12-2014

The Nonlinear Price Fynamics of US Equity ETFs

Gunduz Caginalp
University of Pittsburgh

Mark DeSantis
Chapman University, desantis@chapman.edu

Akin Sayrak
University of Pittsburgh

Recommended Citation
Caginalp, Gunduz, Mark DeSantis, and Akin Sayrak. "The nonlinear price dynamics of US equity ETFs." *Journal of Econometrics* 183.2 (2014): 193-201. doi: 10.1016/j.jeconom.2014.05.009

The Nonlinear Price Fynamics of US Equity ETFs

The Nonlinear Price Dynamics of U.S. Equity ETFs

Gunduz Caginalp[1]
University of Pittsburgh

Mark DeSantis
Chapman University

Akin Sayrak
University of Pittsburgh

Abstract

We investigate the price dynamics of large market-capitalization U.S. equity exchange-traded funds (ETFs) in order to uncover trader motivations and strategy. We show that prices of highly liquid ETFs can deviate significantly from their daily net asset values. By adjusting for changes in valuations, we report the impact of non-classical variables including price trend and volatility using data from 2008 to 2011. We find a cubic nonlinearity in the trend suggesting traders are aware not only of the underreaction of others, but also self-optimize by anticipating others' reactions, and sell when the uptrend is stronger than usual.

Key words: Exchange-traded funds, momentum, volatility, and nonlinear dynamics.
JEL classification: G02, G12, G14, G17

[1] Corresponding Author: Gunduz Caginalp is a Professor of Mathematics at the Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260, E-mail: caginalp@pitt.edu, Phone: 412.624.8339, Fax: 412.624.8397.

1. Introduction

The study of price dynamics, i.e. the relative changes in asset prices, in financial markets poses important challenges for econometricians due to many confounding factors underlying the observed changes. In classical economics, one assumes that prices adjust to changing realities involving the values of assets. The question of whether there exist additional factors in price dynamics beyond valuation is an empirical one. However, testing of these factors is often complicated by the presence of noise, or randomness in news that influences perceptions of valuation (Black, 1986). This introduces considerable difficulty in testing of behavioral effects (e.g., trend) or classical effects (e.g., volatility), since the noise in valuations has the potential to obscure these effects. Furthermore, for ordinary stocks, the empirical measure of valuation is not completely clear or precise.

Exchange-traded funds (ETFs) present a rare opportunity to examine trading price dynamics effectively since their net asset values are known. In this paper we analyze 78 actively traded large market-capitalization ETFs using daily data from 2008 to 2011 with the objective of determining whether linear and non-linear factors beyond valuation play a role in the daily returns of ETFs. The main contributions of this paper are as follows:

1. We determine, as a preliminary goal, the extent to which the prices of a class of large volume and liquid ETFs deviate from their net asset value (NAV) and examine the factors beyond these deviations in the dynamics of observed daily returns.

2. Our methodology establishes that a nonlinear trend term showing both continuation of a trend (i.e., evidence of underreaction) and trend reversal (i.e., evidence of overreaction) are present in the data. Both underreaction and overreaction are distinguished and quantified in terms of the frequency with which a trend of that magnitude is observed.

3. Both short- and long-term volatility exhibit positive regression coefficients, which implies that increased volatility is associated with higher returns.

Existing literature regarding asset prices is primarily centered on multi-factor models based on market equilibrium. Specifically, the Fama-French three-factor (Fama and French, 1993) and the Fama-French-Carhart four-factor models (Carhart, 1997) provide the foundation for most research in this area by incorporating additional variables beyond the standard CAPM. These models and their results have many important theoretical and practical implications, especially with respect to selecting portfolios, measuring performance, and evaluating money managers over broad intervals of time measured in years. Our focus rests on understanding the short-term price dynamics of a specific group of assets, namely ETFs, in terms of both behavioral and classical variables including recent price trend, short- and long-term volatility. Our analysis does not explicitly challenge the notion that valuations are governed by equilibrium models. In essence, we set out to identify the factors behind price dynamics that are not explained by valuations.

In this paper, we account for the random changes in valuation via an appropriately defined valuation variable. We further reduce the "noise" in the data by utilizing a two-way fixed effects model (see Section 3.C) that accounts for between fund differences and contemporaneous correlations. Our methodology allows for the effects of variables such as the recent trend in price, the relative deviation between price and net asset value, and volatility to be illuminated.

The paper is organized as follows. Section 2 describes our approach and model selection. Section 3 provides a description of the data set, variable definitions, and the methodology. Section 4 provides the empirical results, and Section 5 concludes the paper.

2. Analyses and Model Selection

A. The challenge of "noise"

A basic axiom of classical finance is that asset prices should reflect the consensus opinion of the market participants and make rapid adjustments to the changing valuations. This leads immediately to the question of how traders and investors react to new information and make price adjustments.[2] We postulate that traders are aware not only of the stream of information that changes valuations, but also of the motivations of other traders as exhibited by changes in prices, volume and order books.

On the one hand, experimental (Smith, Suchanek, and Williams, 1988) and empirical (Caginalp and Constantine, 1995, and Duran and Caginalp, 2007) studies have suggested that the game theoretic aspect of anticipating others' actions leads to over- and underreactions that are highly significant in the dynamics of price formation. On the other hand, according to efficient market adherents, the fact that any trader can observe the same information suggests that these effects should be negligible. The question of whether such effects are trivialized by the competitive process inherent in markets is an empirical one. Empirical testing of such hypotheses, however, is often difficult due to the multitude of factors that influence valuations.

For example, statistical studies on market price data have often shown a very small or negligible price trend effect (see, for example, Poterba and Summers, 1988). This may be due to the explanation provided by the efficient market hypothesis, or it may be a consequence of the large amount of "noise" due to rapid and random changes in valuations that tend to mask such effects. One way to circumvent this "noise" problem is to examine financial instruments in which the underlying asset value is directly observable. In this way, one can examine the difference between the trading price and the net asset value. In a recent study Caginalp and DeSantis (2011) investigate the nature of the change in the trading price relative to the change in the net asset value in the context of closed-end funds.[3] Using a methodology that compensates for changes in valuation and other variables, they find that recent price trend, examined in a nonlinear context, is a highly significant factor in terms of return.

In this paper, we examine the price dynamics of a set of highly liquid ETFs using a

---

[2] Sturm (2013) presents an overview of the issue of noise and the related literature.

