Lost in the Rising Tide: ETF Flows and Valuation *

Yuan Zou

Columbia Business School

yzou19@gsb.columbia.edu

January 22, 2019

Abstract

The last decade has witnessed a dramatic growth in passive investing via exchange-traded funds (ETFs). To the extent that the demand for stocks via ETF flows is not related to firm-specific fundamental values, large ETF flows may push the price of the underlying stocks away from their fundamentals-based value. In this study I provide evidence consistent with this conjecture. In particular, I first document a positive association between ETF flows and the price-to-fundamentals relation of underlying stocks. Then, by using BlackRock's expansion into the ETF business as an exogenous shock, I provide evidence that the association is likely to be causal rather than reflect some form of endogeneity (i.e., ETFs selecting certain stocks). Also, I find that high-flow firms subsequently underperform low-flow firms in operating and stock performance, consistent with the misvaluation being caused by non-fundamental demand shocks. Cross-sectional tests suggest that the ETF-related misvaluation is stronger for stocks with: a less competitive equity market (i.e., with prices more sensitive to demand shocks), lower ownership by active investors, and more costly arbitrage constraints. Finally, I find that high-flow firms exhibit behavior typically associated with perceived overvaluation (e.g., more secondary equity offerings).

---

* For their guidance and support, I especially thank the members of my dissertation committee: Stephen Penman (chair), Fabrizio Ferri, Lawrence Glosten, Jonathan Glover, and Shiva Rajgopal. I thank Sudipta Basu, Thomas Bourveau, Robert Bushman, Matthias Breuer, Wei Cai, Kent Daniel, Richard Frankel, Pingyang Gao, Trevor Harris, Rachel Hayes, Alon Kalay, Jinhwan Kim, David Kwon, Alvis Lo, Doron Nissim, Ethan Rouen, Robert Stoumbos, Hao Xue, Wenhao Yang, Frank Yu, Ronghuo Zheng, Christina Zhu, workshop participants at Columbia, the CMU Emerging Scholar Session, and the SEC Doctoral Symposium for helpful comments. I also thank Bruiser Liu from PhaseCapital, Alex Schiller from the SEC, and multiple investment professionals for providing institutional insights. This project has also benefited from discussions with my fellow students Serene Huang, Yiran Kang, Venkat Peddireddy, and Kunjue Wang. I thank Martijn Cremers for making active shares data available. I am also thankful for the financial support from the Columbia Accounting Alumni Award, the W. Edwards Deming Center, and the Duncan Merriwether Doctoral Fellowship. All errors are my own.

1. Introduction

The efficient market hypothesis and modern portfolio theory have laid the foundations for index investing and implied that investors should buy well-diversified portfolios with low fees. However, these theories are silent about the impacts of index investing on pricing. This study fills this gap by examining whether exchange-traded fund (ETF) flows induce stock misvaluation by pushing the price of the underlying stocks away from their fundamentals-based value.

The creation of ETFs represents one of the most important financial innovations in recent years. ETFs possess many of the characteristics of what Rubinstein (1989) calls an "ideal market basket vehicle." For example, ETFs provide investors with continuous access to diversified portfolios and with tax benefits.[1] As a result, they have exhibited extraordinary growth over the past decade. From 2007 to 2017, the assets under management by ETFs grew from $608 billion to $3.4 trillion (Investment Company Institute 2018) and are expected to reach $5 trillion by 2020 (PwC 2015). Currently, roughly 30% of U.S. equity trading volume is attributable to ETFs.

The rapid growth in ETFs has raised concerns among policy-makers and practitioners about their pricing impact (see Appendix 1). Textbook theories suggest that in complete and frictionless markets, basket securities such as ETFs are redundant assets with no impact on the prices of the underlying securities. However, ETF flows are mechanically invested in the underlying stocks in proportion to their weights in the index that the ETF replicates.[2] This implies that demand for stocks via ETF flows is not related to firm-specific fundamental values. If individual stocks have downward-sloping demand curves (e.g., Shleifer 1986; Kaul et al. 2000; Wurgler and Zhuravskaya 2002), then large ETF flows may induce non-fundamental demand

[1] ETF shares are created or redeemed "in kind" – that is, to exchange ETF shares for a basket of securities rather than cash. This allows the ETF to avoid selling securities to raise cash to meet redemptions and prevents capital gain distributions. Therefore, ETFs are more tax-efficient than traditional mutual funds.

[2] In my sample 60% of ETFs use market-capitalization weights. Others use price-weighted or equal-weighted schemes.

shocks, causing prices to decouple from fundamental values.[3] Thus, whether and how ETF flows affect the price relative to the fundamental value of the underlying stocks is an empirical question.

To address this question, I start by examining the association between ETF flows and two measures of the price-to-fundamentals relation for the underlying stocks. The first one is the value-to-price ratio (V/P) ratio, which is the ratio of "intrinsic value" (*V*) to market price (*P*) (e.g., Lee and Swaminathan 1999; Frankel and Lee 1998; Bradshaw 2004). *V* is a forward-looking measure of the fundamental value derived from the residual income valuation model. Intuitively, a *lower* V/P ratio indicates that the market price is high relative to the fundamental value. The second one is the "relative growth" ratio, which I define as the ratio between the growth rate implied by the fundamentals (operationalized as the growth rate in residual income implied by analysts' consensus earnings forecasts) and the growth rate implied by the stock price (i.e., the growth rate that reconciles the residual income model to the market price, hereinafter the "price-implied growth rate" (Penman 2011)). Intuitively, a *lower* relative growth ratio suggests that the growth rate expected by the market is high compared to the growth rate based on the firm's fundamentals (as reflected in analysts' forecasts).[4] Thus, for both measures a lower value suggests high market price or market growth expectation relative to the fundamentals-based value or growth.

Using ETF flow data from 2002 to 2016 from Bloomberg, I find negative associations between ETF flows (during a given year) and both V/P and relative growth ratios (measured at the end of the same year), after controlling for firm-level determinants of price-to-fundamental ratios (e.g., operating performance, risk measures) and firm fixed effects. Specifically, a one-standard-

[3] As explained in more details in Section 2.1, if the ETF price deviates from the NAV of the portfolio holdings because of a demand shock, authorized participants and other arbitrageurs trade the underlying securities in the same direction as the initial shock to the ETF. Therefore, the shocks that occur in the ETF market can be transmitted to the underlying stocks.

[4] The value-to-price ratio is perhaps a more direct measure of the price-to-fundamentals relations, but it requires some assumptions about the terminal value (see Section 3.3 for details). Hence, I also examine the relative growth ratio, which does not require those assumptions.

deviation increase in ETF flows is associated with an 8.2%-standard-deviation decrease in V/P and a 5.0%-standard-deviation decrease in relative growth, equivalent to shifting the median stock from the 50th percentile to the 42nd and 45th percentiles of the ratios' distributions, respectively.

This association per se is not evidence of a causal impact of ETF flows on the price-to-fundamentals relation and may be the result of various forms of endogeneity. For example, it may reflect a selection effect. If ETF fund managers create ETFs that tend to cover glamour industries or include stocks that have recently performed well, there may be a spurious correlation between ETF flows and the price-to-fundamentals relation. However, my findings are robust to adding industry-year fixed effects and to excluding ETF flows during the first year after an ETF is created. More importantly, to address endogeneity concerns, I exploit BlackRock's expansion into the ETF business via its acquisition of BGI (Barclays Global Investors) and its iShares ETF unit in 2009. The acquisition was driven by BlackRock's desire to expand into the ETF business and Barclays's desire to raise funds to avoid a possible future bailout by the UK government (Massa et al. 2016; Azar et al. 2018). As a result of the acquisition, iShares ETFs experienced a significant increase in inflows (relative to ETFs not belonging to iShares), due to BlackRock's ability to utilize its extensive resources (e.g., salesforce, branding and scale benefits, and strong relationships within the distribution channels). This setting generates exogeneous variation in ETF flows because the acquisition was not driven by the fundamental characteristics of the stocks with a larger fraction of shares held by iShares (the treated sample) but did increase the ETF flows for these stocks, compared with stocks with a smaller fraction of shares held by iShares (the control sample).

Using a difference-in-differences (DiD) research design, I find a statistically significant decrease in the V/P and relative growth ratios for treated stocks relative to control stocks following the acquisition.[5] To the extent that the iShares acquisition was indeed an exogenous shock (exogenous relative to the portfolio stocks' fundamental characteristics), this analysis supports a causal explanation for the association between ETF flows and the price-to-fundamentals relation documented in my main tests.

A causal effect of ETF flows on the price-to-fundamentals relation is consistent with non-fundamental demand shocks pushing prices away from fundamental values, but may also reflect other mechanisms. In particular, it is possible that ETF flows impound fundamental information better or more promptly than the inputs in the residual income model do. For example, firms with high ETF flows (hereinafter high-flow firms) might have superior expected future performance, which earnings forecasts in the residual income model fail to fully incorporate. In other words, the estimated *V* for high-flow firms may understate the true *V* and the estimated implied growth may overstate the true market growth expectation, and thus artificially induce the documented low V/P and relative growth ratios. However, I find that ETF flows are *negatively* associated with subsequent firm operating performance, sales growth, and stock returns. In particular, a one-standard-deviation increase in ETF flows is associated with a 3.7% decrease in subsequent annual returns. The negative association provides further support for a non-fundamentals-based explanation, since non-fundamental demand shocks should subsequently revert. Along the same lines, high-flow firms may be associated with lower risk, which may not be properly captured by the discount rate in the residual income model (thus, again, the estimated *V* for high-flow firms may understate the true *V* and the estimated implied growth for high-flow firms may overstate the true market growth expectation and thus artificially induce the documented low V/P and relative

[5] In a series of placebo tests, I find no differential changes in V/P and relative growth between treatment and control firms from 2005 to 2008, providing support for the parallel trend assumption underlying the DiD research design.

growth ratios).[6] However, I find no difference in the cost of capital calculated from the Fama-French (2015) five-factor model of high-flow vs. low-flow firms. Also, I control for the Fama-French (2015) five-factor loadings in all the regressions.

Overall, the above analyses suggest that (i) the positive association between ETF flows and the price-to-fundamentals relation is consistent with a causality interpretation (rather than being driven by a selection effect or other forms of endogeneity); (ii) the mechanism underlying this causal effect is a non-fundamental-driven demand shock.

Next, I conduct a series of cross-sectional tests based on three predictions. First, if a stock has a perfectly competitive equity market, then investors are price-takers, and the demand curve for the stock should be horizontal with stock prices immune to demand shocks (Shleifer 1986). Thus, I expect stronger ETF-related misvaluation for stocks with less competitive equity markets, as proxied by the number of shareholders (Armstrong et al. 2010). Second, I expect a stronger ETF-related misvaluation in stocks with lower ownership by sophisticated "active" investors (identified as in Cremers and Petajusto (2009)), i.e., investors that could eliminate this misvaluation. Third, I predict stronger ETF-related misvaluation in stocks without close substitutes (proxied by stocks with higher idiosyncratic volatility, Mashruwala et al. (2006)) and stocks with higher short-selling costs (proxied by stocks with fewer shares held by lendable ETFs), since in these cases arbitrage is more costly and thus the ETF-related misvaluation should be higher. My cross-sectional analyses find support for all three predictions.

Finally, I examine whether high-flow firms exhibit the behavior typically associated with perceived overvaluation, i.e., greater equity issuance, less repurchases, and more insider sales

[6] This may happen under a number of scenarios. For example, ETF inflows may increase investors' attention, which could result in lower cost of capital (Merton 1987). However, I find a negative association between ETF flows and investors' conference call participation (a proxy for investors' attention) (see Section 5.2 for details).

(Baker and Wurgler 2002; Loughran and Ritter 1995; Khan et al. 2012). Indeed, I find that high-flow firms are more likely to issue secondary equity offerings (SEOs), tend to repurchase fewer shares, and have greater insider share sales. I also examine disclosure policies and find that high-flow firms reduce the amount of earnings forecasts, while low-flow firms increase the amount of positive earnings forecasts. One interpretation is that high-flow firms remain silent in an attempt to maintain the optimistic valuation, while low-flow firms increase the positive voluntary disclosure activity to correct the pessimistic valuation. Overall, these analyses indirectly confirm that my price-to-fundamentals measures do indeed capture stock overvaluation. They also suggest that ETF-driven misvaluation may have real effects on firms' policies.

This study contributes to a nascent line of literature on exchange-traded products (Madhavan 2016; Lettau and Madhavan 2018), on which the Securities and Exchange Commission (SEC) has called for more research (Piwowar 2017).[7] Prior studies have examined the effects of ETF ownership on co-movement of the underlying securities (Da and Shive 2015), the volatility of the underlying securities (Ben-David et al. 2017a), and the extent to which prices incorporate earnings information (Israeli et al. 2017; Glosten et al. 2018). Other studies focus on the role ETFs play in transferring industry information impounded in earnings news (Bhojraj et al. 2018). I contribute to this research by documenting another important consequence from a valuation perspective: an increase in ETF flows might push the stock price away from the fundamentals-based value. The results highlight a tension between the benefit of "completing" markets and the cost of a loss of fundamental information in stock prices due to the financial innovation. Such a

[7] A speech by the SEC commissioner Michael Piwowar in September 2017 points out that the effects of exchange traded products is still "a nascent field, and one in which we need more discussion and discovery": https://www.sec.gov/news/speech/speech-piwowar-2017-09-08.

loss of fundamental information might undermine efficient resource allocation and reduce managerial learning from prices (Baker et al. 2003; Chen et al. 2006).[8]

Another important implication of my findings is that ETF investors who are passive in nature may unintentionally end up holding overvalued stocks. In turn, this also suggests that passive investors might create a market for active investors *(*Johnson 2016), who should thus continue to focus on fundamentals and "buy value rather than prospects or popularity" (Graham 1963), in contrast to the recent shift of active management toward more index-like investing (see the AFA presidential address by Stambaugh (2014)). Relatedly, the SEC might consider further relaxing short-sale constraints to mitigate the overvaluation. Some ETFs lend out shares which could reduce the short-sale cost. Thus, the SEC might consider encouraging ETFs to lend out more shares.

