due to a concentration of positive monthly returns, or a result of unrelated secular events that might impact volume.

For completeness, Figure 2 panels (c) and (d) plot the response of *mret* to a one standard deviation shock to *mturn* and *mret*, respectively, along with a two standard error confidence bands. The vertical axis in panels (c) and (d) measures the difference in market return compared to the average return. For example, in Panel (c) the first period impulse-response indicates that a one standard deviation shock to market turnover results in a 20 basis point increase in the next month's return. The two standard error bands suggest this result is weakly significant, but in general the impulse response function coefficients in panel (c) is not significantly different from zero. The market return to market return impulse response function in panel (d) is also unremarkable and consistent with weak-from market efficiency, with the exception of lagged months 5 and 6, based on the anomalous coefficients in the market VAR estimation in Table 3.

4. Security VAR Estimation and Test Results

4.1 Security Vector Autoregressions

As reported in the previous section, we find strong empirical evidence for the overconfidence hypothesis in that market-wide trading activity is positively related to past market returns. However, this finding is also consistent with the disposition effect. Investors may be trading more after high market returns because of an increased confidence in their trading skills, or simply because they enjoy realizing paper gains on individual securities. Alternatively, investors may be trading less after negative market returns because the loss reduces their confidence in the value of trading, or simply because they do not want to sell individual securities and acknowledge the loss.

One distinguishing characteristic of the two alternative explanations is that the Gervais and Odean (2001) overconfidence hypothesis relates to trading in general, whereas the Shefrin and Statman (1985) disposition effect has generally been understood to describe investor attitudes towards specific securities in their portfolio. In this section we attempt to disentangle the two theories by examining the trading activity of individual stocks. If overconfidence plays a role in explaining volume in addition to the disposition effect, we should find positive coefficients in regressions of security turnover on lagged market returns even when lagged returns on that stock are also included. A definitive disentanglement of the overconfidence and disposition hypothesis may prove difficult. For example, one could argue that when investors sell other securities for the disposition effect, they raise cash that can then be used to purchase the security in question, thus increasing its volume. Despite this limitation in interpretation, an examination of individual security volume will at least ensure that the market-wide turnover patterns we have documented are not a simple summation of direct disposition effect evidence in individual security VARs.

As described in Table 2, we use monthly turnover and returns on 1,878 NYSE/AMEX non-fund common stocks that have at least 120 months of contiguous time-series data. We estimate the trivariate VAR in Equation (3) for each security, and report the mean coefficient estimates in Table 4. The security VAR model has three endogenous variables, *turn*, *ret*, and *mret*, and one exogenous variable, *sig*. Detrended log security turnover, *turn*, and security return, *ret*, parallel the endogenous market variables in the bivariate market VAR, and the endogenous variable, security volatility, *sig*, parallels volatility. While the SIC for optimal model selection varies across firms, we choose to abide by the K=10 and L=2 lag specification used in the market VAR for all individual firm VARs in order to ensure consistency in the summary results reported in Table 4. We also find that the specification of K=10, and L=2 yields the minimum average SIC across all securities.

The voluminous output associated with 1,878 separate security VAR estimations requires a compact way to communicate the results. We show cross-sectional mean coefficient in Table 4, along with the p-value of a two sided t-test that the true mean coefficient value is zero. As explained in the methodology section, the significance tests are based on bootstrapped standard errors for each mean coefficient. While the security VAR is trivariate in endogenous variables, we do not report the market return dependent variable regression or the exogenous variable coefficients. As one would expect, the unreported coefficients indicate that nothing on a security specific basis predicts market returns. As with market turnover, the unreported exogenous variable coefficients indicate that security turnover is positively and significantly correlated with contemporaneous security volatility, consistent the volume-volatility relationship documented in prior volume research.

The critical results in Table 4 relate the dependence of security turnover, *turn*, on lagged security and market returns. None of the lagged security return, *ret*, mean coefficients are individually significant, but are notable in that all ten lags have a positive sign. For example, first lagged *ret* mean coefficient is 0.171 and has a p-value of (0.33). On the other hand, the first lagged *mret* mean coefficient is 0.990 and highly significant, as are the second and third lagged mean coefficients. Thus, not only does that impact of past market returns on a typical security's trading activity survive the inclusion of lagged security returns in the same regression, it appears that the lagged market return impact is actually larger (compare 0.990 to 0.171). Note that lagged observations of security turnover are also included in each security VAR, so the estimated impact of market returns can not be dismissed based on the notion that they are a proxy for turnover autocorrelation. We focus further on our second key finding that security turnover depends on both lagged security and market returns in the context of impulse response functions that follow.

4.2 Security Impulse Response Functions

Figure 3 panel (a) shows the security turnover response to a one standard deviation shock in security return. Despite the use of detrended security turnover, the response is consistently positive over many months, consistent with the prediction of the disposition effect. The mean value of the first lagged security return coefficient of 2.29 percent is statistically significant at the (0.02) level (shown later in the top row of Table 5). The relatively large standard error bands in Figure 3 panel (a) suggest that while the mean coefficient is positive, there is considerable variation in the impluse response functions of the 1,878 individual security VAR estimations.

Figure 3 panel (b) shows the average security turnover response to a one standard deviation shock in market return. Although the duration of the market return shock appears shorter (in months) compared to the security return shock, the market return shock has a relatively larger impact on security turnover. The impact is also more uniform across securities as indicated by the relatively narrow standard error bands. The first lagged market return mean coefficient of 4.25 percent is statistically significant at less than the (0.01) level (shown later in the top row of Table 6). The impact of the market return shock remains statically significant for the first six months. In aggregate, the security VAR estimations and associated impulse response functions indicate that both past security and past market wide returns are important predictors of trading activity for the average security. We interpret these results as confirmation of the disposition effect in securities, and a strong validation of the new testable implication that changes in market return induced investor overconfidence impacts realized trading volume.

4.3 Sub-period and Firm Size Sensitively Analysis

We performed a sub-period analysis by estimating the security VAR in Equation (3) using data samples from four 120 month (decade long) sub-periods that span the full 1963 to 2002 sample period. The last five months of 1962 were not included to obtain equal sub-period sample sizes. Our sensitivity analysis focuses on the security turnover impulse responses to both security and market returns, as shown in Tables 5 and 6, respectively.

Table 5 presents the sub-sample results for the first six lags of the security turnover impulse response function to a security return shock. The full-sample results in the top row of Table 5 correspond to Figure 3 panel (a), which we associate with the disposition effect. Most of the sub-period mean coefficient estimates are positive, with the strongest evidence in the second sub-period, 1973 to 1982, where the first two lags are statically significant. The generally positive coefficient values are not evident in the latest sub-period, 1993 to 2002, where there is no evidence of disposition effect trading.

Table 5 breaks the sub-period analysis further into three size-based samples. Securities naturally migrate between size rankings over time based on realized returns, one of the endogenous variables. To avoid unintended biases, we measure a securities market capitalization only once at the beginning of each sub-period in order to assign the security time-series to a size quintile. We label the sample of largest quintile securities as large-cap stocks, and the second quintile as mid-cap stocks. The requirement of 120 months of contiguous time series data (i.e., all possible sub-period observations) excludes a relatively larger proportion of smaller capitalization stocks, so all three of the smaller size quintiles are grouped together in the sample labeled small-cap stocks. This procedure results in approximately equal sample sizes within each size category. For example, the early 1963 to 1972 sub-period sample includes 313 large-cap security VAR estimations, 319 mid-caps, and 337 small-caps. Approximately equal groupings allow for a direct comparison of mean coefficient standard errors, which depend in part on sample size. The impulse response functions in Table 5 are clearly stronger for small-cap securities, with positive value and statistical significance for most of the individual coefficient estimates, even in the most recent sub-period. The mid-cap results are more mixed, and there is little evidence of the disposition effect in the large-cap stocks. In the first sub-period, some of the large-cap coefficients are significantly negative, explaining the weak results in the first sub-period for all stocks.

Table 6 is similar to Table 5, but with impulse response function results for the security turnover impact of lagged market returns instead of lagged security returns. The first row of Table 6 corresponds to Figure 3 panel (b). The results are noteworthy in terms of consistency of the findings that market returns positively impact security turnover, which we interpret as evidence of investor overconfidence. Almost all of the mean coefficients in Table 5 are positive and highly significant including all sub-periods and size groupings. The magnitude of the mean values suggest that overconfidence based trading is more pronounced for smaller capitalization stocks. There also appears to be an almost monotonic decline in mean coefficient magnitudes over time, which the largest values in the first sub-period, and the smallest values in last sub-period. Consequently, the weakest results are in the large-cap sample of the latest 1993-2002 sub-period, where the first lagged coefficient value is only 1.03 percent, and some of the later lags are not statically significant.