[3] Closed-end funds are mutual funds that trade independently of their NAV since investors do not have the option to redeem their shares as they do for open-end mutual funds. Closed-end funds often trade at chronic discounts, the causes of which have been under study for many years (Anderson and Born, 2002). Among these are tax liability, management fees, corporate structure, etc. Although these are all significant, they do not typically vary from day to day.

methodology motivated by Caginalp and DeSantis (2011). The sample of ETFs allows us to test various hypotheses within an environment of financial instruments that are actively traded by both institutional and individual investors. Furthermore, the ETFs have a redemption and creation mechanism (see Section 2.B) that anchors the price close to the net asset value, unlike closed-end funds. By examining price dynamics using high trading volume and high liquidity ETFs, one can attain the potential to provide strong evidence for a spectrum of classical and behavioral factors.

B. ETF discounts/premiums

As we focus on large market-capitalization and high-liquidity U.S. equity ETFs in this paper, we would expect the ETF prices not to deviate significantly or chronically from their NAVs.[4] We find that the ETFs in our sample do exhibit significant discounts/premiums.[5] It is highly unlikely our observation is due to lack of trading volume, as our sample of ETFs is typically quite liquid (see Appendix). For example, the daily trading volume of funds in our sample is more than 5.9 million shares on average. One potential reason might be that investors are utilizing ETFs in their short-term trading strategies. Another possible reason might be due to institutional features of equity markets (Madhavan, 2012). Alternatively, there might be hidden inefficiencies associated with the pricing mechanism for the ETFs in our sample.

ETFs were primarily designed as tradable assets that would closely track the value of their underlying securities. This is accomplished through various institutional investors who assume the role of "authorized participants." These investors are allowed to create and redeem shares of the ETFs as market prices diverge from their true values. Authorized participants realize the creation and

---

[4] While transaction costs of the ETF vary with quantity, we can obtain an order of magnitude estimate for a $1 million trade, for example, in one of the ETFs with volume in the mid-range of our sample, e.g., EWC. The bid-ask spread is almost always a penny for EWC (see also Borkovec and Serbin, 2013). For example, on April, 22, 2013 near the close one could execute a single order (e.g., on INET) for up to 234,500 shares to sell at $26.73, and up to 42,400 shares to buy at $26.74

[5] There are 168 data points with a valuation-spread in excess of 2% (in magnitude) within the data set. Of these observations 140 occurred during the time period 9/15/2008 – 12/15/2008. Further, 77 of the 168 observations were from five funds (EWC, IBB, PHO, UYG, and XME). This suggests that the majority of funds trade fairly close to their NAVs; however, during a crisis (with a large number of negative events) a large number of funds may have large valuation spreads (in magnitude) across several days.

redemption of ETF shares by buying and selling the underlying assets (stocks, bonds, etc.) and exchanging them with the company in charge of managing the ETF for a small fee. Essentially, the authorized participants act as arbitrageurs who trade with the objective of making a profit. Theoretically, this structure implies tight arbitrage bounds on the gap between the price and the value.

This arbitrage bound will depend on a number of factors that can be categorized as direct and indirect costs of arbitrage. The direct cost of arbitrage is related to the creation/redemption fees, which are usually fairly small (in the order of a few basis points for large market capitalization equity ETFs). The indirect costs can be various, but they most notably arise from the volatility and patterns of illiquidity in the underlying securities. The illiquidity issue is a form of the implementation cost related to the actual trading required to realize a perceived arbitrage opportunity. As illiquidity poses a challenge for the arbitrageur, unexpected changes in liquidity further add to the complications related to the execution of trades. Another possible source may be attributed to halts to the creation/redemption process due to regulatory scrutiny.

Suppose, for example, that the authorized participants for a particular ETF with NAV at $ 100 calculate that it is profitable for them to buy at $ 99.50 and sell at $ 100.50, but not otherwise due to trading costs, fees and risks. Then the price dynamics between these two values ($ 99.50 and $ 100.50) will not be affected by the authorized participants, but rather by other traders in the market. This kind of price dynamics is what we focus on in this paper.

During the past two decades, rapid trading (and more recently high-frequency trading) has occupied a greater share of market activity. The explosion of trading volume for some ETFs, for example, SPY, whose daily volume is often more than 50 times that of the most active individual stock, has grown in parallel with the shift toward shorter term trading as a profit mechanism over the traditional longer term bargain hunting for undervalued stocks (Demos, 2012).

The inefficiency that sometimes occurs with ETFs leads to the question of what governs the dynamics of trading prices for these ETFs. We examine this question in the context of our methodology. Through the use of appropriately defined variables and a methodology that accounts for firm heterogeneity and contemporaneous correlations (see Section 3), we find the change in ETF prices consists of either noise or non-classical dynamics including the influence of price trend, volatility change, and other variables.

C. Hypotheses development

The price dynamics of ETFs beyond valuation can be used to examine diverse theories of markets. The prevailing theory of the efficient market hypothesis would stipulate that given the information relating to valuation on day t, the return for day $t + 1$ should be a small constant term representing the risk-free return plus a risk premium, plus noise. This is, in effect, the null hypothesis. Alternatively, the asset flow approach to dynamics (Caginalp and Balenovich, 1999) suggests that factors such as price and volume trend, changes in volatility, etc. should play an important role, as traders observe these changes as cues to the motivations of other traders. Without infinite arbitrage capital, there will be non-trivial changes in price due to these factors, according to this theory.

Theory that has evolved from the laboratory asset markets starting with Smith, Suchanek and Williams (1988) also suggests prices can veer from fundamental value even when there is complete information on valuation, demonstrating that traders are reacting to the anticipated actions of others.

Parallel to these theories are several key assertions of behavioral finance such as under- and overreaction to changes in valuation (Madura and Richie, 2004), and anchoring (George and Hwang, 2004) whereby prices are influenced by significant markers of the past. Thus, the price dynamics of ETFs provide an important test of very different perspectives into financial markets.

3. Data, Variable Definitions, and Methodology

A. Data

As of December 2011 there were 1,374 actively-traded ETFs on the U.S. exchanges. Ninety-seven of these funds are U.S. equity ETFs with market capitalizations of at least $500 million. Due to data availability limitations of the Bloomberg system and to maintain a balanced panel (i.e., the same number of observations for each fund), we restrict our analysis to 78 of these funds. Our final panel consists of 77,376 data points.

Table 1 displays the market capitalization, daily volume, and valuation spread (defined as $(NAV - P)/NAV$, where P is the price) for the 78 ETFs included in this study over the time period January 2, 2008 through December 6, 2011. About half of the funds have a market capitalization of at least $1.75 billion and an average daily trading volume that is greater than 920,000 shares.