Lastly, my paper contributes to the understanding of firms' strategic responses to perceived ETF-related overvaluation. In a similar vein, a recent literature documents that passive investors can affect firm policies and governance through their voting power (e.g., Appel et al. 2016). My paper suggests that ETF investors might have real effects on firms' policies indirectly through their unintended changes in stock prices via the flows into ETFs.

2. Institutional Details and Literature Review

2.1 Institutional Details

ETFs provide investors with the exposure to the broad market, sectors or geographical regions, or specific rule-based investment strategies. ETFs are passively managed in that they leave the selection and the weighting of stocks in the portfolios to the indexes.[9] The closet substitute to

[8] Although premature, one may consider a situation in which during market downturns, appreciation that was driven by the ETF buying is likely to eventually turn out to be rotational. This might pose threats to the financial stability (Anadu et al. 2018).
[9] 2% of ETFs are actively-managed. I exclude them from my sample.

ETFs are open-ended index mutual funds. Unlike index mutual funds, which can only be bought or sold at the end of the day at net asset value (NAV), ETFs are traded throughout the day at market-determined prices.

In addition to continuous access to diversified portfolios, ETFs offer low-fees and tax advantages relative to mutual funds.[10] Many ETFs lend out shares to short-sellers. They also provide investors with access to stocks that were previously difficult to trade.[11] In particular, the ability to exit a position at any time, rather than having to wait until the end of the day, makes the risk management of a bet much easier. These benefits have attracted an increasing demand from both institutional and individual investors.

Figure 1 presents the ownership by ETFs, index mutual funds, and non-index mutual funds for common stocks each year from 2002 to 2017.[12] The average fraction of a stock's shares held by ETFs has risen from 1.45% in 2002 to 6.81% in 2017. The growth of ETFs appears more drastic when compared to non-index mutual funds and index mutual funds. Index fund average stock ownership was more stable in the sample. Non-index mutual fund average stock ownership, although higher than ETFs, has experienced a decline starting from 2007.

The increasing demand for ETF shares can migrate to the underlying stocks because ETFs and the basket securities are tied by arbitrage both on the primary and secondary markets. On the primary market, authorized participants (APs), who are large broker-dealers, create or redeem ETF

---

[10] On average, the ETF structure enables lower fees than traditional mutual funds, as mutual funds may charge fees (such as 12b-1 fees) that ETFs do not charge.

[11] For example, a small-cap ETF, VB, is based on small stocks. However, VB holds close to $21.6 billion net assets as of December 2017 and trades at low costs.

[12] I identify a fund as an index fund if its fund name includes a string that identifies it as an index fund or if the CRSP Mutual Fund Database classifies the fund as an index fund. Specifically, I use the same strings as in Appel et al. (2016) to identify index funds. These strings include: *Index, Idx, Indx, Ind_(where_ indicates a space), Russell, S & P, S and P, SandP, SP, DOW, Dow, DJ, MSCI, Bloomberg, KBW, NASDAQ, NYSE, STOXX, FTSE, Wilshire, Morningstar, 100, 400, 500, 600, 900, 1000, 1500, 2000,* and *5000.*

shares to ensure that intraday prices of the ETF approximate the NAV of the underlying assets. The unique creation/redemption mechanism ensures that ETF shares outstanding expand or contract based on investors' excess demand. Figure 2 uses the creation of ETF shares to illustrate the mechanism. With inflows into an ETF (step 1), the AP short sells the ETF shares to meet investors' demand and buys the underlying securities to hedge the short position (step 2), putting upward pressure on the underlying securities. At the end of the day, the AP transfers the basket of the securities to the ETF sponsor (step 3), and the ETF sponsor creates new ETF shares to cover the AP's short position (step 4). A similar process happens if the AP redeems ETF shares with outflows from the ETF.

On the secondary market, with an increase in investors' demand in ETF shares, arbitrageurs like hedge funds can directly short the ETF and long the underlying securities. Moreover, arbitrageurs can impound the demand shock indirectly. For example, hedge funds can short sell an overpriced stock and hedge the industry risk by going long in the corresponding sector ETFs. This arbitrage activity can put upward pressure on the ETF price, which might be transferred to the underlying securities.

2.2 Literature Review

There is a growing recent literature on ETFs (Lettau and Madhavan 2018; Ben-David et al. 2017b). On the one hand, several studies document negative effects of ETFs. They document a negative association between ETF ownership and underlying stocks' liquidity and a positive association between ETF ownership and return co-movement (Hamm 2014; Da and Shive 2016). Ben-David et al. (2017a) find that ETF ownership increases intraday and daily stock volatility because of the transmission of liquidity shocks. Brown et al. (2017) document that ETF arbitrage activity negatively predicts subsequent monthly returns at the ETF level. However, all of these studies focus on shorter horizons than the annual horizon I use in my study. In the long-run, one would expect short-term mispricing to be arbitraged away.[13]

On the other hand, several early market microstructure studies highlight positive effects of ETFs. Theoretical models in Admati and Pflederer (1988) and Foster and Viswanathan (1990) suggest that the liquidity trading in the basket can induce informed traders to trade on private information. Consistent with these models, Hasbrouck (2003) documents evidence that ETFs contribute to intraday price discovery during March 2000 to May 2000. Yu (2005) finds that a substantial amount of information incorporated into stock prices originates in the ETF market from July 2002 to September 2002. Compared with these studies, I use not only the longer annual horizon, but also the data during the most recent period. This is important given the increasing popularity of ETFs since 2000.

Two recent papers using a relatively long-time horizon and recent sample periods are Glosten et al. (2018) and Israeli et al. (2017). Taken together, they find that an increase in ETF ownership is associated with an increase in the extent to which stock prices incorporate concurrent quarterly systematic earnings information, but a decrease in the extent to which stock prices incorporate future annual earnings information. My study complements these studies by focusing on valuation using a broader set of fundamentals (not only earnings) for the underlying stocks.[14]

This paper is also related to the literature on the effects of fund flows on stock prices (e.g., mutual fund flows, Frazzini and Lamont 2008). While the finding of a price impact is not new, I argue that the passive and trading nature of ETFs warrant separate attention. First, while non-index

[13] One anonymous investment professional states that his prior belief is that the non-fundamental demand shock might exist in the short run, but in the long-term, any mispricing will be arbitraged away by active investors.

[14] It is important to note that the evidence in the three papers is not necessarily mutually exclusive, as Ben-David et al. (2017b) argue: “It is possible that prices more quickly reflect certain pieces of information, and, at the same time, also are more impacted by liquidity shocks.”

mutual fund managers have control over the price impact or their trades, ETF managers have almost no discretion in terms of the timing and composition of their trades. Second, ETFs allow investors to access the market continuously with high liquidity compared with index mutual funds.[15] These benefits likely enable ETFs to attract more non-fundamental demand than mutual funds. In Section 7.2, I show that ETFs have a much larger impact on the price-to-fundamentals relation than index mutual funds and non-index mutual funds.[16]

3. Data, Sample, and Variable Construction

3.1 Data

I follow Brown et al. (2017) and draw information from the Center for Research in Security Prices (CRSP) and ETF.com to identify ETFs. Specifically, I download ETF tickers from ETF.com and merge them with the CRSP database. From Bloomberg, I obtain daily flows for each ETF from 2002 to 2016, which I then aggregate into the monthly level.[17] I focus on ETFs that are listed on U.S. exchanges and whose baskets contain U.S. stocks. I include only "plain vanilla" ETFs that engage in full replication for the underlying indexes. I exclude leveraged and inverse-leveraged ETFs. I merge the ETF data with quarterly ETF holdings data from Thomson-Reuters Mutual Fund holding database (S12) by using the MFLINKS tables. My sample includes 445 ETFs in total, which is comparable to recent studies (e.g., Ben-David et al. 2017a).

Table 1 Panel A presents descriptive statistics for the ETFs used in the paper. On average, an ETF attracts $32.843 million inflows per month. 72% of the ETFs lend out shares (lendable

---

[15] For example, nowadays a lot of insurance companies invest in ETFs due to their high liquidity.

[16] This paper is also indirectly related to the literature on index inclusion (e.g., Shleifer 1986; Wurgler 2011). The trigger for the effect I measure is the extent to which investors buy the entire baskets via ETFs, as opposed to the index inclusion. Index membership matters only in defining the stocks that ETFs affect. Also, index inclusion is usually associated with increased firm disclosures, analyst following, and news coverage which might increase investors' attention (e.g., Boone and White 2015). However, as Section 5.2 and Section 6.2 suggest, ETF flows appear to be negatively associated with investors' attention and firm disclosure activity.

[17] The daily ETF-level flow is defined as the daily change in ETF shares outstanding multiplied by fund NAV.

ETFs). Around 60% of the ETFs use market-capitalization weights and other ETFs use either equal-weighted schemes or price-weighted schemes.

The stock sample includes firms listed on the NYSE, AMEX, or NASDAQ with CRSP share codes 10 or 11. Stock return and accounting data are from the intersection of the CRSP and Compustat datasets. Analyst earnings forecast data are from I/B/E/S. Fama-French factors are from Ken French's website. I exclude stocks with prices less than $1 to mitigate the market microstructure noise. I obtain SEO data from the SDC Platinum database and insider trading data from the Thomson Financial Insider database. Management earnings forecasts data are from I/B/E/S Guidance database. To alleviate the effects of outliers, I winsorize all continuous variables at the 1 and 99 percent levels. My final sample includes 29,590 firm-years with all the financials, analyst data, and ETF flows data available.

3.2 ETF Flow

For each stock *i* in month *m*, I calculate the stock-level ETF flows as the weighted average of the flows for all ETFs holding the stock, where the weight is the stock *i*'s portfolio weight in each ETF *j*. I then aggregate monthly stock-level ETF flows at the annual level, where the flows are accumulated for 12 months from 3 months after the last fiscal year-end to 3 months after the current fiscal year-end. Next, I scale the annual stock-level ETF flows by the annual dollar trading volume.

Specifically, the ETF flows for each stock *i* during each year *t*, $ETF\_Flow_{i,t}$, is calculated as follows:

$$ETF\ Flow_{i,t} = \frac{\sum_m^M \sum_{j=1}^J \$ETF\ Flow_{j,m,t} \times Weight_{i,j,m,t}}{\sum_m^M \$Volume_{i,m,t}}, \tag{1}$$

where $ETF\ Flow_{j,m,t}$ is the dollar amount of flows in ETF *j* in month *m* of year *t* and $Weight_{i,j,m,t}$ is the percentage of stock *i* in the portfolio of ETF *j* held in month *m* of year *t* as reported by the most recent available fund portfolio weight data in Thomson Reuters. $Volume_{i,m,t}$ is the dollar trading volume of stock *i* in month *m* of year *t.* [18] Table 1 Panel B presents descriptive statistics for the common stocks used in the paper. An average stock experiences \$16.431 million ETF inflows per year during the sample period.

3.3 V/P Ratio and Relative Growth

To examine whether ETF flows induce stock misvaluation relative to fundamentals-based valuation, I start by examining the associations between ETF flows and two measures of the price-to-fundamentals relation in the fundamental valuation analysis.

3.3.1 V/P Ratio

The first measure is the Value-to-Price (V/P) ratio, which is the ratio of the "intrinsic value" (*V*) to the market price (*P*) (e.g., Lee et al. 1999; Frankel and Lee 1998; Bradshaw 2004). *V* is a forward-looking measure of the fundamental value derived from the residual income model, and the price is measured at the end of 3 months after the fiscal year-end. Intuitively, a *lower* V/P ratio indicates that the market price is high relative to the fundamental value. In particular, the residual income value *V* is estimated as the sum of book value of equity (*B*), the present value of expected residual income (*RI*) for the next two years, and a terminal value after the second year:

$$V_{i,t} = B_{i,t} + \frac{E_{i,t}[RI_{i,t+1}]}{1+r} + \frac{E_{i,t}[RI_{i,t+2}]}{(1+r)^2} + \frac{E_{i,t}[TV_{i,t+2}]}{(1+r)^2}, \tag{2}$$

where residual income is calculated as:

$$RI_{i,\,t+1} = Earn_{i,t+1} - r \times B_{i,t} = (ROE_{i,t+1} - r) \times B_{i,t}\ , \tag{3}$$

that is, the residual income is earnings less a charge against book value to cover the investor's required return. As the proxy for the market earnings expectation, I use the median of

[18] To make sure the results are not driven by the denominator, I also use market capitalization as the scale variable. The main results are qualitatively similar.

the first forecasts made by analysts following 3 months after the fiscal year-end to make sure the information of financial statements is available to analysts when they form earnings expectations.

As discussed in Penman (1997), the terminal value computation is crucial to the valuation model. Following Lee et al. (1999), I estimate the terminal value under the assumption that the abnormal economic profits for each firm fade toward the long-term industry average. Specifically, I assume that firm ROE fades toward a target industry average ROE since industry average ROE incorporates the degree of accounting conservatism in a given industry.

In particular, I estimate terminal value by assuming firm ROE linearly fades to the median industry ROE for 10 years starting from $t+3$ to $t+12$, after which the residual incomes are assumed to be a perpetuity. I calculate the median industry ROE as the median ROE for all stocks in the same industry according to the Fama-French 48-industry classifications (Fama and French 1997) for the past 10 years.

I estimate the forecasted book value using the clean surplus relation:

$$B_{i,\,t+1} = B_{i,t} + (1-k) \times NI_{i,t+1} = B_{i,t} \times [1 + ROE_{i,t+1} \times (1-k)]\,, \tag{4}$$

where k is the dividend payout ratio, which is estimated by dividing actual dividends in the most recent fiscal year by earnings. Following Lee et al. (1999), for companies experiencing negative earnings and paying dividends, I divided the dividends paid by (0.06×Total Assets) to derive an estimate of the payout ratio. Payout ratios of less than zero or greater than one were assigned a value of 0 or 1.

3.3.2 Relative Growth

One concern with using the V/P ratio is that the estimate of V relies on the terminal value assumption. As an alternative measure, I estimate the “relative growth” ratio, which I define as the ratio between the growth rate implied by the fundamentals (operationalized as the growth rate in residual income implied by analysts' consensus earnings forecasts) and the growth rate implied by the stock price (i.e., the growth rate that reconciles the residual income model to the market price, hereinafter the "price-implied growth rate" (Penman 2011)). Intuitively, a *lower* relative growth ratio suggests that the growth rate expected by the market is high compared to the growth rate based on the firm's fundamentals (as reflected in analysts' forecasts).