Tables 5 and 6 collectively produce our third key empirical finding; the impact of both disposition and overconfidence trading is more pronounced in small capitalization stocks, and appears to be declining over time. Both the sub-period and size regularities may be due to the impact of retail level investors, who own a higher proportion of smaller capitalization stocks, and whose share of total trading volume has declined in recent years (see Gompers and Metrick (2001), among others).

4.4 Security Return Prediction

The prediction of security returns based on prior return and volume data has been a major focus of empirical research for many years. Researchers have examined the data from almost every possible perspective in numerous tests of market efficiency and equilibrium pricing models. Typically, security return prediction is based on powerful cross-sectional tests that sort securities in to factor mimicking portfolios. For example, the Fama and French (1993) size and value factors, and the Jegadeesh and Titman (1993) momentum factors are statistically significant predictors of the cross-sectional variation in security returns. Event studies also bring statistically powerful tests to bear on various corporate actions through tight event windows and large cross-sectional samples.

In contrast, the time-series nature of our methodology is not as powerful, making the turnover findings above even more striking. As mentioned previously, the market-wide results in Table 3 and Figure 2 panels (c) and (d) indicate that market returns conform to weak-form efficiency in that lagged turnover and returns have little predictive power. We explore the predictive power of the security VAR's in Table 7 and Figure 3 panels (c) and (d) to address some of the implications in Daniel, Hirshleifer, and Subrahmanyam (1998) (hereafter DHS). The key implication of DHS is that investor overconfidence in private versus public signals reconciles short-term overreaction (positive autocorrelation in returns) and long-term correction (negative autocorrelation in returns) into one model of security price formation. DHS do not explicitly mention trading volume, although a plausible extension might suggest that private signals generate higher contemporaneous volume than public signals. For example, the public announcement that a firm has accounting regularities should cause a large price drop, although little volume if investors agree on the informational content of the announcement. On the other hand, private signals lead to heterogeneous beliefs about security value. Heterogeneous beliefs with inappropriately high confidence levels lead to differences of opinion and consequently high trading volume.

The first row of Table 7 gives mean security return impulse response coefficients to a one standard deviation shock in security turnover, based on the 1,878 security VAR estimations summarized in Table 4. A one standard deviation shock to the turnover residual results in a small positive mean coefficient in the first month, followed by negative coefficients in subsequent months. For example, the second and fifth lag mean coefficients are significantly negative, as shown by the confidence bands in Figure 3 panel (c). This pattern might be consistent with the notion that a private signal leads to a short-term over reaction and longer-term correction. However, this interpretation assumes that the private signal was favorable for the security (i.e., lead to an upward revision in estimated value).

To test this possibility, we segregate turnover observations into positive-return and negative-return observations, based on the sign on the contemporaneous security return. Higher volume in a month with a contemporaneous positive return indicates a favorable private signal, and higher volume in a month with a contemporaneous negative return indicates an unfavorable private signal. If our interpretation of DHS is correct, the lagged impact of positive-return volume should be negative returns as the market corrects, and the lagged impact of negative-return volume should be positive returns as the market corrects. The second and third rows in Table 7 test this prediction. While the return coefficients in the second (positive-return turnover) row are in fact more negative than the third (negative-return turnover) row, the differences appear to be minor. We interpret this as only a weak validation of the DHS overconfidence prediction.

For completeness, we include an impulse response function plot of security return shocks on security return in Figure 3 panel (d). The first lagged coefficient is negative, although only weakly significant. Subsequent months turn positive although without statistical significance. The results are similar to the well-established cross-sectional finding that stocks exhibit positive momentum in returns up to a year, except for the first month which often exhibits negative autocorrelation perhaps due to bid-ask bounce in price quotations.

5. Summary and Conclusion

An old Wall Street adage warns investors not to "confuse brains with a bull market." More recently, formalized theories of investor overconfidence have been developed that motivate an empirical analysis of time-series patterns in both market-wide and security-specific trading activity. The proposition that some investors become overconfident about the value of active trading after positive portfolio returns, and less overconfident after negative portfolio returns, can be examined by multivariate (turnover and return) time-series analysis using vector autoregressions (VAR) and associated impulse response functions.

We first document a material and statistically significant tendency for market-wide turnover to increase in the months following high market returns, after accounting for contemporaneous and lagged volatility associations. The market-wide results we document can also be interpreted as a manifestation of previously documented disposition effect trading. We extend the VAR and impulse response function specifications to time-series on individual stocks in an attempt to distinguish between the two theories. In support of the overconfidence hypothesis, we find that security turnover levels are responsive to past market returns even when past security returns are included in the model. We show that both overconfidence and disposition effect trading are more pronounced in small-cap stocks and in earlier periods where individual investors hold a greater proportion of shares. We also test the return autocorrelation predictions of formal overconfidence models and find some confirmatory results.

Our central finding of a lead-lag relationship between market returns and turnover confirms the conventional wisdom of market making professionals, as well as formal theories of investor overconfidence. Market makers (i.e., specialists and brokers) typically celebrate market-wide gains because they portend increases in trading activity and market making revenues. The economic significance of high market returns on subsequent volume can be substantial. The estimated VAR indicates that a high market return of say 7 percent compared to -5 percent in a given month leads to a full “extra month” of trading volume spread out over the next six months. The long-run impact on non-detrended share turnover may be much greater.

While our analysis was motivated by theories of investor overconfidence, the finding that trading activity is highly dependent on past returns is an important empirical fact that should be appreciated by theorists and empirical researchers, independent of one’s interpretation. The contemporaneous association between volume and return volatility has been understood in financial economics literature for many years. More recently, the predictability of security returns based on lagged volume has been documented as a possible violation of strict market efficiency. Our research adds the new finding that trading volume is dependent on past returns over many months. We interpret aspects of this finding as confirmation of the overconfidence hypothesis that motivated the study, although precise distinctions between overconfidence trading and the disposition effect are somewhat subjective. Our findings suggest that further theoretical development and empirical research on investor types (i.e., individual versus institutional) may yield finer distinctions and testable implications.

Table 1
Market descriptive statistics

		Full	A	B	C	D
Period:		1962:08-2002:12	1963-1972	1973-1982	1983-1992	1993-2002
Observations:		485	120	120	120	120
Market Turnover	Mean	4.04	1.48	2.38	5.37	7.07
	Std Dev	2.52	0.40	0.99	1.02	1.72
	Min	0.73	0.73	1.06	3.79	4.23
	Max	13.50	2.45	6.14	10.58	13.50
Detrended Log	Mean	0.00	0.23	-0.70	-0.26	0.45
Market Turnover	Std Dev	14.25	15.13	17.70	12.65	10.68
(*mturn*)	Min	-45.52	-45.52	-38.51	-27.71	-22.63
	Max	52.73	35.41	43.27	52.73	34.29
Market Return	Mean	0.94	0.87	0.72	1.31	0.83
(*mret*)	Std Dev	4.35	3.63	5.07	4.49	4.05
	Min	-21.97	-10.52	-11.75	-21.97	-14.96
	Max	16.61	9.42	16.61	13.03	9.69
Market Volatility	Mean	3.70	2.80	4.37	3.67	3.97
(*misg*)	Std Dev	2.06	1.59	1.79	2.40	2.09
	Min	0.58	0.58	1.43	1.20	0.99
	Max	24.52	12.41	11.31	24.52	14.12
Dispersion (*disp*)	Mean	7.10	6.08	7.25	6.79	8.36
	Std Dev	1.88	1.13	1.82	1.01	2.40
	Min	3.93	3.93	4.72	4.59	5.54
	Max	15.76	8.98	14.88	9.93	15.76

This table gives descriptive statistics (reported in percentage points) on market-wide variables, where the market is defined as the value-weighted composite of all NYSE/AMEX non-fund common stocks Market Turnover is the monthly value-weighted turnover (shares traded divided by outstanding shares). Detrended Log Market turnover (*mturn*) is the Hodrick-Prescott (1997) detrended natural log of market turnover. Market Return (*mret*) is the monthly value-weighted market return. Market Volatility (*misg*) is the French, Schwert and Stambaugh (1987) monthly volatility measure based on daily return standard deviation. Dispersion (*disp*) is the monthly cross-sectional standard deviation of security returns. The first column reports on the full sample period, August 1962 to December 2002, and the last four columns report on four 120 month (decade-long) sub-samples.