<< Insert Table 1 >>

As noted in Section 2.B, due to the ability of authorized participants to redeem their shares of an ETF for the underlying securities, theoretically, the price of an ETF should not vary much from its *NAV*. We compute the valuation spread on a daily basis for each of the 78 funds remaining in our sample. In Table 1, we note that the daily valuation spread ranged from a minimum of -13.24% to a maximum of 7.381% with a mean value of 0.001% and median of 0.002% across all funds. The mean valuation spread has a fairly low value, as would be predicted by the no-arbitrage argument.

However, averages could be misleading in this context, as values with opposite signs would cancel each other. To better examine the deviation between price and NAV we consider the magnitude of the deviation. As such, we determine the average of the absolute value of the daily valuation spread for each fund. We find that half of the funds in our sample have an average deviation of 0.093% in magnitude, and one-quarter of the funds have a deviation of at least 0.136%. These values are similar to those of Krause and Tse (2013) who found that the average daily difference in absolute value between prices and *NAV*s of five ETFs was 0.168% of the underlying price.

When the deviation between the ETF's trading price and NAV is sufficiently large, the authorized participants have an incentive to create or redeem shares. Yet, we find evidence consistent with previous research suggesting price deviations from the NAVs are not completely eliminated via arbitrage. Thus, when deviations are smaller than the cost of redemption, there is a price dynamics governed by a number of factors that can be uncovered by appropriate data analysis. That is, by accounting for the differences among funds and the contemporaneous correlations, we are able to measure how variables such as the valuation spread, recent trend in price, and volatility affect returns.

B. Variable Definitions

We utilize the variables detailed in Caginalp and DeSantis (2011) as the basis of this study. We use the daily adjusted-closing price, P_t, which incorporates all dividends as indicated in the "Total Return Index Gross Dividends" field available from the Bloomberg system. The dependent variable in the regressions discussed in Section 4 is the following day's return. Thus, if today is day t, we define the dependent variable as

$$R_{t+1} = \frac{P_{t+1} - P_t}{P_t}.$$

As we have shown that ETFs can trade at a discount/premium, we account for the observed discount/premium in the form of a valuation spread[6], which we define as

$$V_t = \frac{NAV_t - P_t}{NAV_t} - \frac{1}{3.2318}\sum_{k=1}^{10} e^{-0.25k}\,\frac{NAV_{t-k} - P_{t-k}}{NAV_{t-k}}.$$

The rationale for this choice is that, by subtracting a weighted average of past valuation spreads from the current spread, we account for any possible persistence in the discount/premium. We employ the negative exponential to account for the tendency that investors place more weight on recent events.

The relationship between the return and the recent trend in price is explored in detail in Section

[6] Bloomberg's CEF_PCT_PREM is defined as $(P_t - NAV_t)/NAV_t * 100$, where P_t is the closing price and NAV_t is the net asset value of the fund on day t. We multiply this field by -1/100 and utilize it in the definition of the *Valuation* variable.

4. The recent trend in price is defined as

$$T_t = \frac{1}{3.2318} \sum_{k=1}^{10} e^{-0.25k} \frac{P_{t-k+1} - P_{t-k}}{P_{t-k}}.$$

This definition corresponds to a weighted average over the past ten days including the current day. The negative exponential serves a similar purpose as in the definition of the valuation variable. In addition, it helps to serve as a smoothing function by allowing the effect of significant changes in the trend to slowly decay over time. The topic of trend or momentum has been studied previously. See, for example, Asem and Tian (2010), Hong and Stein (1999), Barberis, Shleifer, and Vishny (1998), and Daniel, Hirshleifer, and Subrahmanyam (1998).

Mean-reversion in asset prices over longer time periods has been observed in previous studies. We account for mean-reversion by using a long-term trend variable defined in the following way. First, we fit a linear regression to the past 252 returns including the current day. We define the long-term trend variable by annualizing the slope coefficient multiplying it by 252.

We utilize the M2 Money Supply[7] as reported by the Federal Reserve as a proxy for the amount of money in the system. Support for the consideration of this variable as a factor in returns has been shown in earlier empirical, theoretical, and experimental works (see Caginalp and DeSantis, 2011, and the references therein). Weekly data are available on the Federal Reserve website. We convert this to daily units through a linear interpolation. Then, we define the *Money Supply* variable as

$$MONEYSUPPLY_t = \frac{M2_t - M2_{t-1}}{M2_{t-1}}$$

where$M2_t$ represents the value of the interpolated $M2$ statistic on day *t*.

Classical methods postulate that volatility is an important factor in returns. We distinguish between the effects of two different forms of volatility based on the length of the time period considered. *Short-Term Volatility* (*STV*) corresponds to the standard deviation of returns over the past

[7] Following Caginalp and DeSantis (2011) we utilize an M2 figure that has not been adjusted for seasonality. This variable includes currency, traveler's checks, demand and other checkable deposits, retail money market mutual funds, savings, and small time deposits, and is measured in units of trillions of USD.

ten days, while *Long-Term Volatility* (*LTV*) incorporates volatility over the past year. We define *STV* and *LTV* as

$$STV_t = \left[\frac{1}{10}\sum_{i=t}^{t-10}\left(R_i - Mean\left(R_{[t-10,t]}\right)\right)^2\right]^{1/2}$$

$$LTV_t = \left[\frac{1}{251}\sum_{i=t}^{t-251}\left(R_i - Mean\left(R_{[t-251,t]}\right)\right)^2\right]^{1/2}.$$

Note that these formulae effectively compensate for the price trend. Caginalp and DeSantis (2011) observed an ambiguous relationship between volatility and returns in the context of closed-end funds. For example, one result found that volatility over the long term corresponded to lower returns, while short-term volatility related to higher returns in the context of closed-end funds. As noted in Section 4.A we do not find a similar result with the ETFs. Indeed, both volatility variables exhibit positive coefficients supporting the notion that higher volatility leads to higher returns.