I calculate the relative growth using a two-stage approach. First, I infer the implied growth rate ($g\_Implied_{i,t}$) from the stock price by reverse engineering the growth rates from the residual income valuation model:

$$P_{i,t} = B_{i,t} + \frac{E_{i,t}[RI_{i,t+1}]}{1+r} + \frac{E_{i,t}[RI_{i,t+2}]}{(1+r)(r-g\_Implied_{i,t})}, \tag{5}$$

$$g\_Implied_{i,t} = r - \frac{E_{i,t}[RI_{i,t+2}]}{(1+r)(P_{i,t}-B_{i,t}-\frac{E_{i,t}[RI_{i,t+1}]}{1+r})}. \tag{6}$$

The price-implied growth rate captures the average long-term firm growth rate that the market expects. As equation (6) suggests, the higher discount rate *r* used in the calculation, the higher the implied growth rate will be.

Second, I calculate the fundamental-based residual income growth rate ($g\_Fundamental_{i,t}$) as:

$$g\_Fundamental_{i,t} = \frac{E_{i,t}[RI_{i,t+2}]-E_{i,t}[RI_{i,t+1}]}{E_{i,t}[RI_{i,t+1}]}. \tag{7}$$

The fundamental-based growth rate captures the average short-term firm growth rate that analysts forecast.

The relative growth is calculated as the ratio of 1 plus the fundamental-based growth rate to 1 plus the price-implied growth rate:

$$g\_Relative_{i,t} = \frac{1+g\_Fundamental_{i,t}}{1+g\_Implied_{i,t}}. \tag{8}$$

A low relative growth indicates that the market has higher expectations for firm growth relative to analysts' forecasts. Thus, for both V/P and relative growth, a lower value suggests a high market price or a market growth expectation relative to fundamentals-based value or growth.

In calculating the V/P and relative growth measures, I assume a constant discount rate r = 10%. As discussed in Section 4.2, the results are similar using discount rates calculated from the Fama-French five-factor model (Fama and French 2015). In the analyses when V/P is used, I omit observations with negative V/P. In the analyses where relative growth is used, I omit observations with the implied growth rate higher than the discount rate. In Section 4.3, I discuss these omitted observations and show that they do not induce bias in my sample. The inferences do not change if they are included.

In panel B of Table 1, the average V/P in the sample period is 0.73. The average is below 1, perhaps because of the fade-rate assumption. The average price-implied growth rate is 5.62%, and the average relative growth ratio is 1.36.

4. Empirical Analyses

4.1 ETF Flows and the Price-to-Fundamentals Relation: Portfolio Analysis

In this section, I examine the association between ETF flows and the price-to-fundamentals relation at portfolio levels. I sort firm-year observations into 5 portfolios on the basis of annual ETF flows and report the time-series averages of firm characteristics, V/P, implied growth, and relative growth for each portfolio.

The results are reported in Table 2. In Panel A, I report firm characteristics for stocks in each portfolio. In column (1), high-flow stocks (portfolio 5) are insignificantly larger than low-flow stocks (portfolio 1). In columns (4) and (5), high-flow stocks and low-flow stocks experience similar levels of index mutual fund flows, but high-flow stocks experience larger non-index mutual fund outflows, although the difference is not statistically significant. In column (6), high-flow firms and low-flow firms have a similar cost of capital as calculated from the Fama-French 5-factor model. In column (7), high-flow firms have lower realized sales growth than low-flow firms. As shown in columns (8)-(10), high-flow firms invest less in R&D, SG&A, and capital expenditure than low-flow firms. In column (11), high-flow firms have lower analysts' long-term earnings forecasts. Columns (7)-(11) suggest that high-flow firms seem to have lower growth potential.

Panel B reports the averages of V/P, implied growth, and relative growth for each portfolio, assuming a constant discount rate r = 10%. In columns (1) and (2), high-flow stocks have statistically significant lower forward E/P and B/P levels than low-flow stocks. In column (3), V/P for high-flow stocks and low-flow stocks are 0.660 and 0.804, respectively. Their difference is -0.144, which is statistically significant at the 1% level, suggesting that high-flow stocks have higher prices relative to their fundamental values calculated from the residual income model. In column (4), the price-implied growth rates for high-flow stocks and low-flow stocks are 6.821% and 4.547%, respectively. The difference is 2.274%, which is significant at the 1% level, suggesting that the market expects higher growth rates for high-flow stocks. In column (5), the relative growth for high-flow stocks and low-flow stocks are 1.253 and 1.434, respectively. The difference between the two is -0.181, which is significant at the 1% level, suggesting that the market forecasts higher growth rate relative to analysts' forecasts for high-flow firms than for low-flow firms. In contrast, as columns (6)-(8) report, there are no significant differences in V/P, implied growth, or relative growth between high-flow stocks and low-flow stocks at the beginning of the year.

In summary, these preliminary portfolio sorting results suggest that high-flow firms have lower sales growth, lower analyst long-term growth forecasts, and lower investment levels, but they have a higher price-to-fundamentals relation as captured by lower V/P and relative growth.

4.2 ETF Flows and the Price-to-Fundamentals Relation: Regression Analyses

Both V/P and relative growth are noisy estimates of the price-to-fundamentals relation. It is possible that the residual income model with analyst forecasts and a constant discount rate does not fully account for firm fundamentals and risk. To investigate these potential deficiencies, I conduct a regression analysis and control for a series of variables capturing firm fundamentals and risk:

$$(\frac{V}{P})_{i,t} \; or \; g\_Relative_{\,i,t} = b_0 + b_1 ETF\ Flow_{i,t} + b_2 Controls + \mu_t + \delta_i + \varepsilon_{i,t}\ . \qquad (9)$$

To ease interpretation, I follow Ben-David et al. (2017a) to standardize V/P, relative growth, and ETF flows by subtracting the sample mean and dividing by the sample standard deviation. The control variables include: (1) fundamental-based ratios to account for potential deficiencies by using analyst earnings forecasts: sales growth, operating profit margin, asset turnover, F-score (Piotroski 2000), and analyst long-term growth forecasts;[19] (2) risk measures to account for potential deficiencies of using a constant discount rate: market beta and the other 4-factor loadings from estimating Fama-French five-factor model using daily returns from 3 months after the previous fiscal year-end to 3 months after the current fiscal year-end (same period as the calculation for stock-level ETF flows); (3) beginning-year firm characteristics to control for potential reverse causality: size, book-to-price, past returns of last four quarters, V/P, and relative growth. I also include both firm and year fixed effects in all the specifications.

[19] F-scores are based on nine financial signals designed to measure the overall improvement or deterioration in firms' financial condition and is a leading indicator of future profitability (Piotroski 2000).

The goal is to assess how V/P and relative growth are associated with ETF flows after controlling for these fundamental ratios and risk measures. In regression equation (9), b_1 is the variable of interest. A negative b_1 suggests a negative association between ETF flows and V/P and relative growth.

The results are reported in Table 3. From columns (1)-(2), I infer that the relation between ETF flows and V/P is negative and statistically significant at the 1% level. In column (1), without control variables, a one-standard-deviation increase in ETF flows is associated with 9.9% of a standard deviation decrease in V/P. In column (2), after controlling for fundamental and risk measures, the magnitude of the coefficient on ETF flows decreases to 8.2% of a standard deviation, which is still statistically significant at the 1% level. From column (1) to column (2), the fundamental measures can explain some variation in V/P. For example, operating profit margin and F-score are positively associated with V/P, suggesting that these variables are value drivers for firms (Piotroski 2000). However, these fundamental and risk measures cannot fully explain the negative association between ETF flows and V/P. In terms of economic magnitude, a one-standard-deviation increase in ETF flows shifts the V/P of the median stock from the 50th percentile to the 42nd percentile of the distribution of V/P.

From columns (3)-(4), I infer that the relation between ETF flows and relative growth is negative and statistically significant at the 1% level. In column (3), without a control variable, a one-standard-deviation increase in ETF flows is associated with 6.0% of a standard deviation increase in the relative growth measure. In column (4), after controlling for fundamental and risk measures, the magnitude of the coefficient on ETF flows is 5.0% of a standard deviation, which is still statistically significant at the 1% level. In terms of economic magnitude, a one-standard-deviation increase in ETF flows shifts the relative growth of the median stock from the 50th percentile to the 45th percentile of the distribution of relative growth.

The results in Table 3 indicate that ETF flows are negatively associated with both V/P and relative growth ratios, and such associations cannot be fully explained by firm fundamentals and risk measures.

4.3 Cases with Negative V/P or Growth higher than the Discount Rate

As discussed in Section 3.3, there are firm-year observations with negative *V* and/or the implied growth rate higher than the discount rate. From a valuation perspective, these observations are difficult to interpret at best, so they are omitted from the main results in Table 3 (termed "omitted sample" hereafter). In this section, I discuss whether including or excluding these observations changes the inferences.

Because it is hard to interpret the two main variables for these observations, I adopt a third valuation measure – the value of speculative growth, which is the difference between the price and the value based on a no-growth assumption, scaled by the price. It is the value the market is placing on the expected growth, captured by price-implied growth rate above, but now applying to all firms.

Specifically, the no-growth value is the total of the book value and the value from short-term earnings forecasts assuming no growth and is calculated as:

$$V\_no\ growth_{i,t} = B_{i,t} + \frac{E_{i,t}[RI_{i,t+1}]}{1+r} + \frac{E_{i,t}[RI_{i,t+2}]}{(1+r)r}. \tag{10}$$

The value of speculative growth is calculated as:

$$Speculative\ Growth_{i,t} = \frac{P_{i,t} - V\_no\ growth_{i,t}}{P_{i,t}}. \tag{11}$$

Speculative Growth captures the proportion in the stock price unexplained by accounting and short-term forecasts. A high value of speculative growth indicates that a large portion of the stock price comes from investors' speculation about firm future growth, instead of anchoring on the accounting in the book value and short-term forecasts.

I first sort observations into 5 portfolios on the basis of the annual ETF flows for the full sample and the omitted sample, respectively, and report the time-series averages of the value of speculative growth for each portfolio.

The results are reported in Panel A of Table 4. In both columns (1) and (2), for the full sample and the omitted sample, the market is pricing speculative growth for the high-flow firms significantly higher than the low-flow firms.

Next, I examine if the association between ETF flows and value of speculative growth can be explained by firm fundamentals and risk measures by estimating the following regression:

$$Speculative\ Growth_{i,t} = b_0 + b_1 ETF\ Flow_{i,t} + b_2 Controls + \mu_t + \delta_i + \varepsilon_{i,t}. \qquad (12)$$

The results are reported in Panel B of Table 4. In both columns (1) and (2), the coefficients on *ETF Flow* are positive and significant at the 1% level, suggesting that ETF flows are positively associated with the speculative portion in prices for both the full sample and the omitted sample. The results imply that including or excluding these omitted observations do not change the main inferences – ETF flows are positively associated with price-to-fundamentals relation for underlying stocks.

4.4 Addressing the Endogeneity Concern

4.4.1 ETF Creations

This association per se is not evidence of a causal impact of ETF flows on the price-to-fundamentals relation and may be the result of various forms of endogeneity. In my analysis, I isolate the variation in ETF flows unrelated to observables (by controlling firm operating performance and risk measures) and cross-sectional differences between firms (by controlling firm fixed effects). However, the association may still reflect a selection effect, confounding the interpretation of the results.

One concern is that if sector ETFs cover glamour industries to begin with (e.g., technology stocks), then there may be a spurious correlation between ETF flows and the price-to-fundamentals relation. To address this, I re-run regression (9) and include industry-year fixed effects in addition to firm fixed effects in order to control for time-variant unobservable industry-specific factors that may drive the results. I obtain similar results. Specifically, in column (2) and (4) of Table 3, the coefficients on *ETF Flow* become -0.063 and -0.048, respectively, which are statistically significant at the 1% level (t-stats = -3.75 and -3.76).

Another concern is that ETF fund managers create ETFs which follow stocks that have recently performed well, and these stocks will continue to perform well due to momentum (Jegadeesh and Titman 1993; Daniel and Moskowitz 2016). In all my regression specifications, I control for the past four-quarter returns. To further mitigate this concern, I re-run regression (9) and use mature ETF flows instead of total ETF flows, where mature ETF flows are calculated by excluding ETF flows during the first year after an ETF is created. In column (2) and (4) of Table 3 the coefficients on *ETF Flow* become -0.070 and -0.049, respectively, which are statistically significant at the 1% level (t-stats = -4.37 and -4.44).

In summary, the results are robust to adding industry-year fixed effects and using mature ETF flows, suggesting that the negative associations between ETF flows and V/P and relative growth cannot be fully explained by a selection effect driven by ETF creations.

4.4.2 BlackRock's Expansion into the ETF Business

More importantly, to address endogeneity concerns, I explore an additional setting – BlackRock's expansion into the ETF business via its acquisition of the iShares ETF family from Barclays.

At the end of 2009, BlackRock acquired BGI (Barclays Global Investors) and its iShares unit in order to establish its ETF business.[20] Barclays sold BGI to raise funds to strengthen its balance sheet, hoping to avoid a possible future bailout by the UK government. The acquisition caused BlackRock to increase its asset holdings by 37%, supporting BlackRock's global leadership in ETFs. As a result of the acquisition, iShares ETFs experienced a significant increase in inflows (relative to ETFs not belonging to iShares), due to BlackRock's ability to utilize its extensive resources (e.g., salesforce, branding and scale benefits, and strong relationships within the distribution channels).[21, 22] Figure 3 plots the flows into an average ETF for iShares ETFs and for ETFs not belonging to iShares, respectively. Figure 3 suggests that before 2009, an average iShares ETF received similar flows to an average non-iShares ETF. However, in 2010, an average iShares ETF received more flows than an average non-iShares ETF. This setting generates exogeneous variation in ETF flows because the acquisition was not driven by the fundamental characteristics of the stocks with a larger fraction of shares held by iShares (the treated sample) but did increase the ETF flows for these stocks, compared with stocks with a smaller fraction of shares held by iShares (the control sample).