Table 2
Security Descriptive Statistics

		Full	A	B	C	D
Period:		1962:08-2002:12	1963-1972	1973-1982	1983-1992	1993-2002
Observations:		530,608	121,154	146,367	134,065	134,067
Turnover	Mean	4.69	3.14	3.05	5.54	7.27
	Std Dev	6.14	5.11	3.44	5.98	8.31
	Min	0.00	0.01	0.00	0.01	0.00
	Max	709.23	322.05	137.88	449.23	709.23
Detrended Log	Mean	0.00	1.22	-1.42	0.08	0.40
Turnover (*turn*)	Std Dev	48.41	49.12	49.22	49.72	45.40
	Min	-416.79	-224.01	-364.75	-416.79	-341.69
	Max	464.54	464.54	420.89	384.42	350.37
Return (*ret*)	Mean	1.26	1.19	1.36	1.36	1.12
	Std Dev	11.78	10.27	11.89	12.04	12.75
	Min	-98.13	-56.54	-78.95	-80.00	-98.13
	Max	1100.00	250.00	300.00	1100.00	452.63
Volatility (*sig*)	Mean	9.23	8.57	9.62	9.11	9.60
	Std Dev	6.77	5.50	6.16	7.40	7.76
	Min	0.00	0.00	0.00	0.00	0.00
	Max	624.27	111.85	133.18	624.27	399.06

This table gives descriptive statistics (measured in percentage points) on security variables for all NYSE/AMEX non-fund common stocks with sufficient (120 months of contiguous) time-series data. Turnover is the monthly turnover (shares traded divided by outstanding shares) for each stock. Detrended log turnover (*turn*) is the Hodrick-Prescott (1997) detrended natural log of turnover. Return (*ret*) is the monthly stock return with dividends. Volatility (*sig*) is the French, Schwert and Stambaugh (1987) monthly volatility measure based on the daily security return standard deviation. The first column reports on the full sample period, August 1962 to December 2002, and the last four columns report on four 120 month (decade-long) sub-samples.

Table 3
Market VAR Estimation

Lagged Market Turnover

		$mturn_{t-1}$	$mturn_{t-2}$	$mturn_{t-3}$	$mturn_{t-4}$	$mturn_{t-5}$	$mturn_{t-6}$	$mturn_{t-7}$	$mturn_{t-8}$	$mturn_{t-9}$	$mturn_{t-10}$
$mturn_t$	Coeff.	0.284	0.204	0.089	-0.107	0.109	-0.091	0.000	-0.118	0.114	-0.031
	SE	0.047	0.048	0.040	0.041	0.041	0.041	0.041	0.040	0.040	0.038
	p-value	(0.00)	(0.00)	(0.03)	(0.01)	(0.01)	(0.03)	(0.99)	(0.00)	(0.00)	(0.42)
$mret_t$	Coeff.	0.019	-0.027	0.013	-0.001	-0.017	-0.009	0.012	-0.029	-0.007	0.002
	SE	0.018	0.018	0.015	0.015	0.015	0.015	0.015	0.015	0.015	0.014
	p-value	(0.28)	(0.13)	(0.38)	(0.96)	(0.25)	(0.56)	(0.41)	(0.06)	(0.66)	(0.90)

Lagged Market Return

		$mret_{t-1}$	$mret_{t-2}$	$mret_{t-3}$	$mret_{t-4}$	$mret_{t-5}$	$mret_{t-6}$	$mret_{t-7}$	$mret_{t-8}$	$mret_{t-9}$	$mret_{t-10}$
$mturn_t$	Coeff.	0.819	0.433	-0.065	0.042	0.220	-0.214	-0.172	-0.002	0.149	-0.118
	SE	0.133	0.136	0.128	0.123	0.121	0.122	0.122	0.122	0.121	0.120
	p-value	(0.00)	(0.00)	(0.61)	(0.74)	(0.07)	(0.08)	(0.16)	(0.99)	(0.22)	(0.33)
$mret_t$	Coeff.	-0.024	-0.002	0.045	-0.028	0.125	-0.070	0.024	-0.028	0.050	0.001
	SE	0.050	0.050	0.047	0.046	0.045	0.045	0.045	0.045	0.045	0.045
	p-value	(0.63)	(0.97)	(0.34)	(0.55)	(0.01)	(0.13)	(0.60)	(0.54)	(0.27)	(0.99)

Exogenous Variables

		Constant	$msig_t$	$msig_{t-1}$	$msig_{t-2}$	$disp_t$	$disp_{t-1}$	$disp_{t-2}$
$mturn_t$	Coeff.	-0.080	1.712	-1.033	-0.762	5.024	-1.583	-2.440
	SE	0.024	0.296	0.338	0.339	0.416	0.497	0.470
	p-value	(0.00)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)
$mret_t$	Coeff.	0.005	-0.999	0.069	0.602	0.946	-0.463	-0.264
	SE	0.009	0.110	0.125	0.126	0.154	0.185	0.174
	p-value	(0.60)	(0.00)	(0.58)	(0.00)	(0.00)	(0.01)	(0.13)

This table reports coefficient (Coeff.), coefficient standard errors (SE), and t-statistic significance levels (p-value) from a VAR of Detrended Logged Market Turnover (*mturn*) and Market Return (*mret*), with ten lags, for the full sample shown in Table 1. The VAR also includes contemporaneous and two lags of the exogenous variables Market Volatility (*misg*) and Dispersion (*disp*), as described in Table 1.

Table 4
Security VAR estimations

		Lagged Turnover									
		$turn_{t-1}$	$turn_{t-2}$	$turn_{t-3}$	$turn_{t-4}$	$turn_{t-5}$	$turn_{t-6}$	$turn_{t-7}$	$turn_{t-8}$	$turn_{t-9}$	$turn_{t-10}$
$turn_t$	Mean Coeff.	0.306	0.043	0.039	-0.032	-0.008	-0.003	-0.024	-0.033	0.000	-0.031
	p-value	(0.00)	(0.27)	(0.26)	(0.29)	(0.45)	(0.47)	(0.32)	(0.26)	(0.50)	(0.25)
ret_t	Mean Coeff.	0.005	-0.012	0.000	-0.006	-0.008	-0.003	-0.004	-0.001	-0.004	-0.004
	p-value	(0.29)	(0.11)	(0.49)	(0.27)	(0.19)	(0.37)	(0.32)	(0.45)	(0.33)	(0.33)
		Lagged Return									
		ret_{t-1}	ret_{t-2}	ret_{t-3}	ret_{t-4}	ret_{t-5}	ret_{t-6}	ret_{t-7}	ret_{t-8}	ret_{t-9}	ret_{t-10}
$turn_t$	Mean Coeff.	0.171	0.043	0.030	0.069	0.036	0.071	0.064	0.041	0.049	0.087
	p-value	(0.33)	(0.46)	(0.47)	(0.43)	(0.46)	(0.43)	(0.44)	(0.46)	(0.45)	(0.41)
ret_t	Mean Coeff.	-0.053	-0.021	-0.010	-0.007	0.011	0.014	0.013	0.003	0.024	0.020
	p-value	(0.33)	(0.43)	(0.47)	(0.48)	(0.46)	(0.45)	(0.46)	(0.49)	(0.42)	(0.43)
		Lagged Market Return									
		$mret_{t-1}$	$mret_{t-2}$	$mret_{t-3}$	$mret_{t-4}$	$mret_{t-5}$	$mret_{t-6}$	$mret_{t-7}$	$mret_{t-8}$	$mret_{t-9}$	$mret_{t-10}$
$turn_t$	Mean Coeff.	0.990	0.709	0.324	0.128	0.397	-0.072	-0.034	-0.088	0.254	-0.061
	p-value	(0.00)	(0.00)	(0.01)	(0.18)	(0.00)	(0.31)	(0.41)	(0.27)	(0.03)	(0.33)
ret_t	Mean Coeff.	0.228	-0.031	0.064	-0.036	0.050	-0.072	0.005	-0.080	-0.002	-0.008
	p-value	(0.00)	(0.33)	(0.17)	(0.28)	(0.21)	(0.11)	(0.46)	(0.08)	(0.49)	(0.44)

This table reports mean coefficients (Mean Coeff.) and t-statistic significance levels (p-value) from 1,878 individual security VARs on Detrended Logged Turnover (*turn*), Return (*ret*), and Market Return (*mret*), with ten lags, for the full sample shown in Table 3. The VAR estimations include contemporaneous and two lags of the exogenous variable Volatility (*sig*), as described in Table 3. Exogenous variable and *mret* dependent variable coefficient means are not reported for the sake of brevity. The p-values in this table are based on boot-strap based standard errors, and test the null hypothesis that the mean coefficient estimate across all securities is zero.