An advantage of our approach is its flexibility to accommodate the notion of anchoring in asset prices. That is, we consider the thesis that a recent price may become fixed in a trader's mind (Tversky and Kahneman, 1974), and that the trader may make decisions based on that "anchored" price instead of other factors like valuation. Traders have long noted (while academicians have doubted) the "resistance" offered by a recent high price as prices move closer to this value. This can be regarded as an example of anchoring. We define our *Resistance* variable based upon a recent quarterly high price. Specifically, we assume that *Resistance* has occurred if the current price is within 85%-100% of this recent quarterly high price given by

$$H_t \coloneqq \max(P_s)\,, s \in [t-63, t-16].$$

We then define *Resistance* via the following formula:

$$Resistance = \begin{cases} 1, & conditions\ a\ and\ b\ hold \\ 0, & otherwise \end{cases}$$

where

$$Condition\ a\colon\ for\ s \in [t-15, t-10], P_s \leq 0.85H_t$$

$$Condition\ b\colon\ 0.85H_t \leq P_t \leq H_t.$$

Note that no restriction is placed upon P_s for s in $[t-9, t-1]$. On a longer time scale, the resistance effect is the subject of several recent works. For example, George and Hwang (2004) form three types of self-financing portfolios by ranking stocks on past performance as measured by the ratio of the current stock price to its recent 52-week high price for each month. In their study, the "winners" portfolio consists of the 30% of stocks with the highest ratio, the "losers" portfolio consists of the bottom 30%. They find that their strategy outperforms the momentum strategies of Jegadeesh and Titman (1993) and Moskowitz and Grinblatt (1999). Sturm (2008) extends this work by considering additional time intervals for the recent high. Utilizing a similar methodology and a *Resistance* variable as described above, Caginalp and DeSantis (2011) find marginal support for Resistance having a negative effect on future returns.

C. Methodology

In order to determine whether or not the variables defined in the previous section influence the ETF's daily relative price change, $R_{t+1} = (P_{t+1} - P_t)/P_t$, we examine various combinations of the independent variables in order to determine how they influence the price change. Our data set is a collection of separate time-series for each of the 78 funds under consideration. That is, we have a time-series-cross-sectional data set, which is the combination of time-series for several funds, i.e. cross-sections. This might also be referred to as "panel" data; though panel data typically consists of several groups with a small number of observations per group, whereas we have both a large number of groups (78) and a large number of observations (740) per group.

To facilitate comparisons of the relative importance of these different variables, e.g., price trend and valuation, in terms of the relative price change; we standardize the regression variables (by fund) by subtracting the mean and dividing by a standard deviation.[8] Thus, if two coefficients are approximately equal, it implies that a one standard deviation change in one variable produces the same

[8] We standardize all variables except for *Resistance*. Standardization of the highly skewed binary Resistance variable could overstate the significance of this factor (Gelman, 2008).

relative price change for the following day as a one standard deviation change in the other ceteris paribus. Simply put, this standardization allows the regression coefficients to provide a natural metric for comparing the impact of variables that are measured using very different scales.

To account for the hierarchical nature of the data and the possibility of contemporaneous correlations, we utilize a fixed-effects approach. Specifically, we use the SAS procedure PANEL with the FIXTWO option. This model has the following form[9]:

$$R_{it} = \beta' z_{i,t-1} + u_{it} \tag{1}$$

where $z_{i,t-1}$ is a column vector of independent variables, β' is a row vector of parameters to be estimated, and u_{it} is an idiosyncratic error term drawn from a mean zero distribution. Note that $i = 1, \cdots, 78$, and $t = 1, \cdots, 740$.

A contemporaneous correlation, e.g. a news announcement on day *t* affecting the entire market, occurs when u_{it} and u_{jt} are correlated. That is, even after removing the dependence of R_{it} and R_{jt} on the observed values of $z_{i,t-1}$ and $z_{j,t-1}$, there might be unobserved common components that determine them. Thus, in this case, $Cov(u_{it}, u_{jt}) \neq 0$.

We consider a two-way fixed-effects with the error term specified via:

$$u_{it} = \mu_i + \gamma_t + \varepsilon_{it} \tag{2}$$

where μ_i is a fund-specific effect drawn from a mean zero distribution, γ_t is a time-specific effect drawn from a mean zero distribution, and ε_{it} is an idiosyncratic error term drawn from a mean zero distribution. Thus, our model has the form:

$$R_{it} = \beta' z_{i,t-1} + \mu_i + \gamma_t + \varepsilon_{it} \tag{3}$$

where the SAS procedure estimates each μ_i and each γ_t separately – without imposing any restrictions on their marginal distributions. By including an intercept term in this procedure we ensure each set of effects is estimated subject to the restriction that it sums to zero.

[9] Although this analysis for performed for dependent variable R_{it} and regressors $z_{i,t-1}$, the results in Section 4 are for regressions with dependent variable $R_{i,t+1}$ and regressors $z_{i,t}$.

We consider μ, γ, and ε to be random variables with variances $Var(\mu)$, $Var(\gamma)$, and $Var(\varepsilon)$. Further, we assume these variables are independent of one another and draws of any one of them are i.i.d.

With this approach we have

$$Cov(u_{it}, u_{jt}) = E[(\mu_i + \gamma_t + \varepsilon_{it})(\mu_j + \gamma_t + \varepsilon_{jt})] = Var(\gamma) \tag{4}$$

where the independence assumptions eliminate the cross terms. In addition,

$$Var(u_{it}) = Var(\mu) + Var(\gamma) + Var(\varepsilon). \tag{5}$$

This implies the contemporaneous correlation of u_{it} in a straightforward two-way panel model is given by:

$$\rho = \frac{Var(\gamma)}{Var(\mu)+Var(\gamma)+Var(\varepsilon)}. \tag{6}$$

It is commonly accepted that stock prices exhibit "heavy tails." We winsorize the data on a per fund basis in order to mitigate the effect of a few dramatic price and/or NAV changes on the *Valuation* variable. That is, for each independent variable (excluding *Resistance*) for each fund we replace any values greater/less than the mean +/- three standard deviations with the value at +/- three standard deviations.[10]

4. Results and Discussion

A. Regression Results

First, we consider the following two variables, *Valuation* and *Price Trend*. Model 1 in Table 2 serves as our baseline analysis.[11] We observe that the coefficient of the *Valuation* is statistically significant (t-value of 9.48) and positive, while the *Price Trend* variable's coefficient is statistically

[10] Prior to winsorizing the data, the standardized valuation variable had a mean (across all funds) kurtosis of 13.78 and mean (across all funds) range of 13.18. After this process, the mean kurtosis for this variable is 2.26 with a mean range of 6.

[11] Due to the definitions of the *Long-Term Volatility* and *Long-Term Trend* variables, the regressions are executed over the time period December 31, 2008 through December 6, 2011.

significant (t-value of -8.42), negative, and slightly larger in magnitude (0.81 > 0.585).

<< Insert Table 2 >>

A negative trend term in the regression for returns has been suggested as evidence for mean reversion, implying that prices are returning to a more realistic value after an overreaction, for example, to a news announcement. Conversely, a positive coefficient for the trend suggests underreaction, i.e. prices are reacting too slowly to a particular development. Our approach helps to show that the positive and negative effects of trend are present simultaneously.