[20] The acquisition was announced on June 11, 2009 and was completed in December of 2009. Thus, I expect the effects of increased ETF inflows to start manifesting itself in stock prices from the beginning of 2010.

[21] For example, the assets under management of iShares increased by 19% from 2009 to 2010 (Blackrock 2010).

[22] During a conference call on June 12, 2009, BlackRock emphasized these points to expand the iShares business. For example, Larry Fink, the Chairman of BlackRock talked about the distribution channel with Bank of America Merrill Lynch: "…with our relationship with that entity, with our mutual fund platform there we hope that we could use those relationships to even build a more robust platform for the iShares business."

To verify that treated stocks indeed experience a greater increase in ETF flows compared with control stocks, I estimate the following DiD regression:

$$ETF\ Flow_{i,t} = b_0 + b_1 TREAT \times POST + b_2 Other\ ETF\%_{\ i,t} + b_3 Controls + \mu_t + \delta_i + \varepsilon_{i,t}\ , \quad (13)$$

where *TREAT* equals 1 if the percentage of shares by iShares is above the sample median before the acquisition and equals 0 for other stocks.[23] *POST* equals 1 for 2010 and 0 for 2009. $ETF\ Flow_{i,t}$ is the sum of the 12-month stock-level ETF flows scaled by total dollar trading volume in year 2009 or 2010. To mitigate the concern that the results might be driven by ETFs other than iShares ETFs, I also control for *Other ETF%,* which is the ownership by ETFs not belonging to iShares. I also include both firm and year fixed effects in all the specifications. The results are reported in columns (1) and (2) of Table 5. Across two columns, I find consistent results that treated stocks experience a greater increase in ETF flows compared with control stocks. The economic magnitude seems large. As compared with other stocks, the treated stocks experience 19.7% of a standard deviation increase in ETF flows following the acquisition.

I then adopt the same DiD setting comparing the changes in V/P and relative growth one year before and one year after the acquisition. I estimate the following regressions:

$$(\frac{V}{P})_{i,t}\ or\ g\_Relative_{i,t} = b_0 + b_1 TREAT \times POST + b_2 Other\ ETF\%_{\ i,t} + b_3 Controls + \mu_t + \delta_i + \varepsilon_{i,t}. \quad (14)$$

The results are reported in column (3) for V/P and column (4) for relative growth in Table 5. In column (3), compared with control stocks, the treated stocks experience a 4.9% of a standard deviation decrease in V/P, with t-statistic of -2.40. In column (4), compared with control stocks,

[23] It is identified as the last available weight in Thomson Reuters before the fund management name changed from "Barclays Global Fund Advisors" to "BlackRock Fund Advisors."

the treated stocks experience a 2.4% of a standard deviation decrease in relative growth, with a t-statistic of -1.91.

A potential concern is that V/P and relative growth may have evolved differently prior to the acquisition, which would violate the parallel trend assumption of the difference-in-differences design and make it hard to attribute the changes in V/P and relative growth of treated firms to the increased ETF flows following the acquisition. To address this concern, I re-estimate model (14) using pseudo-event dates prior to 2009. Specifically, I re-run regression (14) by adjusting the event date to 2005, 2006, 2007, and 2008, respectively. In an untabulated analysis, I find that all the coefficients on *TREAT×POST* are statistically insignificant. These results alleviate concerns related to a violation of the parallel trend assumption driving my results.

Another concern is that iShares ETFs may have selected stocks in glamour industries which non-ishares ETFs did not cover before the acquisition. In untabulated results, most industries (based on the Fama-French's 48 industry classification) are present in both samples, except for "Candy & Soda" and "Tobacco Products," which are only included in the treated sample, and "Defense" which is only included in the control sample. However, the main inference does not change if I exclude these firms from the sample. In Column (3) and (4) the coefficients of *ETF Flow* are -0.048 and -0.023, and statistically significant at the 5% and 10% level, respectively (t-stats = -2.36 and -1.83).

In summary, I find statistically significant decreases in V/P and relative growth for treated stocks relative to control stocks following the iShares acquisition. To the extent that the acquisition was indeed an exogenous shock (exogenous relative to the portfolio stocks' fundamental characteristics), the association supports a causal interpretation between ETF flows and the price-to-fundamentals relation documented in my main tests.[24]

5. Exploring the Mechanism: Non-fundamental Demand Inducing Misvaluation

Section 4 suggests that ETF flows increase the price-to-fundamentals relation for underlying stocks. The effects of ETF flows are consistent with non-fundamental demand shocks pushing prices away from the fundamental values, but might be also consistent with other mechanisms. In this section, I explore the alternative mechanisms.

5.1 ETF Flows and Future Firm Performance

It is possible that ETF flows impound fundamental information better or more promptly than the inputs in the residual income model do. For example, high-flow firms might have superior expected future performance or increased future investment levels that analysts fail to fully incorporate into their earnings forecasts. In other words, the estimated *V* for high-flow firms may understate the true *V* and the estimated implied growth may overstate the true market growth expectation, and thus artificially induce the documented low V/P and relative growth ratios.

To investigate this mechanism, I first examine the association between ETF flows and subsequent firm performance as captured by operating performance, sales growth, and investments. Next, relying on the premise that non-fundamental demand shocks subsequently revert, whereas fundamental information leads to permanent price changes, I examine the association between ETF flows and subsequent stock returns.

Specifically, I estimate the following regressions:

[24] One caveat with this analysis is the time period immediately followed the financial crisis. To alleviate this concern, I re-run regression (14) and define the post period as 2010-2014, and the pre period as 2005-2009, to smooth out the effects of the financial crisis. The results are qualitatively similar.

$$RNOA_{i,t+1} \text{ or } Sales\ Growth_{i,t+1} \text{ or } INVST_{i,t+1} \text{or } Return_{i,t+1} = b_0 + b_1 ETF\ Flow_{i,t} + b_2 Controls + \mu_t + \delta_i + \varepsilon_{i,t}\ , \quad (15)$$

where $RNOA_{i,t+1}$ is the return on net operating assets in the subsequent year, which is calculated as operating income divided by net operating assets. $Sales\ Growth_{i,t+1}$ is the sales growth rate in the subsequent year. $INVST_{i,t+1}$ is one of the following variables in the subsequent year: 1) CAPEX scaled by beginning-year total assets, 2) R&D scaled by beginning-year total assets, 3) SG&A scaled by beginning-year total assets. To ease interpretation, I standardize *RNOA*, *Sales Growth*, *INVST* and *ETF Flow* by subtracting the sample mean and dividing by the sample standard deviation. $Return_{i,t+1}$ is the buy-and-hold stock return during the subsequent year.

The results are reported in Table 6. In columns (1) and (2), the coefficients on *ETF Flo*w are negative and significant at the 10% and 5% levels, respectively, suggesting that ETF flows are negatively associated with subsequent operating performance and sales growth.[25] In columns (3) and (4), the coefficients on *ETF Flo*w are insignificant, suggesting that ETF flow are not significantly associated with subsequent investments in capital expenditure and R&D. In column (5), the coefficient on *ETF Flo*w is significant at the 1% level, suggesting that ETF flows are negatively associated with future firm investments in SG&A. More importantly, in column (6), the coefficient on *ETF Flo*w is negative and statistically significant at the 1% level (t-stat = -3.7), suggesting that ETF flows are negatively associated with subsequent stock returns. A one-standard-deviation increase in ETF flows is associated with a 3.7% decrease in subsequent annual stock returns.

[25] Due to conservative accounting, if a high-flow firm invest more, it might earn lower RNOA subsequently. However, columns (9)-(11) in Table 2 suggest that high-flow firms invest less in R&D, SG&A, and CAPEX.

In summary, this section suggests that an increase in ETF flows is not associated with superior future firm performance and investments, and is negatively associated with future stock returns. These findings are more consistent with the conjecture that the non-fundamental demand driven by ETF flows push prices away from fundamental values than the conjecture that ETF flows impound fundamental information more promptly than the inputs in the residual income model do.

5.2 ETF Flows and Risk

It is possible that ETF inflows are associated with lower risk. If so, then the constant discount rate used in the paper might underestimate the true discount rate and thus the intrinsic value *V* and might overestimate the implied growth rate for high-flow firms. However, as column (6) of Panel B in Table 2 suggests, there is no significant difference in the cost of capital calculated from the Fama-French (2015) five-factor model of high-flow vs. low-flow firms, and I control for the Fama-French (2015) five-factor loadings as risk measures in all the regressions. Nevertheless, I re-run regression (9), using the cost of capital calculated from the five-factor model, and find that the coefficients on *ETF Flow* in Columns (2) and (4) of Table 3 become -0.062 and -0.037, respectively, which are still significant at the 1% level (t-stats = -3.82 and -3.64).

Another possible alternative explanation under the risk hypothesis comes from Merton (1987), where increased investors' attention can reduce the cost of capital. Lehavy and Sloan (2008) find empirical evidence consistent with the Merton model. If ETF flows increase investors' attention, then the effect of ETF flows is more driven by increased investors' attention rather than the non-fundamental demand mechanism. To examine this possibility, I estimate the following regression:

$$Attention_{i,t} = b_0 + b_1 ETF\ Flow_{i,t} + b_2 Controls + \mu_t + \delta_i + \varepsilon_{i,t}. \quad (16)$$

I use investors' participation during conference calls to capture investors' attention. More specifically, I use the two proxies for attention: (1) *#Questions* is the total number of questions being asked at conference calls during the year and (2) *Question Length* is the total number of words in questions being asked at conference calls during the year. I interpret more questions or longer questions during conference calls as higher investors' attention. I obtain conference call transcripts from the Thomson Reuters Street Events database.[26] The results are reported in Table 7. The results suggest that ETF flows are associated with fewer questions and shorter questions from investors during conference calls, implying that ETF flows are associated with lower investors' attention. Therefore, it seems less likely that my main results are driven by increased investors' attention with ETF flows.

Notwithstanding the above arguments and tests, a caveat in interpreting my study is that I cannot fully rule out the possibility that ETF flows may affect the price-to-fundamentals relation through a risk channel.

5.3 Cross-sectional Analyses

I conduct a series of cross-sectional analyses to provide indirect evidence on the role of ETF flows in stock misvaluation based on three predictions. First, I expect that the demand shocks have a stronger effect on stock prices when the stock has an imperfectly competitive equity market. If a stock has a perfectly competitive market, investors are price-takers, and the demand curve for the stock should be horizontal with stock prices immune to demand shocks (Shleifer 1986). Therefore, I expect the main results in Table 3 to be stronger for stocks with less perfectly competitive equity markets, as proxied by the number of shareholders (Armstrong et al. 2010). To test this conjecture, I conduct the following regression analysis:

[26] I conduct this analysis for S&P 1500 companies.

$$(\frac{V}{P})_{i,t} or\ g\_Relative_{i,t} = b_0 + b_1 ETF\ Flow_{i,t} \times Few\ SH_{i,t-1} + b_2 ETF\_Flow_{i,t} + b_3 Few\ SH_{i,t-1} + b_4 Controls + \mu_t + \delta_i + \varepsilon_{i,t}\ , \quad (17)$$

where $Few\ SH_{i,t-1}$ equals 1 if the number of shareholders is below the median at the beginning of year *t.* A negative b_1 suggests that the ETF-related misvaluation is stronger for stocks with less perfectly competitive markets. The results are reported in column (1) and column (2) of Table 8. The coefficients on *ETF_Flow× Few SH* are both negative and statistically significant, consistent with the conjecture that the ETF-related misvaluation are stronger with less perfectly competitive markets.

Next, I expect sophisticated active investors to exploit and eliminate this misvaluation. To proxy for active investors, I use non-index mutual funds with high active shares as calculated by Cremers and Petajusto (2009). As Cremers and Petajusto (2009) find, non-index mutual funds with higher active shares outperform those with lower active shares in terms of benchmark-adjusted returns before and after expenses. Active shares of a fund represent the shares of portfolio holdings that differ from the benchmark index holdings. Specifically, the active share of a fund is calculated as follows:

$$Active\ Share = \frac{1}{2} \sum_{i=1}^{N} |w_{fund,i} - w_{index,i}|\ , \quad (18)$$

where $w_{fund,i}$ and $w_{index,i}$ are the portfolio weights of stock *i* in the fund and in the index, and the sum is taken over all stocks. I define non-index funds with at least 60% *Active Share* as active investors and other non-index funds as closet indexers, since Cremers and Petajisto suggest that level as a rough break-point for closet indexing. I conjecture that stocks with lower ownership by active investors (i.e., higher ownership by closet indexers) at the beginning of the year exhibit a stronger ETF-related misvaluation. To test this conjecture, I estimate the following regression equation:

$$\frac{V}{P}_{i,t} \; or \; g\_Relative_{i,t} = b_0 + b_1 ETF\ Flow_{i,t} \times Low\ Active_{i,t-1} + b_2 ETF\_Flow_{i,t} + b_3 Low\ Active_{i,t-1} + b_4 Controls + \mu_t + \delta_i + \varepsilon_{i,t}\ , \tag{19}$$

where $Low\ Active_{i,t-1}$ equals 1 if ownership by active investors is below the median at the beginning of year *t*. The results are reported in Column (3) and Column (4) of Table 8. The coefficients on *ETF Flow*×*Low Active* are both negative and statistically significant, suggesting that the ETF-related misvaluation is stronger with less activity by active investors.