Table 5
Security Turnover Impulse Response Function: Security Return Shock

	Lags:	*1*	*2*	*3*	*4*	*5*	*6*
Full sample:	Mean Coeff.	2.29%	1.45%	1.03%	0.96%	0.62%	0.99%
	p-value	(0.02)	(0.13)	(0.26)	(0.27)	(0.47)	(0.24)
All stocks							
A: 1963 - 1972	Mean Coeff.	1.60%	0.50%	0.02%	0.22%	0.68%	0.48%
	p-value	(0.02)	(0.44)	(0.98)	(0.71)	(0.25)	(0.41)
B: 1973 - 1982	Mean Coeff.	4.16%	3.38%	2.35%	1.20%	0.69%	1.44%
	p-value	(0.00)	(0.02)	(0.08)	(0.35)	(0.58)	(0.24)
C: 1983 - 1992	Mean Coeff.	1.41%	0.41%	0.42%	1.01%	0.28%	0.36%
	p-value	(0.18)	(0.69)	(0.67)	(0.29)	(0.76)	(0.70)
D: 1993 - 2002	Mean Coeff.	-0.13%	-0.17%	-0.09%	0.48%	0.55%	0.75%
	p-value	(0.88)	(0.85)	(0.91)	(0.54)	(0.48)	(0.32)
Large-cap Stocks							
A: 1963 - 1972	Mean Coeff.	-2.71%	-2.64%	-2.09%	-1.02%	-0.74%	-1.09%
	p-value	(0.00)	(0.00)	(0.00)	(0.06)	(0.17)	(0.04)
B: 1973 - 1982	Mean Coeff.	2.11%	1.93%	1.90%	0.32%	0.07%	0.74%
	p-value	(0.14)	(0.17)	(0.16)	(0.81)	(0.96)	(0.55)
C: 1983 - 1992	Mean Coeff.	0.39%	-0.28%	-0.59%	0.30%	-0.20%	0.01%
	p-value	(0.70)	(0.78)	(0.54)	(0.75)	(0.82)	(1.00)
D: 1993 - 2002	Mean Coeff.	-1.34%	-0.86%	-0.41%	-0.06%	0.10%	-0.09%
	p-value	(0.13)	(0.33)	(0.62)	(0.94)	(0.89)	(0.91)
Mid-cap stocks							
A: 1963 - 1972	Mean Coeff.	3.12%	1.56%	0.23%	0.42%	0.72%	0.83%
	p-value	(0.00)	(0.01)	(0.69)	(0.45)	(0.19)	(0.11)
B: 1973 - 1982	Mean Coeff.	4.48%	3.81%	2.32%	1.70%	1.02%	2.12%
	p-value	(0.00)	(0.01)	(0.09)	(0.19)	(0.42)	(0.09)
C: 1983 - 1992	Mean Coeff.	0.95%	0.10%	-0.08%	0.91%	0.03%	-0.38%
	p-value	(0.37)	(0.92)	(0.93)	(0.34)	(0.97)	(0.68)
D: 1993 - 2002	Mean Coeff.	-1.04%	-0.67%	-0.68%	0.00%	0.01%	0.63%
	p-value	(0.23)	(0.43)	(0.40)	(1.00)	(0.99)	(0.40)
Small-cap stocks							
A: 1963 - 1972	Mean Coeff.	4.17%	2.43%	1.78%	1.19%	1.98%	1.61%
	p-value	(0.00)	(0.00)	(0.01)	(0.07)	(0.00)	(0.01)
B: 1973 - 1982	Mean Coeff.	7.71%	5.27%	3.42%	1.65%	1.08%	0.92%
	p-value	(0.00)	(0.00)	(0.01)	(0.20)	(0.39)	(0.46)
C: 1983 - 1992	Mean Coeff.	2.70%	1.27%	1.73%	1.68%	0.92%	1.35%
	p-value	(0.01)	(0.22)	(0.08)	(0.08)	(0.32)	(0.14)
D: 1993 - 2002	Mean Coeff.	2.93%	1.55%	1.33%	1.97%	2.02%	1.97%
	p-value	(0.00)	(0.07)	(0.10)	(0.01)	(0.01)	(0.01)

This table reports mean impulse response functions coefficients (Mean Coeff.) for a security turnover response to a security return shock. The full-sample base case in the first row is shown in Figure 3 panel (a), based on the VAR estimation in Table 4. The other rows provide sensitivity analysis by time-period for all stocks, and then by time-period for large-cap (top size quintile), mid-cap (second size quintile), and small-cap (bottom three size quintiles) stocks. The significance levels (p-value) of the mean coefficient estimates are based on boot-strapped standard errors.

Table 6
Security Turnover Impulse Response Function: Market Return Shock

		1	*2*	*3*	*4*	*5*	*6*
Full sample:	Mean Coeff.	4.25%	4.65%	2.96%	1.97%	2.62%	0.99%
	p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
All stocks							
A: 1963 - 1972	Mean Coeff.	7.44%	8.15%	3.68%	4.38%	3.51%	4.02%
	p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
B: 1973 - 1982	Mean Coeff.	5.18%	5.87%	3.01%	3.22%	3.70%	1.00%
	p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.08)
C: 1983 - 1992	Mean Coeff.	3.81%	3.34%	2.53%	-0.25%	1.25%	0.27%
	p-value	(0.00)	(0.00)	(0.00)	(0.12)	(0.00)	(0.10)
D: 1993 - 2002	Mean Coeff.	2.39%	1.95%	2.34%	0.37%	1.63%	-0.27%
	p-value	(0.00)	(0.00)	(0.00)	(0.08)	(0.00)	(0.20)
Large-cap Stocks							
A: 1963 - 1972	Mean Coeff.	5.94%	6.39%	2.41%	2.92%	2.61%	2.89%
	p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
B: 1973 - 1982	Mean Coeff.	4.35%	4.00%	2.32%	3.31%	2.91%	0.18%
	p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.61)
C: 1983 - 1992	Mean Coeff.	2.31%	2.30%	1.72%	-1.10%	1.19%	0.34%
	p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.17)
D: 1993 - 2002	Mean Coeff.	1.03%	1.00%	1.56%	0.20%	1.20%	0.14%
	p-value	(0.00)	(0.00)	(0.00)	(0.35)	(0.00)	(0.51)
Mid-cap stocks							
A: 1963 - 1972	Mean Coeff.	7.16%	8.17%	3.72%	4.66%	3.57%	3.73%
	p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
B: 1973 - 1982	Mean Coeff.	5.64%	6.48%	3.53%	2.90%	3.89%	0.82%
	p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)
C: 1983 - 1992	Mean Coeff.	3.55%	3.65%	2.16%	0.02%	1.47%	0.82%
	p-value	(0.00)	(0.00)	(0.00)	(0.94)	(0.00)	(0.00)
D: 1993 - 2002	Mean Coeff.	2.50%	2.21%	2.47%	0.46%	1.83%	-0.16%
	p-value	(0.00)	(0.00)	(0.00)	(0.03)	(0.00)	(0.45)
Small-cap stocks							
A: 1963 - 1972	Mean Coeff.	9.10%	9.77%	4.80%	5.46%	4.28%	5.34%
	p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
B: 1973 - 1982	Mean Coeff.	5.62%	8.17%	2.99%	3.98%	4.90%	3.37%
	p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
C: 1983 - 1992	Mean Coeff.	5.30%	3.88%	3.56%	0.17%	1.08%	-0.32%
	p-value	(0.00)	(0.00)	(0.00)	(0.51)	(0.00)	(0.21)
D: 1993 - 2002	Mean Coeff.	3.84%	2.66%	3.06%	0.41%	1.82%	-0.98%
	p-value	(0.00)	(0.00)	(0.00)	(0.05)	(0.00)	(0.00)

This table reports mean impulse response functions coefficients (Mean Coeff.) for a security turnover response to a market return shock. The full-sample base case in the first row is shown in Figure 3 panel (b), based on the VAR estimation in Table 4. The other rows provide sensitivity analysis by time-period for all stocks, and then by time-period for large-cap (top size quintile), mid-cap (second size quintile), and small-cap (bottom three size quintiles) stocks. The significance levels (p-value) of the mean coefficient estimates are based on boot-strapped standard errors.