Indeed, linear regressions can be used to understand the nonlinear effects of an independent variable by including in the regression the square and cube (and possibly higher powers) of the variable.[12] In Model 2, we augment Model 1 with the second and third powers of the *Price Trend* and *Valuation* terms together with the cross terms, e.g. *Price Trend* multiplied by *Valuation*, *Price Trend* multiplied by the square of *Valuation*, etc. Indeed, the model is given by Equation 1,

$$R_{t+1} = \alpha_0 + \alpha_1 T_t + \alpha_2 T_t^2 + \alpha_3 T_t^3 + \alpha_4 V_t + \alpha_5 V_t^2 + \alpha_6 V_t^3 + \alpha_7 P_t V_t + \alpha_8 P_t^2 V_t + \alpha_9 P_t V_t^2. \quad (7)$$

The coefficients of the *Valuation*, *Price Trend*, *Price Trend* squared, *Price Trend* cubed, and the interaction of *Price Trend* with *Valuation* variables are statistically significant. A plot of the next day's *Return* versus the variables in Model 2 is included in Figure 1. This figure clearly shows the nonlinear relationship between the *Return* and the *Price Trend*.

<< Insert Figure 1 >>

To further examine this relationship one may take cross-sections of Figure 1 holding the *Valuation* variable fixed. For example, setting the *Valuation* to zero yields the following relationship

$$R_{t+1} = 0.902 T_t - 0.58 T_t^2 - 0.46 T_t^3, \quad (8)$$

[12] If the intrinsic relationship between two variables is nonlinear, for example of the form $z = x(x-1)(x+1)$, then the best straight line approximation to this relationship would be the line $z = 0$. If one is testing the hypothesis that z increases with increasing x when x is small; but for larger x, further increase in x results in smaller z, then a linear regression is inappropriate. Instead one needs to have x, x^2 and x^3 terms in the regression. Similarly, if there are two independent variables, x and y, whose higher order terms need to be considered, then one also needs to consider for consistency the cross terms xy, xy^2 etc. In particular, we would like to examine whether the return and trend exhibit this type of nonlinear relationship.

which is plotted (as the cubic curve) in Figure 2. For smaller *Price Trend* values the relationship is fairly linear. However, once a change in *Price Trend* is greater than 0.491 or less than -1.33 standard deviations, then the nonlinearity is quite apparent. Indeed, a positive *Price Trend* supports higher returns until this variable reaches 0.491 standard deviations. This seems to be the point of diminishing returns in that higher *Price Trend* values result in lower returns. Once the *Price Trend* is greater than 0.905 standard deviations, the following day's return turns negative.

<< Insert Figure 2 >>

A similar phenomenon occurs with negative *Price Trend* values; however, it is slightly skewed to the left. That is, increasingly negative *Price Trends* correspond to increasingly negative *Returns* until the *Price Trend* is equal to -1.33. As the *Price Trend* decreases below this value, returns increase. The return is zero for a *Price Trend* of -2.17 and is positive for *Price Trend* of less than -2.17. So, while a *Price Trend* of approximately 0.96 leads to a negative return, it takes a negative *Price Trend* of more than twice that value to produce a positive return. This might suggest that traders were less likely to buy on downtrends (i.e., seek bargains) than they were to sell (i.e., realize profits) on uptrends.

There is an inflection point at a *Price Trend* value of -0.42, which indicates the value (in standard deviations) at which the slope of the curve, i.e. the value at which the rate of change of return versus trend, is maximized. Indeed, the significance of the inflection point is that it represents the value at which the price trend has its greatest effect on the return. Note the intercept term is set to zero in Figure 2 (Figure 1). This term would (in this case) shift the plot "up" by approximately 1.57 units. While this would change the x-intercepts, it would not change the points at which the Price Trend's influence changes (i.e., the local maximum and minimum values of 0.491 and -1.33, respectively). Moreover, if all other variables are set to zero, then Figure 2 demonstrates that a fund with a *Price Trend* of 0.525 will always have a greater expected return the next day than a fund with a *Price Trend* value of 1.

A plot of both the nonlinear (Model 3) and linear (Model 1) relationships between tomorrow's return and the recent trend in price is included in Figure 2. Observe that if one were to fit a straight line to this curve, then it would have a negative slope. This imposition of linearity on a nonlinear effect explains why the coefficient of the *Price Trend* variable in Model 1 is negative. This might also help to explain why other studies have found the recent trend to be either not statistically significant or so small as to be negligible in practical terms.

The relationship between tomorrow's return and *Valuation* may also be analyzed in a similar manner. Assuming the *Price Trend* variable is zero, this relationship is given by

$$R_{t+1} = 0.540V_t + 0.020V_t^2 + 0.029V_t^3. \tag{9}$$

While the curve is not a straight line, the relationship is fairly linear, especially for small values of *Valuation*. Also, notice the curve is always increasing with no local maximum or minimum values. Thus, larger *Valuation* values correspond to higher expected returns, while smaller *Valuation* values correspond to lower future returns.

In Model 3, we augment Model 2 with the money supply, volatility, long-term trend, and resistance variables. Note that the *Valuation* and *Price Trend* variables have almost the same coefficients as in Model 2. Both the *Short* and *Long-Term Volatility* variables have positive coefficients, though only the *Short-Term Volatility* variable's coefficient is significant. Prior studies (see Section 3.B) have demonstrated that an increasing money supply has a positive effect on the return. While the *Money Supply's* positive coefficient is consistent with these findings, it is not statistically significant in any of the models. The *Long-Term Trend* variable has a negative coefficient suggesting mean reversion over the longer (annual) time scale; however, it is also not statistically significant.

As discussed above, market participants are aware of recent price history as a reflection of the anticipated strategies of other traders. In Model 3, we see that the *Resistance* variable's coefficient is negative with a t-value of -1.48. Note that 859 (out of 57,720) records satisfy the Resistance criteria.

Thus there is marginal support for the belief held by many practitioners that as prices approach a recent high price, there is additional selling that tends to reduce the return. The motivation for such selling is that a yearly or recent high price is "anchored" in the minds of investors who have subsequently seen a lower valuation of their asset. Thus, the investor has observed that others are selling at this price (whatever their reason) and is more likely to sell in anticipation of prices pulling back from this level once again. One explanation for this behavior is that traders seek to avoid the "regret" they would encounter upon seeing the stock retreat a second time from the price it attained. *Resistance* is a non-standardized, binary variable. Thus, although the coefficient of the *Resistance* variable is larger than those of the other variables, one must exercise caution in performing any type of comparison.[13]

B. Discussion of key results

The results of the regressions demonstrate that price trend is a significant effect. An interesting feature of this variable is the nonlinearity that it exhibits. When the price trend is sufficiently small that it occurs less than one standard deviation from the mean trend, there is a positive impact on the following day's return. As the trend increases to a level that is observed less frequently, it has a negative impact. We also see a great deal of asymmetry between the positive and negative trends. For negative trends, the change in the sign of the return does not occur until the downtrend is a two standard deviation event (i.e., occurring 2.5% of the time).