Third, I predict ETF-related misvaluation to be stronger when it is more costly for arbitrageurs, as captured by: 1) lack of close substitutes and 2) short-selling costs. First, in a riskless hedge, the residual variance of returns to the zero-investment hedge after netting out the long and short position is zero (Wurgler and Zhuravskaya 2000). The arbitrageur reduces the residual variance of returns in the hedge portfolio if he can find close substitute stocks to the stocks subject to ETF-related misvaluation. Therefore, I conjecture the ETF-related misvaluation to be stronger for stocks without close substitutes. Following Mashruwala et al. (2006), I use idiosyncratic volatility to capture whether a stock is likely to have close substitutes. Second, with more shares owned by lendable ETFs, the number of shares of underlying stocks available for shorting goes up, possibly reducing the cost of shorting those shares (Glosten et al., 2018), which might reduce the overvaluation. Therefore, I conjecture the ETF-related overvaluation to be stronger for stocks with fewer shares held by lendable ETFs. To test these conjectures, I estimate the following regressions:

$$\frac{V}{P}_{i,t} \; or \; g\_Relative_{i,t} = b_0 + b_1 ETF\ Flow_{i,t} \times High\ IVOL_{i,t-1} + b_2 ETF\_Flow_{i,t} + b_3 High\ IVOL_{i,t-1} + b_4 Controls + \mu_t + \delta_i + \varepsilon_{i,t}\ , \tag{20}$$

$$\frac{V}{P}_{i,t} \; or \; g\_Relative_{i,t} = b_0 + b_1 ETF\ Flow_{i,t} \times Low\ Lend_{i,t-1} + b_2 ETF\_Flow_{i,t} +$$

$$b_3 Low\ Lend_{i,t-1} + \; b_4 Controls + \mu_t + \delta_i + \; \varepsilon_{i,t} \; , \tag{21}$$

where $High\,IVOL_{i,t}$ equals 1 if idiosyncratic return volatility is above the median and idiosyncratic return volatility is calculated as the standard deviation of residuals from a market model (with value-weighted CRSP index) that uses 252 daily returns during year *t-1*. $Low\ Lend_{i,t-1}$ equals 1 if ownership by lendable ETFs is below the median. The results are reported in columns (5)-(8) of Table 8. In column (5) and column (6), the coefficients on *ETF Flow*× *High IVOL* are negative and significant at the 5% level. In column (7) and column (8), the coefficients on *ETF Flow*× *Low Lend* are negative and significant at the 5% level. These results are consistent with the conjecture that the ETF-related misvaluation is stronger for stocks without close substitutes or with high short-selling costs.

6. Indirect Evidence on ETF-related Misvaluation: Firm Policies

Finally, I examine whether high-flow firms exhibit behavior typically associated with perceived overvaluation.

6.1 Share Issuances/Repurchases

Baker (2009) suggests that firms supply more shares as a response to non-fundamental investor demand. Loughran and Ritter (1995) document that firms take advantage of mispricing by issuing equity when it is overpriced and buying it back when it is undervalued. If ETF flows induce stock misvaluation, then I expect that high-flow firms are more likely to issue seasoned equity offers (SEOs) and/or repurchase fewer shares than low-flow firms. Relatedly, I expect high-flow firms have greater insider sales. To test these conjectures, I estimate the following regressions:

$$SEO_{i,t+1} \; or \; Adjusted\ Repurchases_{i,t+1} \; or \; Insider\ Sales_{i,t+1} = b_0 +$$

$$b_1 ETF\ Flow_{i,t} + b_2 Controls + \mu_t + \delta_i + \; \varepsilon_{i,t}, \tag{22}$$

where $SEO_{i,t+1}$ equals one if the firm has an SEO during the subsequent year and zero otherwise. $Adjusted\ Repurchases_{i,t+1}$ equals the total amount of share repurchases after removing the effects of share issuances (Fama and French 2001; Skinner 2008) and potential repurchases for exercised options (Ferri and Li 2018), scaled by beginning-year total assets.[27] $Insider\ Sales_{i,t+1}$ is the ratio of shares sold to the sum of shares sold and purchased by all insiders during the subsequent year (Khan et al. 2012; Priotroski and Roulstone 2005). To ease interpretation, I standardize *Adjusted Repurchases*, *Insider Sales*, and *ETF Flow* by subtracting the sample mean and dividing by the sample standard deviation.

The results are reported in the Panel A of Table 9. In column (1), the coefficient on *ETF Flow* is positive and significant at the 10% level (*t-stat* = 1.85), consistent with the conjecture that with high ETF flows, firms are more likely to issue SEOs during the next year. In column (2), the coefficient on *ETF Flow* is negative and significant at the 10% level (*t-stat* = - 1.89), consistent with the notion that with high ETF inflows, firms repurchase fewer shares during the next year. In column (3), the coefficient on *ETF Flow* is positive and significant at 1% level (*t-stat* = 2.99), consistent with the conjecture that firm insiders subsequently sell more shares following an increase in ETF flows. These results suggest that when investors indirectly buy more stock shares of a specific firm (via ETF flows), the firm is more likely to issue SEOs, tends to repurchase fewer shares, and has greater insider sales.

[27] The repurchase after removing the effect of share issuances is calculated as follows: if the firm uses the treasury stock method for repurchase, then it equals the increase in common treasury stock. If the firm uses the retirement method, then it equals total repurchases minus stock issuance and change in preferred stock redemption value. The effect of repurchasing exercised options equals the product between the number of options exercised and average stock price during the year. The number of option exercised data is available on Compustat starting from 2004.

6.2 Firm Disclosures

If ETF flows induce misvaluation, then I expect firms' disclosures to vary with ETF flows. For example, high-flow firms may reduce voluntary disclosures about their future earnings and remain silent in an attempt to maintain the optimistic valuation and low-flow firms may increase voluntary disclosures in an attempt to correct the pessimistic outlook (Bergman and Roychowdhury 2008). Following Bergman and Roychowdhury (2008), I focus on long-horizon earnings forecasts, since over short horizons, disclosure policy is greatly affected by the disciplining effect of earnings announcements.

To test these conjectures, I estimate the following regression:

$$\#Guidance_{i,t+1} \ or \ \#WalkUp_{i,t+1} \ or \ \#WalkDown_{i,t+1} = b_0 + b_1 ETF\ Flow_{i,t} + b_2 Controls + \mu_t + \delta_i + \varepsilon_{i,t} \ , \qquad (23)$$

where $\#Guidance_{i,t+1}$ is the number of long-horizon management earnings forecasts issued during the subsequent year. $\#WalkUp_{i,t+1}$ and $\#WalkDown_{i,t+1}$ are the numbers of long-horizon, walk-up and walk-down earnings forecasts issued during the subsequent year, respectively. A long-horizon forecast is one made at least 90 days before the forecast period end date, and a walk-up or walk-down forecast is one that is higher or lower than the consensus estimate, respectively. To ease interpretation, I standardize *#Guidance*, *#WalkUp*, *#WalkDown*, and *ETF Flow* by subtracting the sample mean and dividing by the sample standard deviation.

The results are reported in the Panel B of Table 9. In column (1), the coefficient on *ETF Flow* is negative and significant at 1% level (*t-stat* = -2.59), suggesting a negative association between ETF flows and firm disclosures. Columns (2) and (3) suggest that the negative association between ETF flows and number of forecasts are driven by walk-up forecasts. The results imply that firm voluntary disclosures depend on ETF flows in a way that seems consistent with ETF flows inducing misvaluation, as high-flow firms reduce the amount of voluntary disclosures and low-flow firms increase the amount of positive voluntary disclosures. I interpret these results as high-flow firms remaining silent in an attempt to maintain the optimistic valuation and low-flow firms increasing the amount of voluntary disclosures in an attempt to correct the pessimistic valuation.

Overall, the results in Section 6.1 and 6.2 provide indirect support for the notion that the price-to-fundamentals measures do indeed capture stock misvaluation. They also suggest that ETF-driven misvaluation may have real effects on firm policies.

7 Additional Analyses

7.1 Cross-section of Stock Returns

If high-flow stocks are overvalued, then active investors should be able to form trading strategies to exploit the misvaluation. To test this conjecture, at the *beginning* of every calendar month, stocks are ranked in ascending order based on the last available 12-month ETF flows. In Table 10, I report the monthly weighted-average portfolio raw returns in column (1) and Fama-French five-factor alphas in column (2). I also report the zero-cost portfolio returns from buying the top 20% stocks and shorting the lowest 20% stocks. The results indicate that stocks with high-ETF flows underperform low-flow stocks by around -0.7% per month, consistent with the conjecture that active investors could potentially earn abnormal profits by trading against high ETF flows.

7.2 Mutual Fund Flows

In this section, I include non-index mutual fund flows and index mutual fund flows in the regression (9) to alleviate the concern that the effects of ETF flows are associated with and might be subsumed by the traditional mutual fund flows. Non-index and index mutual fund flows at the fund level are calculated as:

$$NonIndex\ MF_{j,\,t} or\ Index\ MF_{j,\,t} = TNA_{j,\,t} - TNA_{j,t-1} \times (1 + R_{j,t}), \tag{24}$$

where $TNA_{j,t}$ is total net assets and $R_{j,t}$ is the fund return during quarter *t*. Since mutual fund flows are not directly observable, I follow prior literature to infer flows by assuming that inflows and outflows occur at the end of the quarter and that existing investors reinvest dividends and other distributions in the fund (Zheng 1999). I then calculate the stock-level mutual fund flows in the same way as I calculate the stock-level ETF flows and re-run the main regression (9). The results are reported in Table 11. Non-index mutual fund flows have no significant effects on V/P and relative growth, perhaps because some non-index mutual funds are skilled in picking stocks with strong fundamental values (higher *V*) and they have control over the price impact of their trades as described in Section 2.2. Compared with index mutual fund flows, ETF flows have much larger effects on V/P and relative growth, perhaps because ETFs impose larger non-fundamental demand on stock shares. These results suggest that the effects of ETF flows are not subsumed by mutual fund flows and, in fact, have larger effects on the price-to-fundamentals relation for underlying stocks.

8. Conclusion

The tremendous growth of ETFs can be explained by the fact that they offer investors continuous access to diversified and inexpensive investment vehicles. However, ETFs may also attract non-fundamental demand, which is passed on to the underlying stocks. This mechanism can push stock prices away from fundamental values, inducing stock misvaluation. Consistent with this conjecture, I find ETF flows are positively associated with the price-to-fundamentals relation for underlying stocks. To investigate whether these documented effects of ETF flows are causal, I use BlackRock's expansion into the ETF business as an exogenous shock supporting the causal interpretation of the results. High-flow firms tend to underperform low-flow firms in operating performance, sales growth, investment levels, and stocks returns during the subsequent year, which is consistent with non-fundamental demand shocks inducing misvaluation. Further supporting this view, I show that the effects of ETFs are stronger in stocks with: a less competitive market (i.e., with prices more sensitive to demand shocks), lower ownership by active investors, and more costly arbitrage constraints (i.e., stock where misvaluation is less likely to be corrected). Lastly, I find high-flow firms exhibit behavior typically associated with perceived overvaluation, i.e., greater equity issuances, less repurchases, and greater insider sales.

In addition to contributing to research on index investing and the literature on ETFs, my study highlights the importance of focusing on firm fundamentals particularly with the increase in passive investments in current financial markets. Regulators might consider further relaxing short-sale constraints and encouraging ETFs to lend out more shares.

REFERENCES

Admati A., Pfleiderer, P. A theory of intraday patterns: volume and price variability. *Review of Financial Studies* 1: 33-40.

Anadu, K., Kruttli, M., McCabe P., Osambela, E., and Shin, C., 2018. The shift from active to passive investing: potential risks to financial stability. Finance and Economics Discussion Series 2018-060. Washington: Board of Governors of the Federal Reserve System, https://www.federalreserve.gov/econres/feds/files/2018060pap.pdf.

Appel, I., Gormley, T., and Keim, D., 2016. Passive investors, not passive owners. *Journal of Financial Economics* 121(1): 111-141.

Armstrong, C., Core, J., Taylor, D, Verrecchia, R.E. 2011. When Does Information Asymmetry Affect the Cost of Capital? *Journal of Accounting Research*, 49 (1), pp. 1-40.

Azar, J., Schmalz, M., and Tecu, I., 2018 Anticompetitive effects of common ownership. *Journal of Finance* 73(4): 1513-1565.

Baker, M., 2009. Capital market-driven corporate finance. Annual Review of Finance and Economics 2009.1: 181-205.

Baker, M., Stein, J., Wurgler, J., 2003. When does the market matter? Stock Prices and the Investment of Equity-Dependent Firms. *The Quarterly Journal of Economics,* 118(3): 969-1005.

Baker, M. and Wurgler, J., 2002. Market Timing and Capital Structure. *Journal of Finance*, 57(1): 1-32.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2017a. Do ETFs Increase Stock Volatility? *Journal of Finance*, Forthcoming.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2017b. Exchange traded funds. *Annual Review of Financial Economics* 9, 165-185.

Bergman, N., and Roychowdhury S., 2008. Investor sentiment and corporate disclosure. *Journal of Accounting Research* 46(5), 1057-1083.

Bhojaj, S., Mohanram, P., and Zhang, S., 2018. ETFs and information transfer across firms. Working paper. Cornell University, University of Toronto, and Iowa State University.

BlackRock, Inc., Annual Report 2010.

Boone, A., and White, J., 2015. The effect of institutional ownership on firm transparency and information production. *Journal of Financial Economics* 117(3): 508-533.

Bradshaw, M., 2004. How do analysts use their earnings forecasts in generating stock recommendations? *The Accounting Review* 79: 25-50.

Brown, D., Davis, S., and Ringgenberg, M., 2018. ETF arbitrage and return predictability. Working paper. University of Arizona, University of Colorado at Boulder and University of Utah.

Chen, Q., Goldstein, I., and Jiang, W., 2006. Price informativeness and investment sensitivity to stock price. *Review of Financial Studies* 20(3): 619-650.

Daniel, K., and Moskowitz, T., 2016. Momentum crashes. *Journal of Financial Economics* 122(2): 221-247.

Da, Zhi, and Sophie Shive, 2016, Exchange Traded Funds and Asset Return Correlations, *European Financial Management* 24: 136-168.

Fama, E., French, K., 1997. Industry Costs of Equity. *Journal of Financial Economics* 43(2): 153-193.

Fama, E., French, K., 2001. Disappearing dividends: changing firm characteristics or lower propensity to pay? *Journal of Financial Economics* 60, 3–43.

Fama, E., French, K., 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116, 1-22.

Ferri, F. and Li, N., 2018. The effect of option-based compensation on payout policy: evidence from FAS123R. *Journal of Financial and Quantitative Analysis*, Forthcoming.