Table 7
Return Impulse Response Function: Unsigned and Signed Turnover Shock

Lags:	*1*	*2*	*3*	*4*	*5*	*6*
	Lagged turnover					
Mean Coeff.	0.21%	-0.47%	-0.20%	-0.31%	-0.46%	-0.32%
p-value	(0.23)	(0.01)	(0.29)	(0.10)	(0.01)	(0.08)
	Lagged turnover with positive returns					
Mean Coeff.	-0.01%	-0.47%	-0.21%	-0.31%	-0.33%	-0.26%
p-value	(0.95)	(0.03)	(0.36)	(0.17)	(0.21)	(0.24)
	Lagged turnover with negative returns					
Mean Coeff.	0.14%	-0.22%	-0.18%	-0.27%	-0.39%	-0.27%
p-value	(0.41)	(0.23)	(0.36)	(0.15)	(0.07)	(0.14)

This table reports mean impulse response function coefficients (Mean Coeff.) for a security return response to a security turnover shock across the 1,878 VAR estimations summarized in Table 4. The first row is the security return response to security turnover, as shown in Figure 3 panel (c). The second (third) row is for the security return response to security turnover that is associated with a contemporaneous positive (negative) security return.

Figure 1
Monthly turnover for NYSE/AMEX market with trend line

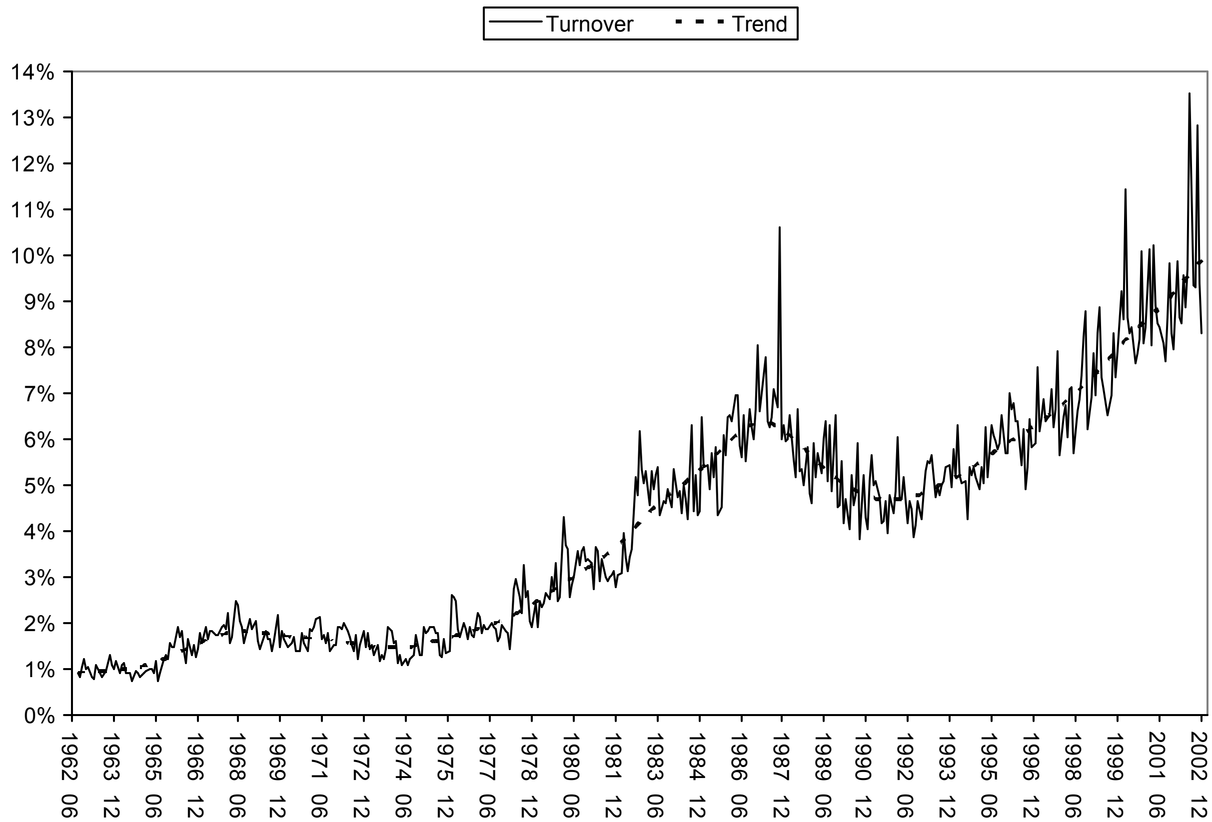
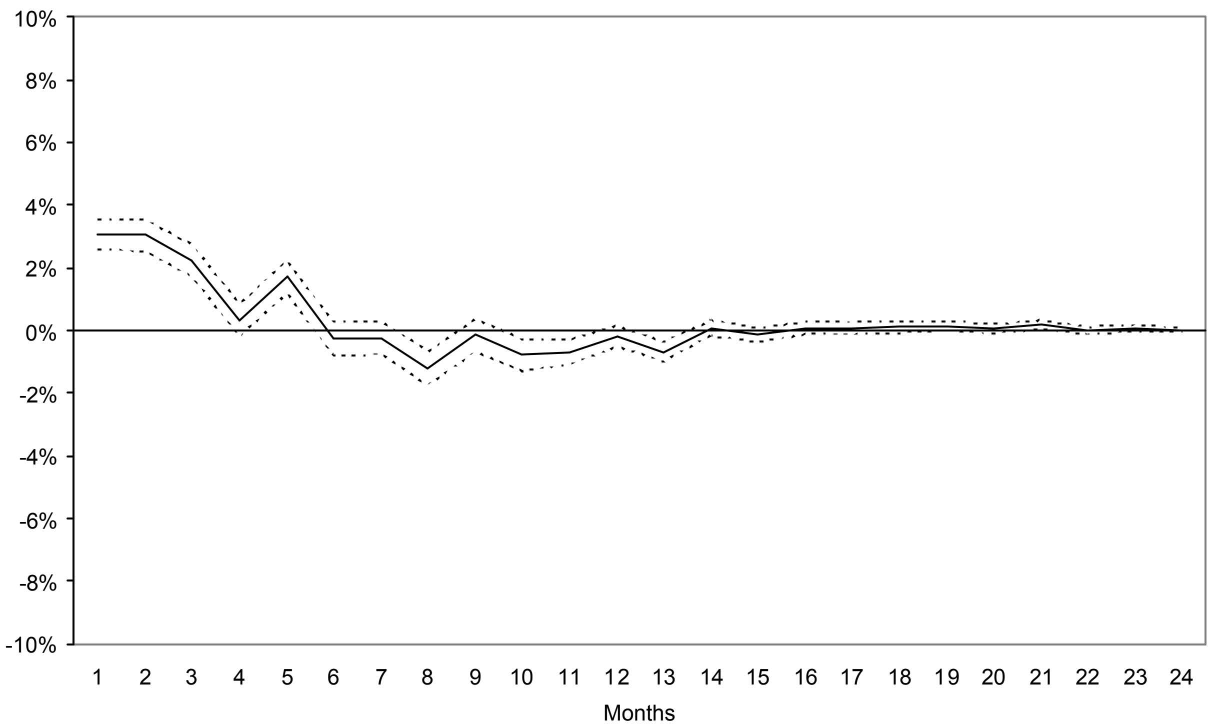