The issue of momentum, or trend, has been examined from a variety of perspectives. Market participants have long observed that trends tend to persist. This forms a key pillar of technical analysis that gives a qualitative description in forms that are often difficult to test (see Magee and Edwards, 1948). Among the early works, Caginalp and Ermentrout (1990) modeled asset prices with a trend term in addition to valuation influencing demand and supply. More recently, Jegadeesh and Titman (1993) and Carhart (1997) have examined the significance of momentum in characterizing asset returns. The

[13] The issue of standardized trading volume is discussed in the Appendix.

underlying causes for the existence of an uptrend could be many. A possible reason is that the market underreacts to changes in valuations. In this interpretation, the existence of a positive coefficient for trend can be regarded as support for underreaction.

Another possible reason for a trend is that a trader observes that prices are moving in a particular direction. This provides information regarding the beliefs/actions of other market participants to the trader, who adjusts his/her strategy accordingly. For example, a fund manager with a large cash position during a strong uptrend may be induced to increase the fund's stock allocation to avoid missing on the opportunity.

Conversely, a positive (negative) return following a negative (positive) change in the trend has been attributed to overreaction (see, for example, Duran and Caginalp, 2007, Madura and Richie, 2004, and Sturm, 2003). Thus, our study of ETFs provides a unified perspective into these phenomena. When prices rise gradually, traders tend to follow. When traders notice that the uptrend is larger than is usually observed, they anticipate that other traders will take profits. A basic game theoretical analysis suggests that selling should dominate at some point. This aspect of trading is similar to classical games such as Prisoner's Dilemma. While the traders with long positions would be better off if none of them sold, the result of iterated dominance is that selling occurs when there is a substantial profit for each of the traders in that situation.

Both short- and long-term volatilities have strong impacts on returns, and their coefficient estimates have magnitudes comparable to that of the *Price Trend*. The signs in both cases are positive. This means that given two ETFs, which are similar in other aspects (e.g., long-term trend, value etc.,) the one with higher volatility exhibits a higher return on average the next day. A higher ex-post return in the short-term following a high volatility period is essentially a different issue from the equilibrium concept of the high-risk/high expected-return paradigm. According to the standard model, increased risk accompanies a reduction (contemporaneously) in intrinsic value and therefore a diminished return. The residual price dynamics we investigate in this paper involves the dynamic issue of a return to equilibrium, and essentially is in a different domain from that of the equilibrium risk/return models.

5. Conclusion

A salient feature of the efficient market hypothesis (EMH) is that once one knows the fundamental value of a financial instrument – or the payout of an asset in a laboratory experiment – one has a complete description of the price dynamics, which is simply the fundamental value plus the random noise. In such a setting, the mathematical description of price dynamics does not require any empirical testing. Therefore, regressions that utilize a suitable valuation variable in addition to other hypothesized independent variables can determine the extent to which a financial market adheres to efficiency. Moreover, these regressions establish empirically the underlying forces in the dynamics of asset prices as they deviate from efficiency.

A closely related issue is that while classical economics focuses on equilibrium, or near equilibrium, there is a stream of news announcements that influences valuation and variables that may influence traders' strategies. The process by which market participants react to news and anticipate the actions of others leads to a price dynamics that requires new methodology. ETFs offer an opportunity to study this dynamics. As discussed above, there is a range of prices within the arbitrage bounds for which it is not economically attractive for the authorized participants to trade when an ETF market price deviates from its fundamental value. We use this approach to delve into the motivations and strategies involved in trading by using market prices of ETFs in conjunction with their net asset values.

While a common assumption is that ETFs precisely track the market, there is also the possibility that the trading in large ETFs moves the overall market in individual equities. By reducing the "noise" in pricing data using appropriately defined variables, we have provided strong empirical evidence for a number of factors in the dynamics of stock prices. These results can be examined in the context of particular theories on the strategies and motivations of traders. For example, Smith, Suchanek, and Williams (1988) observed that, since all participants in their experiments possessed the same information about valuations, any uncertainty involved the actions of others. In our empirical study, we are in a position to make a similar statement: Since the regressions compensate for the changes in valuations, the remaining (statistically significant) dynamics can be attributed to traders' anticipation of the actions of others. Thus, the data involving ETFs provide us with a laboratory on which theoretical ideas can be tested to refine economic theories of markets.

ETFs offer investors a convenient tool for adopting or hedging large equity positions very quickly. The huge market capitalizations and volumes of these instruments together with the possibility of large, rapid trades means that ETFs are a major factor in the dynamics of financial markets. Evidence in our regressions suggests that ETFs may be utilized in trend-based strategies. If value-based investors trade the stocks underlying the ETFs more, then this has implications for the stability (and instability) of markets. Modeling of asset price dynamics using differential equations has shown (DeSantis, Caginalp, and Swigon, 2012) that market stability is related to the strategies and motivations of the traders, as, in particular, trend-based investors comprise a larger fraction of the trading relative to the fundamentalists.

One of the strengths of our methodology is related to using standardized independent variables. For example, by stratifying our sample in time, we can track changes in investor motivations leading to an interesting research extension, where we would investigate the difference between the 1998-2000 and 2001-2003 periods (i.e., when the high/tech boom turned into a bust) to determine whether the cubic effect was more symmetric during the former period. In addition, our methodology can easily be used for trading a particular group of stocks (e.g., by industry) or individual stocks. This provides a unique insight into the relative importance of variables such as price trend, volatility and money supply. Our approach is not restricted to these important variables, and can be implemented to test and calibrate any hypothesis (behavioral or classical) that can be quantified.

Acknowledgement: We thank Professors Alok Bhargava, Nathaniel Wilcox and Steve Gjerstad for helpful comments and suggestions. In addition, we are grateful to three anonymous referees for their detailed analyses and comments. We also acknowledge data support from Bloomberg L.P.

Appendix: Analysis of daily trading volume

We have two objectives in this appendix. The first is to determine the role, if any, of the trading volume, as a proxy for liquidity of the sampled ETFs, in the daily price dynamics. We will show that there is none. The second is to understand the role of the trading volume in the magnitude of the relative spread (difference between price and NAV) and the absolute value of the return. Classical finance would suggest that both of these qualities would be inversely related to trading volume, with the idea that inefficiency born out of illiquidity.