Foster F., Viswanathan S., 1990. A theory of the interday variations in volume, variance, and trading costs in securities markets. Review of Financial Studies 3(4): 593-624.

Frankel, R., and C. M. C. Lee. 1998. Accounting valuation, market expectation, and cross-sectional stock returns. *Journal of Accounting and Economics* 25 (June): 283-320.

Frazzini, Andrea, and Owen A. Lamont, 2008, Dumb Money: Mutual Fund Flows and the Cross-Section of Stock Returns, *Journal of Financial Economics* 88, 299–322.

Glosten, L., Milgrom, P., 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics* 14, 71-100.

Glosten, L., Nallareddy, S., and Zou, Y., 2018, ETF Trading and Informational Efficiency of Underlying Securities, Working Paper, Columbia University and Duke University.

Hamm, S., 2014. The effect of ETFs on stock liquidity. Working paper.

Investment Company Institute, 2018. Investment Company Fact Book, available at: https://www.ici.org/pdf/2018_factbook.pdf.

Israeli, D., Lee, C., Sridharan, S., 2017. Is there a dark side to exchange traded funds (ETFs)? An information perspective. *Review of Accounting Studies* 22: 1048-1083.

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

Jensen, M., 2004. Agency costs of overvalued equity. Financial Management 34(1).

Johnson M., 2016. Why active fund managers should cheer the rise of ETFs. *Financial Times*, available at: https://www.ft.com/content/df1d95ec-bac4-11e6-8b45-b8b81dd5d080.

Khan, M., Kogan, L., and Serafeim G., 2012. Mutual fund trading pressure: firm-level stock price impact and timing of SEOs. *Journal of Finance* 67(4): 1371-1395.

Lee, C. and Swaminathan, B., 1999. Valuing the Dow: a bottom-up approach. Financial Analysts Journal 55, 4-23.

Lehavy, R., and Sloan R., 2008. Investor recognition and stock returns. *Review of Accounting Studies* 13:327-361.

Lettau, M., Madhavan, A., 2018. Exchange-Traded Funds 101 for Economists. *Journal of Economic Perspectives*, 32 (1): 135-54.

Loughran T., and Ritter, J., 1995. The new issues puzzle. *Journal of Finance* 50(1), 23-51.

Madhavan, Ananth N., 2016, *Exchange-Traded Funds and the New Dynamics of Investing*, Oxford University Press.

Mashruwala, C., Rajgopal, S., and Shevlin, T., 2006. Why is the accrual anomaly not arbitraged away? The role of idiosyncratic risk and transaction costs. *Journal of Accounting and Economics* 42: 3-33.

Massa, M., Schumacher, D., and Wang, Y., 2016. Who is afraid of BlackRock? Working Paper, INSEAD, McGill University, and McMaster University.

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

Penman, S.H., 2011. Accounting for Value. Columbia University Press.

Penman, S. H., 1991. An evaluation of accounting rate-of-return. *Journal of Accounting, Auditing, and Finance* (Spring): 233-255.

Penman, S.H., 1997. A synthesis of equity valuation techniques and the terminal value calculation for the DDM. *Review of Accounting Studies* 2, 303-323.

Piotroski, J., 2000. Value investing: the use of historical financial statement information to separate winners from losers. *Journal of Accounting Research* 38, 1-41.

Piotroski, J. and Roulstone, D., 2005. Do insider trades reflect both contrarian beliefs and superior knowledge about future cash flow realizations? Journal of Accounting and Economics 39, 55-82.

Piwowar M., 2017. SEC-NYU dialogue on exchange-traded products.

PricewaterhouseCoopers, 2015. ETF 2020: Preparing for a new horizon. Available at: https://www.pwc.com/gx/en/industries/financial-services/asset-management/publications/etf-2020.html.

Rubinstein, M. (1989). Market basket alternatives. Financial Analysts Journal, 45, 20–29.

Shleifer, A., 1986. Do Demand Curves for Stocks Slope Down? *Journal of Finance* 41, 579–90.

Skinner, D., 2008. The evolving relation between earnings, dividends, and stock repurchases. *Journal of Financial Economics* 87, 582-609.

Stambaugh, R. 2014. Presidential address: investment noise and trends. *Journal of Finance* 69: 1415-1453.

Stambaugh, R., Yu, J., and Yuan, Y., Arbitrage asymmetry and the idiosyncratic volatility puzzle. *Journal of Finance* 70.

Wurgler, Jeffrey, and Ekaterina Zhuravskaya, 2002, Does Arbitrage Flatten Demand Curves for Stocks, *Journal of Business* 75(4), 583–608.

Wurgler, J. 2000. Financial markets and the allocation of capital. Journal of Financial Economics, 58, 187-214.

Wurgler, J., 2011, *Challenges to Business in the Twenty-First Century: The Way Forward*. Chap. On the Economic Consequences of Index-Linked Investing (American Academy of Arts and Sciences).

Yu, L. 2005. Basket securities, price formation, and informational efficiency. Working paper: University of Notre Dame.

Zheng, L. 2002. Is money smart? A study of mutual fund investors' fund selection ability. Journal of Finance 54(3).

Appendix 1: Examples of critics on ETFs from Practitioners:

1. *"... as ETFs attract capital, they have to buy large amounts of these stocks, further fueling their rise… Is Apple a safe stock or a stock that has performed well of late? Is anyone thinking about the difference?"*

   "Yes. They are great companies. But the ETFs probably have accentuated the flow of capital into those stocks. Nothing works forever. Things that are most hyped and usually the things that are most loved produce the most disappointment and the most pain."

   – Howard Marks, Co-Founder and Co-Chairman of Oaktree Capital Management

2. *"The weapons of mass destruction during the Great Financial Crisis were three-letter words: CDS (credit default swap), CDO (collateralized debt obligation), etc. The current weapon of mass destruction is also a three-letter word: ETF (exchange-traded fund). When the world decides that there is no need for fundamental research and investors can just blindly purchase index funds and ETFs without any regard to valuation, we say the time to be fearful is now."*

   – FPA Capital Fund, Inc. in *Letters to Shareholders*

3. *"ETFs occasionally swing to illogical extremes when popularity leads to overpriced and overweighted constituents."*

   – Bruce Berkowitz, Founder of Fairholme Capital Management

4. *"So what's been going on for almost a decade now is a constant flow of money out of actively managed securities which have acted in essence as the bank to fund inflows, into equity ETFs ... that is what forms the foundation of the bubble activity we have today."*

   – Steven Bregman, Co-Founder of Horizon Kinetics LLC

5. *"Passive investing is the bubble right now, and that's a great danger."*

   "But in the market today, the danger is that you have all this money pouring in America into ETFs and ETFs are sort of almost blind buying. You just buy these ETFs and I always question the fact that if you're buying these stocks and you really don't know what you own, you're prone to these periods of time – there could be some kind of crisis and there could be a problem."

   – Carl Icahn, Chairman of Icahn Enterprises, advising U.S. President Donald Trump on Financial Regulation

6. *"The inherent irony of the efficient market theory is that the more people believe in it and correspondingly shun active management, the more inefficient the market is likely to become."*

   – Seth Klarman, Founder of the Baupost Group

Source:

1. There They Go Again… Again: Memos from Howard Marks (2017): https://www.oaktreecapital.com/docs/default-source/memos/there-they-go-again-again.pdfs

   Howard Marks Blames ETFs for Overpriced FANG Stocks (*Yahoo! Finance* 2018): https://finance.yahoo.com/news/howard-marks-blames-etfs-overpriced-165248183.html

2. FPA Capital Fund Letter to Shareholders (2017): https://fpa.com/docs/default-source/funds/fpa-capital-fund/literature/fpa-capital-fund-ann-rpt-3-31-17.pdf?sfvrsn=4.

3. Bruce Berkowitz – ETF's Occasionally Swing To Illogical Extremes (*ValueWalk* 2017): https://www.valuewalk.com/2017/05/bruce-berkowitz-etfs/

4. Steven Bregman: America's Hottest Investment Product has Created "the Greatest Bubble Ever" (*Business Insider* 2017): https://www.businessinsider.com/steven-bregman-etfs-passive-investing-greatest-bubble-ever-2017-6

5. "These Are the Rumblings." Carl Icahn Warns the Stock Market Will Implode (*Fortune* 2018): http://fortune.com/2018/02/06/carl-icahn-stock-market-warnings/

   Will the Stock Market Eat Itself? (*Forbes* 2017): https://www.forbes.com/sites/robisbitts2/2017/02/13/will-the-stock-market-eat-itself/#61a1beac4958

6. A Quiet Giant of Investing Weighs In on Trump (*The New York Times* 2017): https://www.nytimes.com/2017/02/06/business/dealbook/sorkin-seth-klarman-trump-investors.html

Appendix 2: Variable Definitions

Variable Definitions		
Variable	***Variable Name***	***Description***
Main Variables		
ETF flows	ETF Flow	• Annual stock-level ETF flows, which are accumulated for 12 months from 3 months after the last fiscal year-end to 3 months after the current fiscal year-end, scaled by the dollar trading volume during the year. • The monthly ETF flows are calculated as the weighted average of the flows for all ETFs holding the stock, where the weight is the stock *i*'s portfolio weight in each ETF. • The data are from ETF.com, CRSP, and Thomson-Reuters Mutual Fund holding database.
Value-to-price ratio	V/P	• The Value-to-Price ratio, which is the ratio of the "intrinsic value" (*V*) to the market price (*P*). • *V* is a forward-looking measure of fundamental value derived from the residual income model, and the price is the price at the end of 3 months after the fiscal year-end. • The data are from Compustat, I/B/E/S, and CRSP.
Relative growth ratio	Relative Growth	• The ratio of 1 plus the growth rate implied by fundamentals to 1 plus growth rate implied by the stock price. • The fundamental residual income growth rate is calculated as the growth rate impeded in analyst forecasts. • The implied firm's long-run growth rate is calculated from the stock price by reverse engineering the growth rate from the residual income valuation model. • The data are from Compustat, I/B/E/S, and CRSP.
Value of speculative growth	Speculative Growth	• The difference between the price and the value based on the no-growth assumption, scaled by the price. • The data are from Compustat, I/B/E/S, and CRSP.
Control Variables		
Sales growth rate	Sales Growth	• Changes in sales, scaled by sales from last fiscal year. • The data are from Compustat.
Operating Profit margin	Profit Margin	• The ratio of operating income to sales. • Operating income is calculated as earnings (net income - preferred dividends + change in marketable securities adjustment + change in cumulative translation adjustment) plus net interest expense (after-tax interest

		expense + preferred dividends – after-tax interest income + minority interest in income – change in marketable securities adjustment). • The data are from Compustat.
Asset turnover	Asset Turnover	• The ratio of sales to net operating assets (NOA) • NOA is the sum of accounts receivable (item RECT), inventory (item INVT), other current assets (item ACO), property, plant, and equipment (item PPENB), intangible assets (item INTAN) and other long-term assets (item AO), minus the sum of accounts payables (item AP), other current liabilities (item LCO) and other long term liabilities (item LO). • The data are from Compustat.
Analyst long-term growth forecast	Analyst LT Forecast	• The median of analyst long-term growth forecasts made 3 months after the fiscal year-end. • The data are from I/B/E/S.
F Score	F_Score	• The sum of the nine binary signals, and is designed to measure the overall improvement, or deterioration, in firms' financial conditions, as developed by Piotroski (2000). • The data are obtained from Compustat.
Market beta and other risk factor loadings	Beta_Market Beta_HML Beta_SMB Beta_RMW Beta_CMA	• The regression coefficients from regressing excess daily returns for a firm on excess market returns over the period 3 months after the previous fiscal year-end to 3 months after the current fiscal year-end. • The market return and risk-free rate are obtained from Ken French's website. • The return data are from CRSP.
Firm size	Size	• Market capitalization of the firm measured 3 months after the last fiscal year-end. • The market capitalization equals the share price × the number of shares outstanding as reported by Compustat. • I use the natural log of firm size in all the regressions.
Book-to-price	Book-to-Price	• Ratio of book value of equity to the price, both measured at 3 months after the last fiscal year-end, from Compustat and CRSP.
Stock Return during 1 quarter, 2 quarter, 3 quarter, and 4 quarter before ETF flows	Ret (-3,-1) Ret (-6,-4) Ret (-9,-7) Ret (-12,-10)	• Buy-and-hold returns 1 quarter, 2 quarter, 3 quarter, and 4 quarter before ETF flows are calculated. • The return data are from CRSP.
Return on net operating assets	RNOA	• A measure of the operating profitability of the firm, as measured by operating income scaled by NOA. • The data are from Compustat.

Forward returns	Return (t+1)	• Annual buy-and-hold returns beginning three months after the fiscal year-end, calculated from CRSP as compounded monthly returns. • The data are from CRSP.
R&D expense	R&D	• R&D expenditures scaled by beginning-year total assets, set to zero if missing. • The data are from Compustat.
SG&A expense	SG&A	• Selling, General, and Administrative Expenses scaled by beginning-year total assets, set to zero if missing. • The data are from Compustat.
Capital expenditure	CAPEX	• Capital expenditures scaled by beginning-year total assets, set to zero if missing. • The data are from Compustat.
Firms with low active investor activity	Low Active%	• Equals 1 if ownership by active investors is below the median at the beginning of the year. • Active investors are non-index mutual funds with at least 60% active shares as calculated by Cremers and Petajusto (2009). • The data are from Martijn Cremers website: https://activeshare.nd.edu/data/.
Firms with a less competitive market	Few Shareholders	• Equals 1 if the number of shareholders is below the median at the beginning of the year. • The data are from Compustat.
Firms less likely to have close substitutes	High IVOL	• Equals 1 if the idiosyncratic return volatility is above the median, and idiosyncratic return volatility is calculated as the standard deviation of residuals from a market model (with value weighted CRSP index) that uses 252 daily returns during the last year. • The data are from CRSP.
Firms with fewer shares owned by lendable ETFs	Low Lend%	• Equals 1 if ownership by lendable ETFs is below the median. • A lendable ETF is one that lend out shares. • The data are from Bloomberg.
Secondary equity offering issuance	SEO	• Equals one if the firm has an SEO during subsequent year, and zero otherwise. • The data are from SDC plantium database.
Adjusted repurchases/Total Assets	Adjusted Repurchases	• Equals the total amount for share repurchases after removing the effect of share issuances and the effect of repurchases for exercised options, scaled by beginning-year total assets (Ferri and Li 2018). • The data are from Compustat.
Insider sales	Insider sales	• The ratio of shares sold to the sum of shares sold and purchased by all insiders during the subsequent year. • The data are from Thomson Financial Insider database.