Figure 2
Market impulse response functions with two-standard error bands

(a) Market turnover response to a market turnover shock



(b) Market turnover response to a market return shock

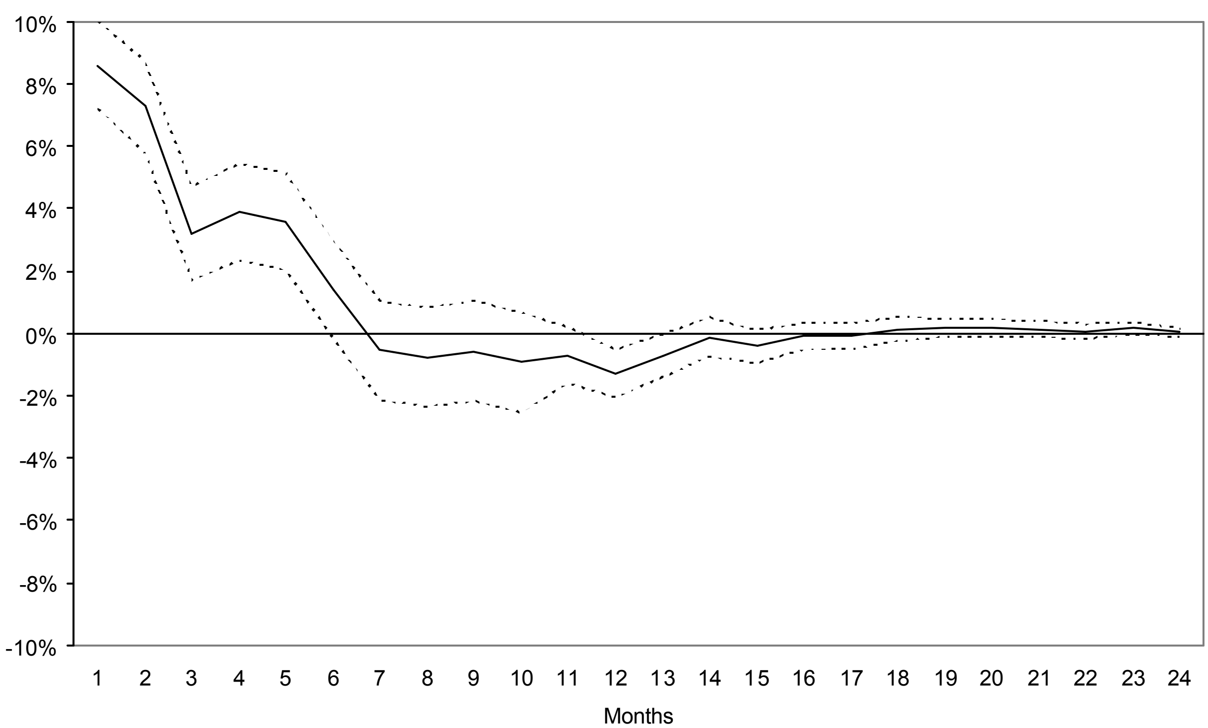
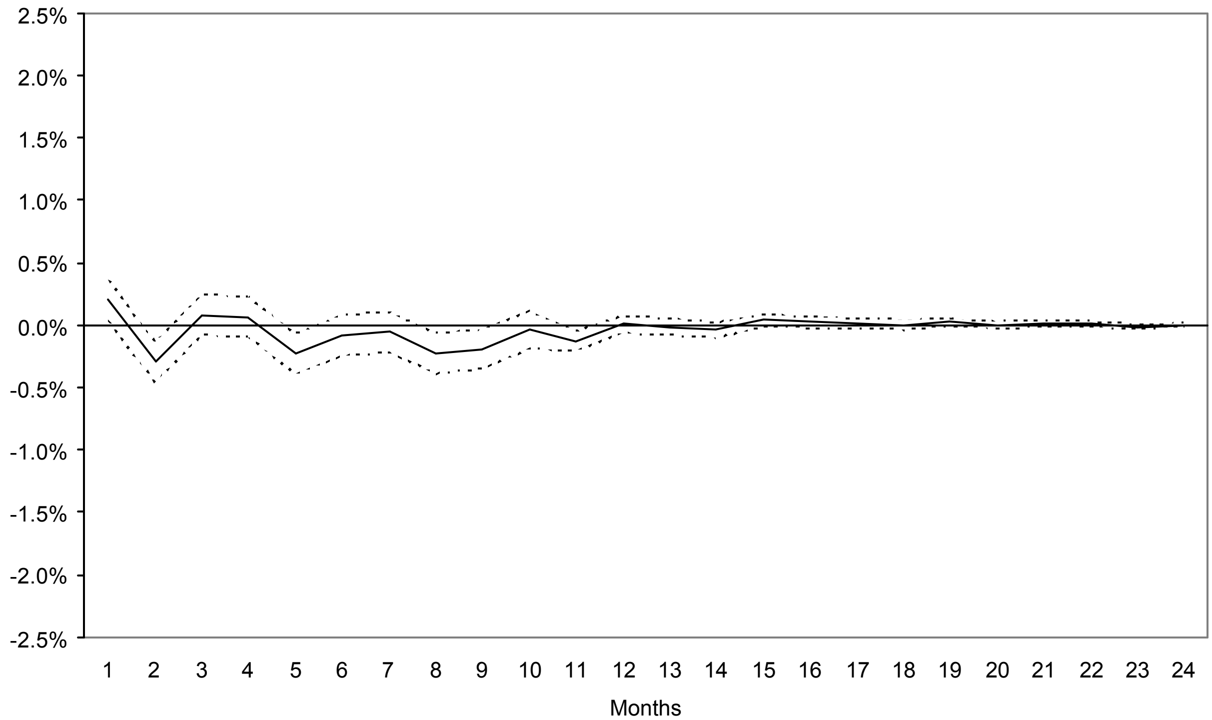


Figure 2 (continued)

(c) Market return response to a market turnover shock



(d) Market return response to a market return shock

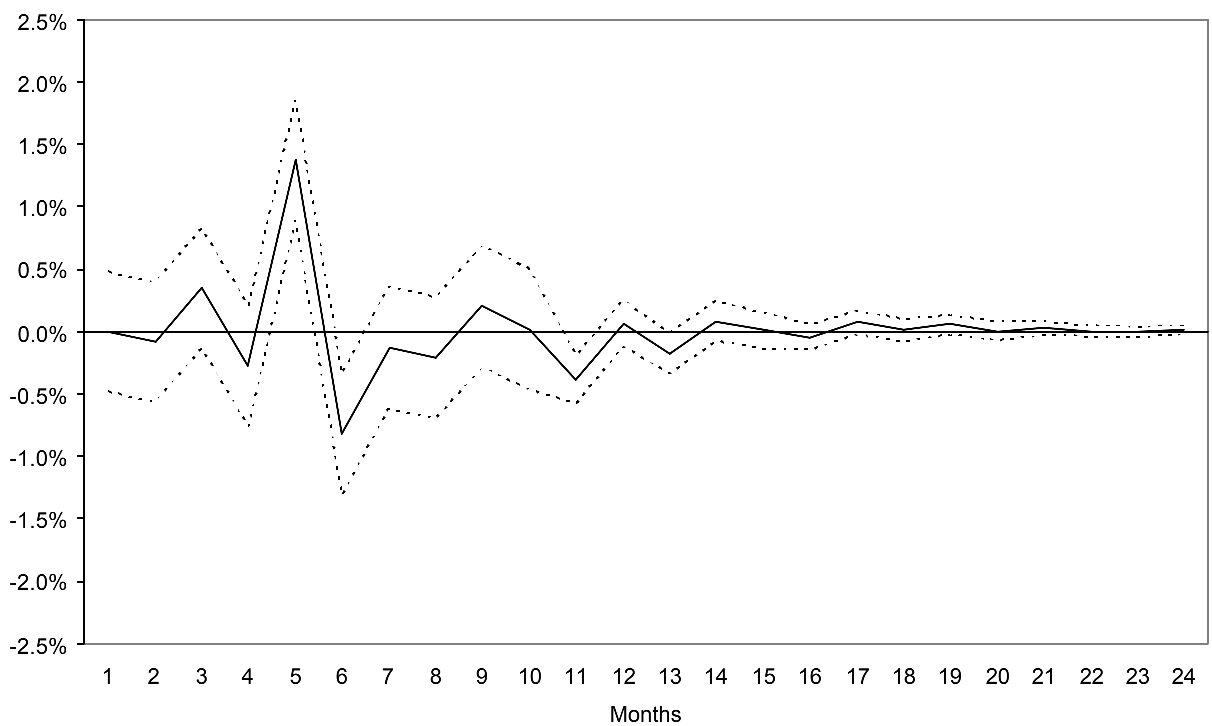
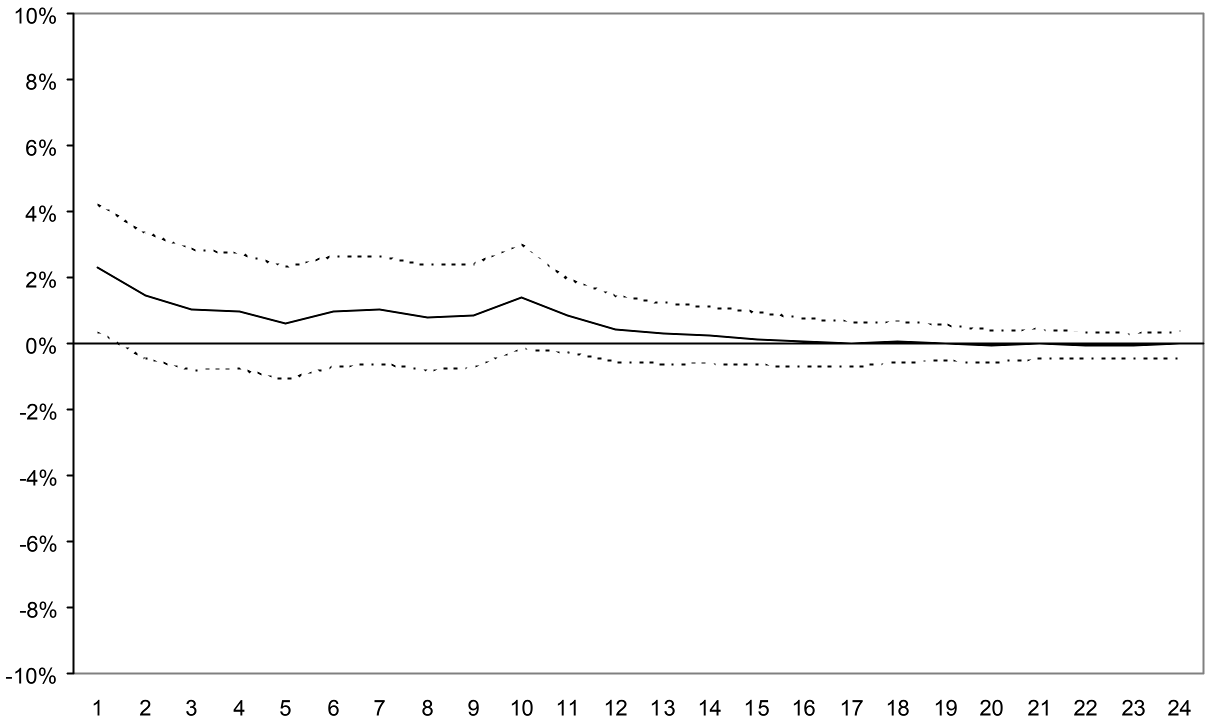