Let $V_{f,t}$ be the trading volume for fund f on day t. We standardize this variable by day by subtracting the mean trading volume of all funds on day t and then dividing by the standard deviation. Using the methodology described in Section 3 and R_{t+1} as the dependent variable, we obtain (Table 2, Model 4) an insignificant coefficient of 0.000013 (t-value equals 0.06) for this daily volume variable. Augmenting Model 3 with volume (Table 2, Model 5) also yields an insignificant coefficient of 0.000078 (t-value equals 0.38), essentially rendering very little change on the coefficient estimates of Model 3. Hence, these results suggest the return is unaffected by the daily trading volume of a fund.

In terms of the second objective[14], we consider whether funds with higher daily trading volumes have larger return magnitudes (positive or negative). We regress the absolute value of the following day's return against the daily (standardized) trading volume variable, and find this variable has statistically significant and positive coefficient of 0.001721 (14.47). This implies funds with higher volume have larger return magnitudes.

We next consider whether liquidity, as measured by trading volume, plays a role in determining the valuation spread. That is, one might conjecture that ETFs with higher daily trading volumes have smaller valuation spreads. We test this assertion with our methodology. We use the absolute value of the daily valuation spread as the dependent variable, and the standardized daily trading volume variable

[14] The results for the following two regressions are discussed in the text, but not included in Table 2.

as the independent variable. This variable has a statistically significant coefficient of 0.000085 (5.21). This result suggests that ETFs with higher daily trading volume actually have a larger (in magnitude) daily valuation spread than lower volume ETFs. This is contrary to the expectation one might have based on the notion that inefficiency as measure by deviation from valuation is due to lack of trading volume. Note that we have chosen only high trading volume ETFs to be included in our sample. Hence, one cannot draw any conclusions regarding ETFs with lower trading volume.

References

Anderson, S.C. and J.A. Born, 2002, Closed-end fund pricing: Theories and evidence, Kluwer.

Asem, E. and G. Tian, 2010, Market dynamics and momentum profits. Journal of Financial and Quantitative Analysis 45, 1549-1562.

Barberis, N.; A. Shleifer; and R. Vishny, 1998, A model of investor sentiment. Journal of Financial Economics 49, 307-343.

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

Borkovec, M. and V. Serbin, Create or Buy: A comparative analysis of liquidity and transaction costs for selected U.S. ETFs. Journal of Portfolio Management (posted online July 2013).

Caginalp, G. and D. Balenovich, 1999, Asset flow and momentum: Deterministic and stochastic equations. Philosophical Transactions of the Royal Society 357, 2119-2113.

Caginalp, G. and G. Constantine, 1995, Statistical inference and modeling of momentum in stock prices. Applied Mathematical Finance 2, 225-242.

Caginalp, G. and M. DeSantis, 2011, Nonlinearity in the dynamics of financial markets. Nonlinear Analysis: Real World Applications 12, 1140-1151.

Caginalp, G. and G.B. Ermentrout, 1990, A kinetic thermodynamics approach to the psychology of fluctuations in financial markets. Applied Math Letters 4, 17-19.

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

Daniel, K.D.; D. Hirshleifer; and A. Subrahmanyam, 1998, Investor psychology and security market under- and overreaction. Journal of Finance 53, 1839-1885.

Demos, T., 2012, April 24, ‘Real’ investors eclipsed by fast trading. Financial Times, Retrieved from http://www.ft.com.

DeSantis, M.; G. Caginalp; and D. Swigon, 2012, Nonlinear dynamics and stability in a multi-group asset flow model. SIAM Journal on Applied Dynamical Systems 11, 1114:1148.

Duran, A. and G. Caginalp, 2007, Overreaction diamonds: Precursors and aftershocks for significant price changes. Quantitative Finance 7, 321-342.

Edwards, R.D. and J. Magee, 1948, Technical Analysis of Stock Trends, New York Institute of Finance.

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

Gelman, A., 2008, Scaling regression inputs by dividing by two standard deviations. Statistics in Medicine 27, 2865-2873.

George, T.J. and C.Y. Hwang, 2004, The 52-week high and momentum investing. Journal of Finance 59, 2145-2176.

Hong, H. and J.C. Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets. Journal of Finance 54, 2143-2184.

Jegadeesh, N. and S. Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency. Journal of Finance 48, 65-91.

Krause, T. and Y. Tse, 2013, Volatility and return spillovers in Canadian and U.S. industry ETFs. International Review of Economics and Finance 25, 244-259.

Madhavan, A., 2012, Exchange-traded funds, market structure, and the flash crash. Financial Analysts Journal 68, 20-35.

Madura, J. and N. Richie, 2004, Overreaction of exchange-traded funds during the bubble of 1998-2002. Journal of Behavioral Finance 5, 91-104.

Moskowitz, T. and M. Grinblatt, 1999, Do industries explain momentum? Journal of Finance 54, 1249-1290.

Poterba, J.M. and L.H. Summers, 1988, Mean reversion in stock prices: Evidence and Implications. Journal of Financial Economics 22, 27-59.

Smith, V.L.; G.L. Suchanek; and A.W. Williams, 1988, Bubbles, crashes and endogenous expectations in experimental spot asset markets. Econometrica 56, 1119-1151.

Sturm, R.R., 2003, Investor confidence and returns following large one-day price changes. Journal of Behavioral Finance 4, 201-216.

Sturm, R.R., 2008, The 52-week high strategy: Momentum and overreaction in large firm stocks. The Journal of Investing 17, 55-67.

Sturm, R.R., 2013, Market efficiency and technical analysis; can they coexist? Research in Applied Economics 5, 1-16.

Theil, H., 1961, Economic forecasts and policy, Second Edition, North-Holland, Amsterdam.

Tversky, A. and D. Kahneman, 1974, Judgment under uncertainty: Heuristics and biases. Science 185, 1109-1186.

Tables

Table 1. Sample Descriptive Statistics (Data Range: 01/02/2008-12/06/2011)
This table provides the descriptive characteristics of the ETF sample. The Min/Max of the Minimum/Maximum Valuation Spread corresponds to the lowest (highest) valuation spread observed for any ETF during the sample period. The Mean Absolute Value Valuation Spread considers only the magnitude of the spread – not the sign.