Number of earnings guidance	#Guidance	• The number of long-horizon management earnings forecasts during the subsequent year. • A long-horizon forecast is one made prior to 90 days before the forecast period end date. • The data are from I/B/E/S Guidance database.
Number of walk-up earnings guidance	#WalkUp	• The number of long-horizon and walk-up earnings forecasts during the subsequent year. • A walk-up forecast is one that are higher than the consensus estimate. • The data are from I/B/E/S Guidance database.
Number of walk-down earnings guidance	#WalkDown	• The number of long-horizon and walk-down earnings forecasts during the subsequent year. • A walk-down forecast is one that are lower than the consensus estimate. • The data are from I/B/E/S Guidance database.
Investor Attention	#Question Question Length	• *#Questions* is the total number of questions being asked at conference calls during the year. • *Question Length* is the total number of the words in questions being asked at conference calls during the year. • I obtain conference call transcripts from the Thomson Reuters Street Events database.
Non-index and index Mutual fund flows	NonIndex MF Flow Index MF Flow	• Annual stock-level non-index and index mutual fund flows, which are accumulated for 12 months from 3 months after the last fiscal year-end, scaled by the dollar trading volume during the year. • The monthly mutual fund flows are calculated as the weighted average of the flows for all non-index and index mutual funds holding the stock, respectively, where the weight is the stock *i*'s portfolio weight in each mutual fund. Mutual fund flows at the fund level are calculated as: $TNA_{j,\,t} - TNA_{j,t-1} \times (1 + R_{j,t})$. • I identify a fund as an index fund if its fund name includes a string that identifies it as an index fund or if the CRSP Mutual Fund Database classifies the fund as an index fund. Specifically, I use the same strings as in Appel et al. (2016) to identify index funds. These strings include: *Index, Idx, Indx, Ind_(where_ indicates a space), Russell, S & P, S and P, SandP, SP, DOW, Dow, DJ, MSCI, Bloomberg, KBW, NASDAQ, NYSE, STOXX, FTSE, Wilshire, Morningstar, 100, 400, 500, 600, 900, 1000, 1500, 2000*, and *5000.* • The data are from CRSP Mutual Fund Database.

Figure 1 ETF Ownership Trend

Figure 1 exhibits the time trend for stock ownership by ETFs, index mutual funds, and non-index mutual funds, respectively. For each year, I average the percentage of shares for each stock owned by different types of funds, respectively.

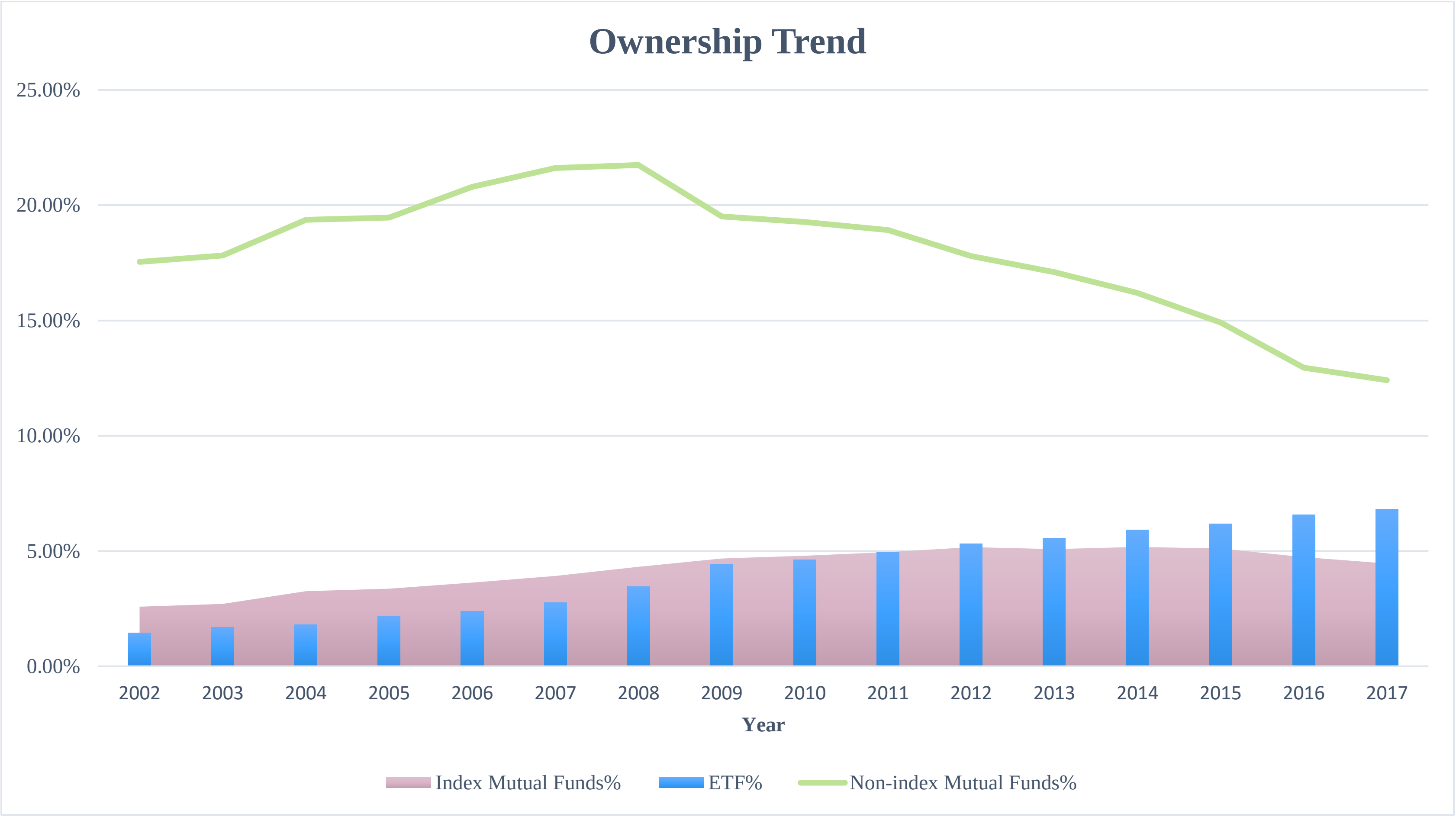

Figure 2 ETF Creation/Redemption Mechanism

This figure exhibits the creation of ETF shares. When there is a large inflow into ETF shares (step 1), the AP short sells the ETF shares to meet investors' demand and buys the underlying securities to hedge the short position (step 2), putting upward pressure on the underlying securities. At the end of the day, the AP transfers the basket of the securities to the ETF sponsor (step 3), and the ETF sponsor creates new ETF shares to cover the AP's short position (step 4). A similar process happens if there is an AP redeeming ETF shares when there are outflows from an ETF.

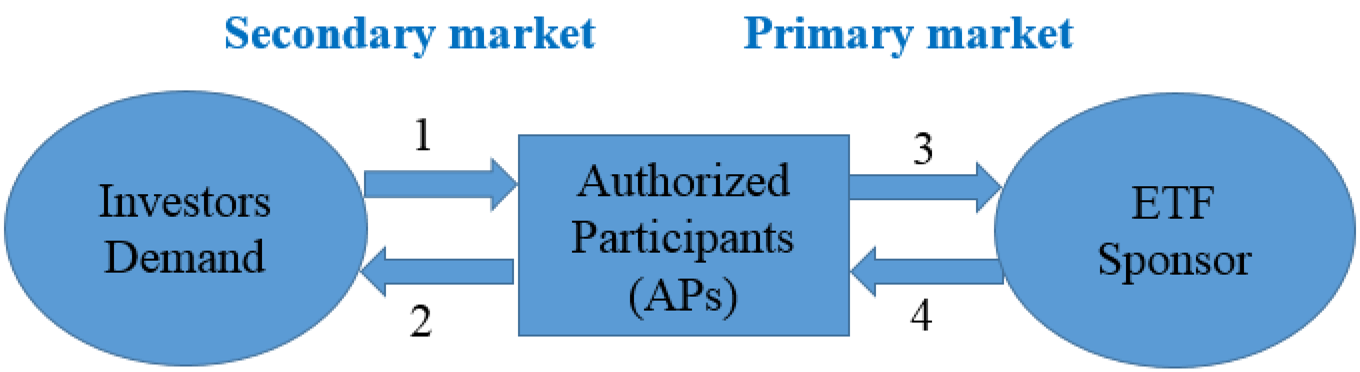

Figure 3 ETF Flow for iShares and Other ETFs

This figure exhibits ETF flows for iShares ETFs and ETFs not belonging to iShares (non-iShares) from 2005 to 2010. For each year, I average the annual ETF flows for iShares ETFs and non-iShares ETFs, respectively.

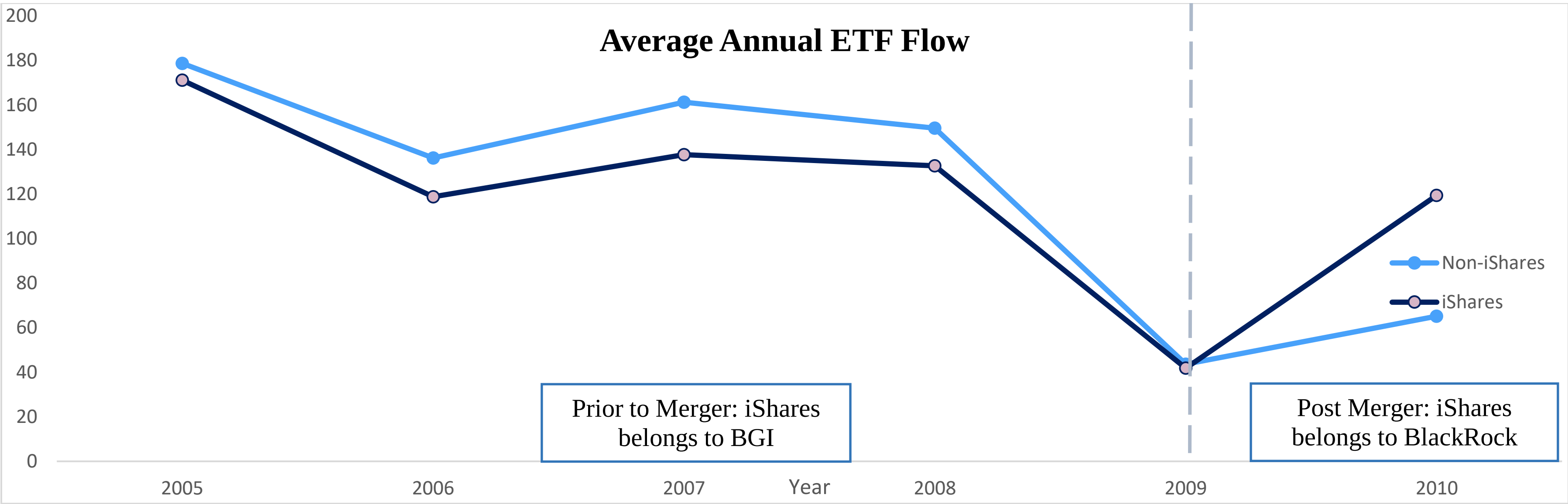

Table 1 Descriptive Statistics

Table 1 presents descriptive statistics of the ETF characteristics and stocks characteristics in my sample. The sample spans from 2002 to 2016.

Panel A ETF Characteristics

	p25	Median	p75	Mean	Std.dev
Monthly Flow (Million)	-8.255	3.886	32.552	32.843	537.277
ETF Market Cap (Million)	83.879	346.084	1,645.253	3,067.940	10,977.583
Monthly Turnover	0.115	0.211	0.479	0.712	6.835
Net Asset Value (Million)	24.565	40.856	62.615	49.381	34.462
Securities Lending Indicator	0.000	1.000	1.000	0.719	0.450
Market-Cap Weight Indicator	0.000	1.000	1.000	0.597	0.491

Panel B Stock Characteristics

	p25	Median	p75	Mean	Std.dev
Market Cap (Million)	318.299	943.018	3,053.086	4,735.040	12,625.420
Book-to-Price	0.311	0.495	0.746	0.582	0.400
ETF Flows (Million)	0.157	2.197	9.155	16.431	32.812
Index Mutual Fund Flows (Million)	-0.001	0.052	2.898	8.394	28.873
Non-index Mutual Fund Flows (Million)	-0.871	-0.004	0.329	-3.969	37.414
ETF Flows*100/$Volume	0.017	0.068	0.165	0.114	0.152
ETF Ownership	0.818%	2.448%	5.925%	4.589%	7.552%
V/P	0.340	0.578	0.877	0.730	0.732
Price-implied Growth Rate	3.565%	6.185%	7.408%	5.620%	10.240%
Relative Growth	1.005	1.112	1.408	1.360	1.352
Speculative Growth	0.111	0.396	0.640	0.300	0.516
Analyst LT Forecast	10.000%	13.000%	17.500%	14.293%	9.171%
F-Score	5.000	6.000	7.000	5.626	1.551
Sales Growth	-0.990%	7.550%	18.542%	10.851%	22.769%
Operating Profit Margin	1.380%	7.893%	13.572%	5.379%	18.215%
Asset Turnover	0.643	1.493	2.687	1.949	3.446
Annual Return	-0.140	0.089	0.326	0.122	0.426
RNOA	-0.027	0.079	0.185	0.100	0.769
#Shareholder (Thousand)	0.290	1.570	7.500	11.360	7.500
Ownership by mutual funds with active share>=60%	0.000%	0.011%	5.579%	4.317%	8.700%
Idiosyncratic volatility (IVOL)	1.400%	1.976%	2.737%	2.198%	1.061%
Lendable ETF%	0.000%	0.900%	3.961%	2.319%	2.978%
SEO Indicator	0.000	0.000	0.000	0.094	0.292
Insider Sales	0.000	0.608	0.994	0.366	0.460
Adjusted Repurchases	0.000%	0.000%	0.714%	1.629%	4.760%
#Guidance	0.000	0.000	3.000	1.350	2.252
#WalkUp	0.000	0.000	1.000	0.502	1.012
#WalkDown	0.000	0.000	1.000	0.789	1.523

Table 2 ETF Flows and the Price-to-Fundamentals Relation: Portfolio Analysis

Table 2 presents the results for portfolio analysis. Monthly stock-level ETF flows are calculated as the weighted average of the flows for all ETFs holding the stock, where the weight is the stock's portfolio weight in each ETF. Each year, all the firms are sorted into 5 portfolios on the basis of the annual ETF flows, which are calculated by adding up monthly stock-level ETF flows for 12 months from 3 months after the previous fiscal year-end, scaled by the total dollar trading volume during the year. Time-series averages of firm characteristics, V/P, and relative growth measures (assuming a constant discount rate 10%) for each portfolio are reported. ***, **, and * represent significance at the 0.01, 0.05 and 0.10 levels, respectively.