Figure 3
Security average impulse response functions with two-standard error bands

(a) Security turnover reponse to a security return shock



(b) Security turnover reponse to a market return shock

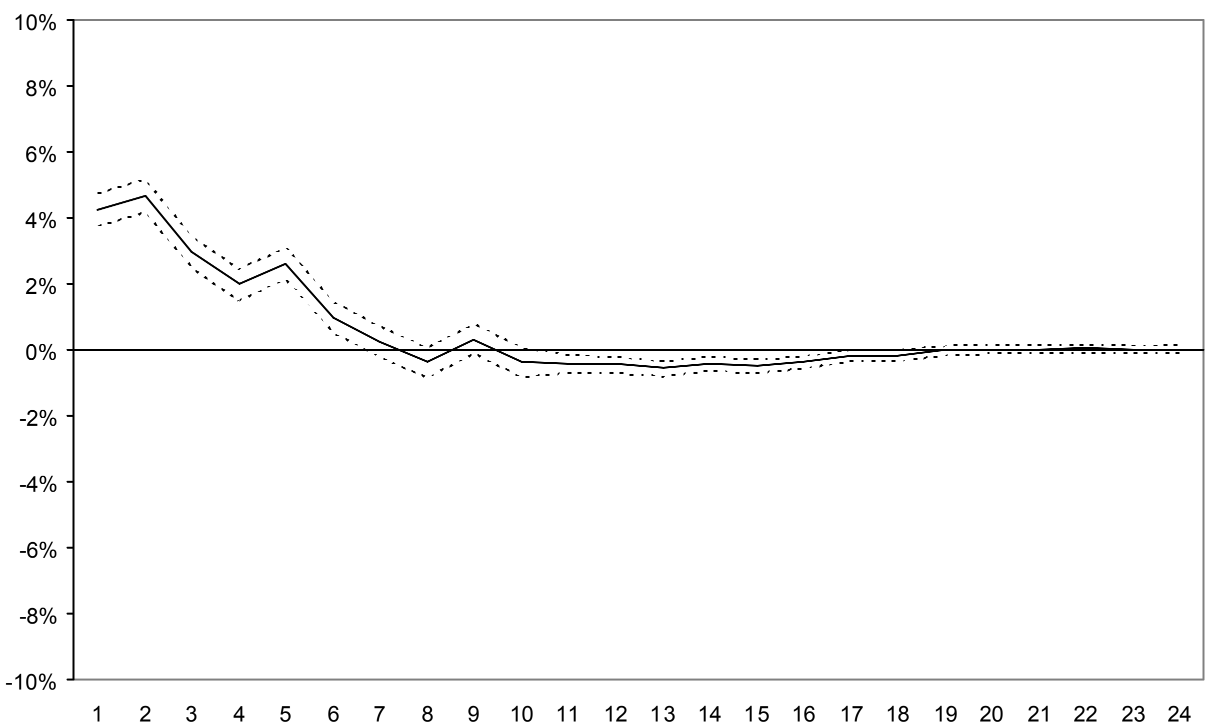
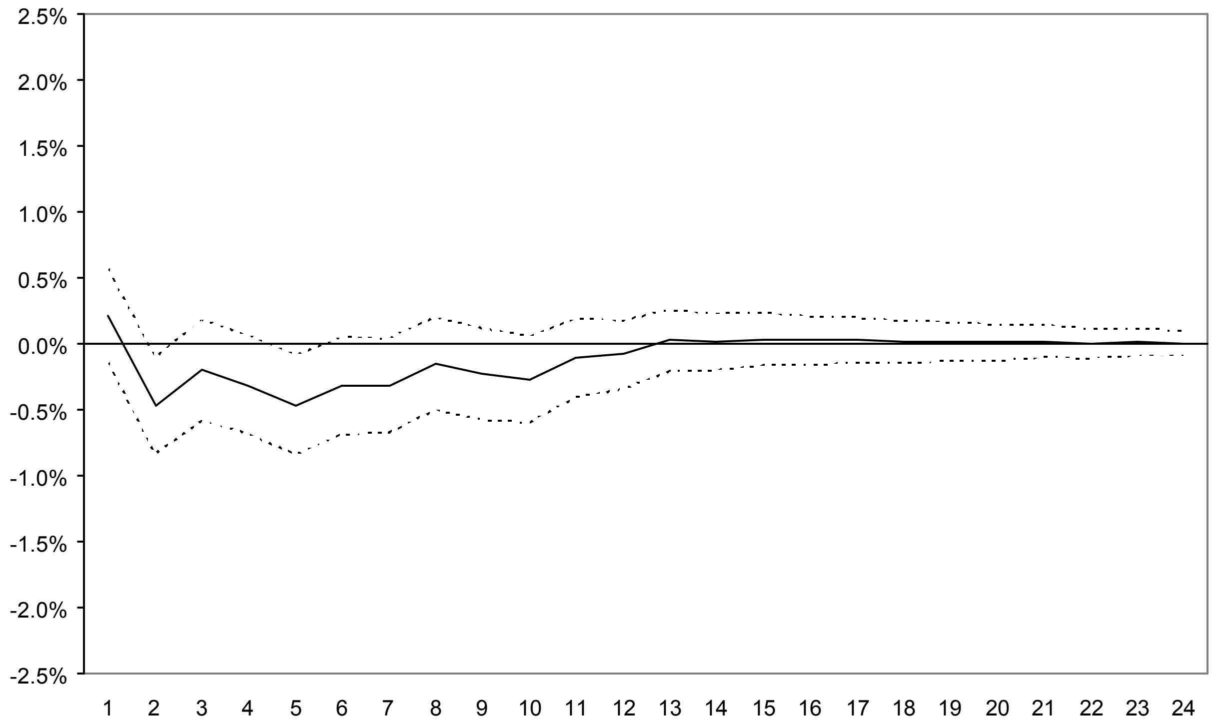
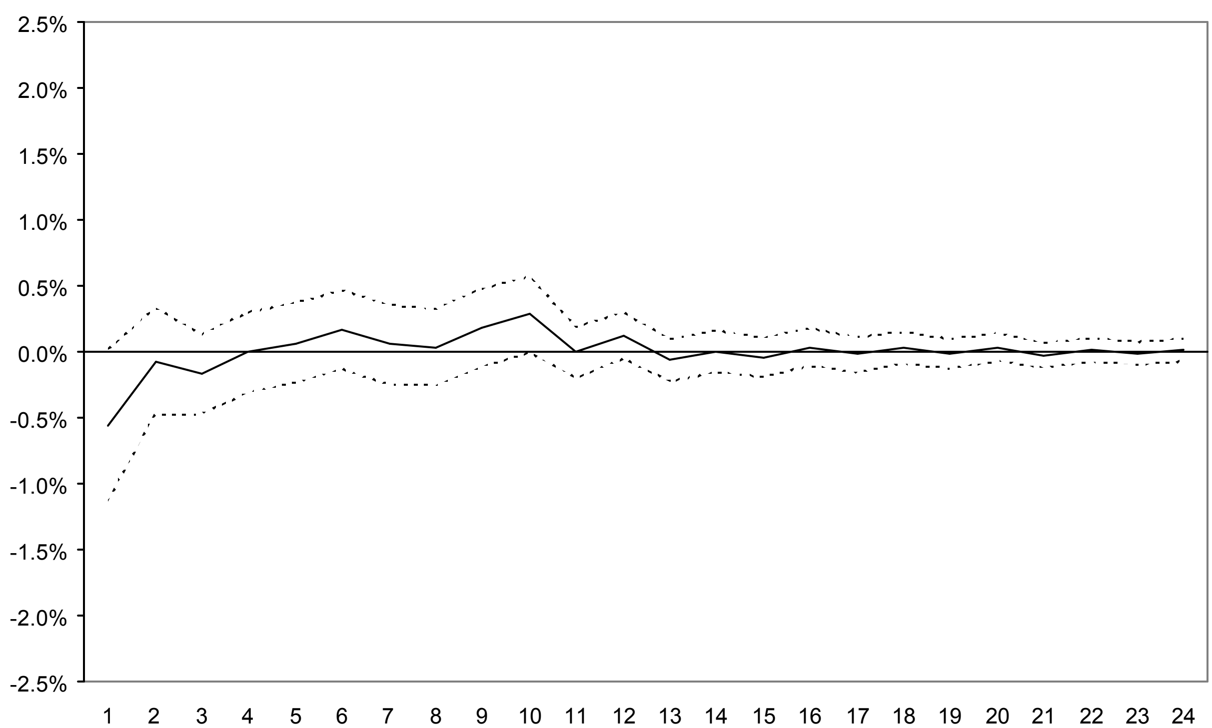