Date Range: 1/2/2008 - 12/7/2011	Mean	St. Dev.	Min	First Quartile	Median	Third Quartile	Max
Mean Market Capitalization[15] (Millions)	2,815	3,395	219	724	1,755	3,195	20,899
Mean Daily Volume (# of shares in Millions)	5.941	17.402	0.065	0.221	0.924	3.636	129.675
Mean Valuation Spread (percent)	0.001	0.023	-0.053	-0.009	0.002	0.010	0.119
Minimum Valuation Spread (percent)	-	-	-13.24	-	-	-	-
Maximum Valuation Spread (percent)	-	-	-	-	-	-	7.38
Mean Absolute Valuation Spread (percent)	0.117	0.06	0.0465	0.078	0.093	0.136	0.386
Number of ETFs in the sample	78						

[15] The Bloomberg system was missing market capitalization data for VAW for the timeframe 1/3/2008-12/22/2008 (inclusive) as well as 12/24/2008. As such, VAW is not included in the Mean Market Capitalization data.

Table 2. Regression Results
The two-way fixed effects methodology was applied to a balanced panel of 78 funds each with 740 daily observations. Notice the coefficient of the *Price Trend* variable is (1) statistically significant (Models 1-3) and (2) positive when the higher order *Price Trend* and *Valuation* variables are included (Models 2 and 3). Models 4 and 5 are discussed in the Appendix.

	Model 1	Model 2	Model 3	Model 4	Model 5
(Intercept)	2.037	1.571	0.402	1.173	0.392
	1.41	(1.09)	(0.02)	(0.81)	(0.02)
Valuation	0.585*	0.54*	0.535*		0.535*
	(9.48)	(5.38)	(5.33)		(5.32)
Price Trend	-0.81*	0.902*	0.909*		0.909*
	(-8.42)	(6.40)	(6.38)		(6.38)
Money Supply			1.399		1.345
			(0.03)		(0.03)
ST Volatility			1.176*		1.171*
			(6.48)		(6.44)
LT Volatility			0.8		0.803
			(1.56)		(1.57)
LT Trend			-0.02		-0.02
			(-0.20)		(-0.20)
$(Price\ Trend)^2$		-0.58*	-0.68*		-0.68*
		(-7.74)	(-8.83)		(-8.83)
$(Price\ Trend)^3$		-0.46*	-0.46*		-0.46*
		(-16.41)	(-16.48)		(-16.48)
$Valuation^2$		0.02	0.014		0.013
		(0.52)	(0.36)		(0.35)
$Valuation^3$		0.029	0.03		0.03
		(1.57)	(1.63)		(1.63)
$Price\ Trend \times Valuation$		-0.019*	-0.2*		-0.2*
		(-3.74)	(-3.83)		(-3.83)
$(Price\ Trend)^2 \times Valuation$		-0.04	-0.04		-0.04
		(-1.47)	(-1.53)		(-1.53)
$Price\ Trend \times Valuation^2$		-0.04	-0.04		-0.04
		(-1.42)	(-1.49)		(-1.49)
Resistance			-0.72		-0.72
			(-1.48)		(-1.48)
Volume				0.013	0.078
				(0.06)	(0.38)

Table 2. (continued)

	Model 1	Model 2	Model 3	Model 4	Model 5
R-Square	0.5928	0.5951	0.5955	0.5916	0.5955
No. Observations	57,720	57,720	57,720	57,720	57,720
No. Groups/Funds	78	78	78	78	78
No. Days (per Fund)	740	740	740	740	740
F Test for No Fixed Effects	101.25*	100.46*	100.20*	101.01*	100.20*

Note:

a. t-values are reported in parentheses.
b. * indicates P < 0.01.
c. Coefficient values have been multiplied by 1,000 for exposition.
d. The reported R-Square value corresponds to the R-Square measure developed by Theil (1961).

Figures

Figure 1. Plot of tomorrow's return versus the recent trend in price and valuation.
The results from Model 2 are utilized to illustrate the nonlinear relationship between tomorrow's return and the recent trend in price.

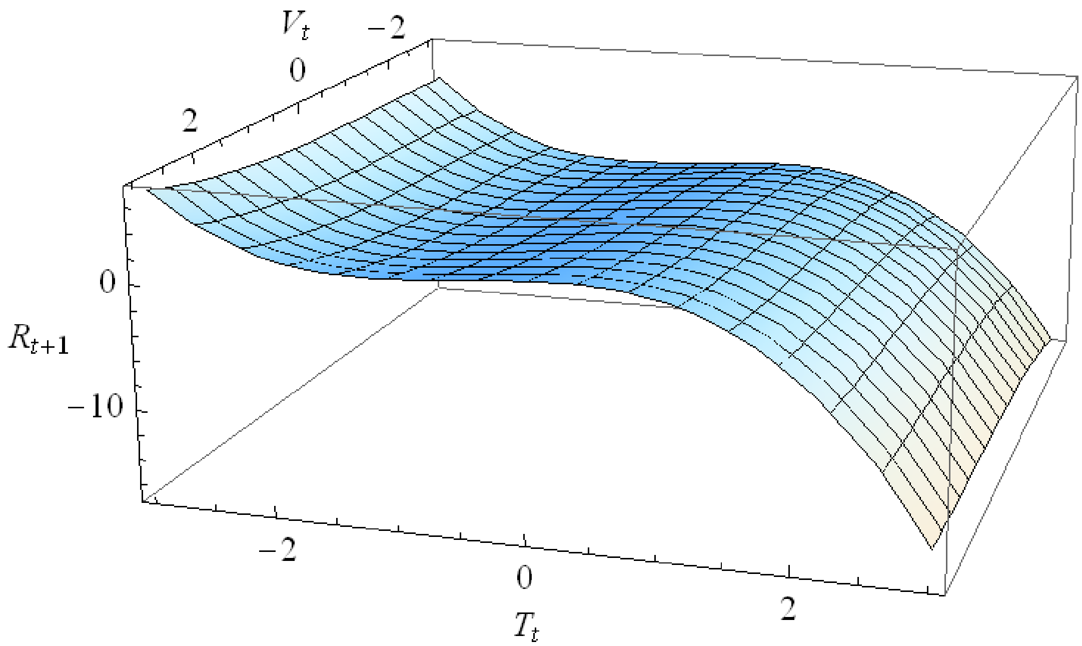

Figure 2. Plots of tomorrow's return, R_{t+1}, versus the recent trend in price.
The nonlinear curve, produced via Model 2, demonstrates the nonlinear nature of the relationship between these two factors. The straight line is the product of Model 1, where the *Price Trend* variable has a negative coefficient. This provides a graphical representation of what can happen when a linearity constraint is imposed upon a nonlinear relationship. We assume that the *Valuation* variable is set to zero and ignore the intercept term for clarity.