Figure 3 (continued)

(c) Security return response to a security turnover shock



(d) Security return response to a security return shock



References

Atkins, A., and E. Dyl, 1997, "Market structure and reported trading volume: NASDAQ versus NYSE," *Journal of Financial Research*, 20, 291-304.

Banz, R., 1981, "The relation between return and market value of common stocks," *Journal of Financial Economics*, 9, 3-18.

Bessembinder, H., K. Chan, and P. Seguin, 1996, "An empirical examination of information, differences of opinion, and trading activity," *Journal of Financial Economics*, 40, 105-134.

Bessembinder, H., and P. Seguin, 1993, "Price volatility, trading volume, and market depth: Evidence from the futures markets," *Journal of Financial and Quantitative Analysis*, 28, 21-39.

Black, F., 1986, "Noise," *The Journal of Finance*, 41, 529-543.

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

Chordia, T., R. Roll, and A. Subrahmanyam, 2000, "Market Liquidity and Trading Activity," UCLA working paper.

Chordia, T., and B. Swaminathan, 2000, "Trading Volume and Cross-Autocorrelation is Stock Returns," *The Journal of Finance*, 55, 913-935.

Cooper, M., (1999), "Filter rules based on price and volume in individual security overreaction," *The Review of Financial Studies*, 12, 901-935.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam, 1998, "Investor psychology and security market under- and over-reaction," *The Journal of Finance*, 53, 1839-1885.

Dhar, R., and N. Zhu, 2002, "Up close and personal: An individual level analysis of the disposition effect," Yale School of Management working paper.

Ferris, S., R. Haugen, and A. Makhija, 1988, "Predicting contemporary volume with historic volume at differential price levels: Evidence supporting the disposition effect," *The Journal of Finance*, 43, 677-697.

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

French, K., G. Schwert, and R. Stambaugh, 1987, "Expected stock returns and volatility," *Journal of Financial Economics*, 17, 5-26.

Gallant, R., P. Rossi, and G. Tauchen, 1992, "Stock prices and volume," *The Review of Financial Studies*, 5, 199-242.

Gervais, S., and T. Odean, 2001, "Learning to be overconfident," *The Review of Financial Studies*, 14, 1-27.

Gervais, S., R. Kaniel, and D. Mingelgrin, 2001, "The high-volume return premium," *The Journal of Finance*, 56, 877-919.

Glaser, M., and M. Weber, 2003, "Overconfidence and trading volume," *Center for Economic Policy* working paper.

Goetzmann, W., and M. Massa, 2003, "Disposition matters: Volume, volatility and price impact of a behavioral bias," NBER Working paper No. W9499.

Gompers, P., and A. Metrick, 2001, "Institutional investors and equity prices," *The Quarterly Journal of Economics*, 116, 229-259.

Grinblatt, M., and B. Han, 2002, "The disposition effect and momentum," NBER working paper No. W8734.

Hamilton, J., *Time Series Analysis*, (1994), Princeton University Press, Princeton, NJ.

Harris, M., and A. Raviv, 1993, "Differences of opinion make a horse race," *The Review of Financial Studies*, 6, 473-506.

Heath, C., S. Huddart, and M. Lang, 1999, "Psychological factors and stock option exercise," *Quarterly Journal of Economics*, 114, 601-627.

Hodrick, R.J. and E.C. Prescott, 1997, "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit, and Banking*, 29, 1-16.

Jegadeesh, N., and S. Titman, 1993, "Returns to buying winner and selling losers; Implications for stock market efficiency," *The Journal of Finance*, 48, 65-91.

Kahneman, D., and A. Tversky, 1979, "Prospect theory: An analysis of decision under risk," *Econometrica*, 47, 263-291.

Karpoff, J., 1987, "The relation between price changes and trading volume: A survey," *Journal of Financial and Quantitative Analysis*, 22, 109-126.

Lakonishok, J., and S. Smidt, 1986, "Volume for winners and losers: Taxation and other motives for trading," *The Journal of Finance*, 41, 951-974.

Lee, C., and C Swaminathan, 2000, "Price Momentum and Trading Volume," *The Journal of Finance*, 55, 2017-2069.

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

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

Lutkepohl, H., 1991, *Introduction to Multiple Time Series Analysis*, Springer-Verlag.

Odean, T., 1998, "Volume, volatility, price, and profit when all traders are above average," *Journal of Finance*, 53,1887-1934.

Odean, T., 1998b, "Are investors reluctant to realize losses?" *Journal of Finance*, 53, 1775-1798.

Phillips, P.C.B. and P. Perron, 1988, "Testing for a Unit Root in Time Series Regression," Biometrika, 75, 335-346.

Ranguelova, E., 2001, "Disposition effect and firm size: New evidence on individual trading activity," NBER working paper.

Runkle, D., 1987, "Vector Autoregressions and Reality," Journal of Business and Economic Statistics, 5: 437 – 442.

Reinganum, M., 1983, "The anomalous stock market behavior of small firms in January: empirical tests for tax-loss effects," *Journal of Financial Economics*, 12, 80-104.

Smidt, S., 1990, "Long-run trends in equity turnover," *The Journal of Portfolio Management*, Fall, 66-73.

Shalen, C., 1993, "Volume, volatility, and the dispersion of beliefs," *Review of Financial Studies*, 6, 405-434.

Shefrin, H., and M. Statman, 1985, "The disposition to sell winners too early and ride losers too long: Theory and evidence," *Journal of Finance*, 40, 777-791.

Stoll, H. and R. Whaley, 1987, "Program trading and expiration day effects," *Financial Analysts Journal*, 43, 16-28.