Panel A ETF Flows and Firm Characteristics

ETF Rank	(1) Market Cap	(2) ETF Flows*100/ $Volume	(3) ETF Flows	(4) Index MF Flows	(5) Non-index MF Flows	(6) Cost of Capital	(7) Sales Growth	(8) R&D/Total Assets	(9) SG&A/ Total Assets	(10) CAPEX/ Total Assets	(11) Analyst LT Forecast
1-Low	4,214.1	-0.032	-0.137	8.212	-2.889	10.891%	13.778%	0.039	0.213	0.044	16.244%
2	3,741.7	0.019	9.226	5.473	-3.412	10.839%	13.845%	0.044	0.233	0.050	15.513%
3	4,568.0	0.063	11.082	6.740	-3.580	11.003%	11.638%	0.036	0.213	0.048	14.703%
4	6,687.2	0.148	22.896	8.725	-6.893	11.048%	9.320%	0.026	0.196	0.045	13.343%
5-High	5,734.7	0.489	28.525	9.040	-9.573	10.906%	6.697%	0.022	0.175	0.039	12.017%
High-Low	1,520.6	0.521***	28.662***	-0.828	-6.684	0.015%	-7.081%***	-0.017***	-0.038***	-0.005**	-4.227%***

Panel B ETF Flows, V/P, and Relative Growth

ETF Rank	(1) Book-to-price	(2) Forward E/P	(3) V/P	(4) Implied Growth	(5) Relative Growth	(6) V/P (t-1)	(7) Implied Growth (t-1)	(8) Relative Growth (t-1)
1-Low	0.642	0.079	0.804	4.547%	1.434	0.788	5.465%	1.425
2	0.620	0.078	0.769	5.536%	1.347	0.714	5.337%	1.332
3	0.568	0.070	0.726	5.605%	1.303	0.788	5.268%	1.399
4	0.578	0.068	0.715	6.097%	1.280	0.761	5.718%	1.251
5-High	0.557	0.064	0.660	6.821%	1.253	0.713	5.967%	1.398
High-Low	-0.085*	-0.015**	-0.144***	2.274%***	-0.181***	-0.075	0.502%	-0.027

Table 3 ETF Flows and the Price-to-Fundamentals Relation: Regression Analyses

This table reports the results from estimating the following regressions: $(\frac{V}{P})_{i,t}\ or\ g\_Relative_{i,t} = b_0 + b_1 ETF\ Flow_{i,t} + b_2 Controls + \mu_t + \delta_i + \varepsilon_{i,t}$. Monthly stock-level ETF flows are calculated as the weighted average of the flows for all ETFs holding the stock, where the weight is the stock's portfolio weight in each ETF. Annual ETF flows are calculated by adding up monthly stock-level ETF flows for 12 months from 3 months after the previous fiscal year-end, scaled by the total dollar trading volume during the year. I standardize the V/P, relative growth, and ETF flows by subtracting the sample mean and dividing by the sample standard deviation. Firm and year fixed effects are included. Standard errors are double clustered at the firm and year level. t-statistics are presented in parentheses. ***, **, and * represent significance at the 0.01, 0.05 and 0.10 levels, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	V/P	V/P	Relative Growth	Relative Growth
ETF Flow	**-0.099*** **	**-0.082*** **	**-0.060*** **	**-0.050*** **
	(-7.18)	**(-5.57)**	**(-5.86)**	**(-4.88)**
Sales Growth		0.299***		0.090*
		(7.48)		(2.03)
Profit Margin		0.464***		0.425***
		(5.47)		(3.13)
Asset Turnover		-0.004		0.003
		(-1.36)		(0.83)
Analyst LT Forecast		0.689***		0.076
		(6.95)		(0.49)
F_Score		0.050***		0.011*
		(8.53)		(1.86)
Beta_Market		-0.164***		0.049
		(-3.46)		(0.64)
Beta_HML		-0.266*		0.254
		(-2.04)		(1.23)
Beta_SMB		0.218*		-0.529***
		(1.76)		(-3.31)
Beta_RMW		-0.014		0.050
		(-0.10)		(0.18)
Beta_CMA		-0.214		-0.190
		(-1.41)		(-0.91)
Size (t-1)		-0.313***		-0.001
		(-10.53)		(-0.05)
Book-to-Price (t-1)		-0.272***		-0.272***
		(-4.20)		(-3.99)
Ret(-3,-1)		0.053		0.049
		(1.47)		(1.48)
Ret(-6,-4)		-0.014		0.013
		(-0.20)		(0.25)
Ret(-9,-7)		-0.014		0.024

		(-0.49)		(0.78)
Ret(-12,-10)		0.010		0.052*
		(0.34)		(1.76)
V/P (t-1)		0.076*		
		(2.06)		
Relative Growth (t-1)				0.065***
				(3.05)
Observations	27,327	27,327	24,795	24,795
Fixed Effects	Firm&Year	Firm&Year	Firm&Year	Firm&Year
Adj. R-squared	0.608	0.702	0.199	0.231

Table 4 Cases with Negative V/P or Growth Higher than the Discount Rate

This table reports the analyses for observations with either negative V/P or the implied growth rate higher than the discount rate. Monthly stock-level ETF flows are calculated as the weighted average of the flows for all ETFs holding the stock, where the weight is the stock's portfolio weight in each ETF. Annual ETF flows are calculated by adding up monthly stock-level ETF flows for 12 months from 3 months after the previous fiscal year-end, scaled by the total dollar trading volume during the year. In Panel A, firms are sorted into 5 portfolios on the basis of the annual ETF flows, and the time-series averages of the value of speculative growth for each portfolio are reported. Panel B reports the results from estimating the following regression: $Speculative\ Growth_{i,t} = b_0 + b_1 ETF\ Flow_{i,t} + b_2 Controls + \mu_t + \delta_i + \varepsilon_{i,t}$. I standardize the value of speculative growth and ETF flows by subtracting the sample mean and dividing by the sample standard deviation. Firm and year fixed effects are included. Standard errors are double clustered at the firm and year level. t-statistics are presented in parentheses. ***, **, and * represent significance at the 0.01, 0.05 and 0.10 levels, respectively.

Panel A ETF Flows and Speculative Growth: Portfolio Analysis

	(1)	(2)
ETF Rank	Full Sample	Omitted Sample
1-Low	0.242	0.490
2	0.274	0.494
3	0.317	0.517
4	0.330	0.575
5-High	0.417	0.636
High-Low	0.175***	0.146***

Panel B ETF Flows and Speculative Growth: Regression Analyses

VARIABLES	(1) Full Sample	(2) Omitted Sample
ETF Flow	**0.046*** **	**0.033*** **
	(5.09)	**(3.60)**
Sales Growth	-0.036	-0.088**
	(-1.12)	(-2.34)
Profit Margin	-0.242***	-0.271***
	(-4.10)	(-4.24)
Asset Turnover	0.000	0.001
	(0.20)	(0.35)
Analyst LT Forecast	-0.418***	-0.203**
	(-7.16)	(-2.53)
F_Score	-0.030***	-0.018***
	(-8.44)	(-4.09)
Beta_Market	0.084**	0.015
	(2.16)	(0.48)
Beta_HML	0.096	0.028
	(1.14)	(0.89)
Beta_SMB	-0.070	-0.015
	(-0.87)	(-0.38)
Beta_RMW	-0.038	0.011
	(-0.71)	(0.42)
Beta_CMA	0.133*	0.018
	(1.78)	(0.64)
Size (t-1)	0.202***	-0.034
	(12.39)	(-1.56)
Book-to-Price (t-1)	0.153***	-0.108
	(4.30)	(-1.51)
Ret(-3,-1)	-0.036	-0.103**
	(-0.95)	(-2.69)
Ret(-6,-4)	0.064	-0.169***
	(0.54)	(-4.49)
Ret(-9,-7)	-0.014	-0.031
	(-0.29)	(-1.14)
Ret(-12,-10)	0.025	-0.019
	(1.13)	(-0.77)
Speculative Growth(t-1)	0.410***	0.112***
	(9.27)	(3.95)
Observations	29,590	5,865
Fixed Effects	Firm&Year	Firm&Year
Adj. R-squared	0.648	0.814

Table 5 BlackRock's Expansion into the ETF Business

This table reports the results for BlackRock's expansion into the ETF business setting. Column (1) and column (2) report the results from estimating the following regression: $ETF\ Flow_{i,t} = b_0 + b_1 TREAT \times POST + b_2 Other\ ETF\%_{\ i,t} + b_3 Controls + \mu_t + \delta_i + \varepsilon_{i,t}$, where *TREAT* equals 1 if the percentage of ownership by iShares is above the sample median before the event, and equals 0 for other stocks. $POST$ equals 1 for 2010, and 0 for 2009. Column (3) and column (4) report the results from estimating the following regression: $(\frac{V}{P})_{i,t}\ or\ g\_Relative_{i,t} = b_0 + b_1 TREAT \times POST + b_2 Other\ ETF\%_{\ i,t} + b_3 Controls + \mu_t + \delta_i + \varepsilon_{i,t}$ I standardize V/P, relative growth, and ETF flows by subtracting the sample mean and dividing by the sample standard deviation. Firm and year fixed effects are included. Standard errors are clustered at the firm level. t-statistics are presented in parentheses. ***, **, and * represent significance at the 0.01, 0.05 and 0.10 levels, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	ETF Flow	ETF Flow	V/P	Relative Growth
TREAT*POST	**0.281*** **	**0.197*** **	**-0.049** **	**-0.024* **
	(4.34)	**(4.21)**	**(-2.40)**	**(-1.91)**
Sales Growth		-0.002	0.018**	-0.005
		(-0.93)	(2.35)	(-0.57)
Profit Margin		-0.122	0.372***	0.118*
		(-1.33)	(3.31)	(1.84)
Asset Turnover		-0.009	0.000	0.000
		(-1.19)	(0.09)	(0.21)
Analyst LT Forecast		0.141	-0.350**	0.051**
		(0.55)	(-2.40)	(2.02)
F_Score		-0.021*	0.000	0.002
		(-1.85)	(0.08)	(0.84)
Beta_Market		-0.372***	0.120	-0.045
		(-4.96)	(1.00)	(-1.57)
Beta_HML		-0.259***	-0.009	0.026
		(-5.45)	(-0.13)	(1.04)
Beta_SMB		-0.072	-0.030	-0.018
		(-1.38)	(-0.40)	(-0.47)
Beta_RMW		-0.088***	-0.019	0.034*
		(-2.61)	(-0.51)	(1.65)
Beta_CMA		-0.103***	-0.010	0.012
		(-4.17)	(-0.21)	(0.96)
Size (t-1)		0.071	-0.022	0.001
		(1.51)	(-0.56)	(0.37)
Book-to-Price (t-1)		0.033	-0.070	0.004
		(1.50)	(-1.57)	(1.14)
Ret(-3,-1)		0.110	0.086	0.014
		(1.40)	(0.92)	(1.02)
Ret(-6,-4)		0.614***	0.097	-0.018
		(5.51)	(0.97)	(-1.33)

Ret(-9,-7)		-0.179*	-0.349***	-0.022
		(-1.73)	(-2.68)	(-1.14)
Ret(-12,-10)		0.141	-0.160	0.005
		(1.65)	(-1.59)	(0.43)
V/P (t-1)		-0.011	0.034	
		(-0.53)	(0.57)	
Relative Growth (t-1)		-0.000		0.013
		(-0.45)		(1.36)
Other ETF%		-12.996***	3.008	-0.371
		(-2.83)	(1.42)	(-1.41)
Observations	3,974	3,974	3,755	3,470
Fixed Effects	Firm&Year	Firm&Year	Firm&Year	Firm&Year
Adj. R-squared	0.519	0.594	0.906	0.356

Table 6 ETF Flows and Future Firm Performance

This table reports the results from estimating the following regression equation: $RNOA_{i,t+1}\ or\ Sales\ Growth_{i,t+1}\ INVST_{i,t+1} or\ Return_{i,t+1} = b_0 + b_1 ETF\ Flow_{i,t} + b_2 Controls + \mu_t + \delta_i + \varepsilon_{i,t}$, where $RNOA_{i,t+1}$ is return on net operating assets in the subsequent year, which is calculated as operating income divided by net operating assets. $Sales\ Growth_{i,t+1}$ is the sales growth rate in the subsequent year. $INVST_{i,t+1}$ is one of the following variables in the subsequent year: 1) CAPEX scaled by beginning-year total assets 2) R&D scaled by beginning-year total assets 3) SG&A scaled by beginning-year total assets. $Return_{i,t+1}$ is the buy-and-hold stock return in the subsequent year. I standardize RNOA, sales growth, investments, and ETF flows by subtracting the sample mean and dividing by the sample standard deviation in the entire sample. Firm and year fixed effects are included. Standard errors are double clustered at the firm and year level. *t*-statistics are presented in parentheses. ***, **, and * represent significance at the 0.01, 0.05 and 0.10 levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	RNOA(t+1)	Sales Growth(t+1)	CAPEX(t+1)	R&D(t+1)	SG&A(t+1)	Return(t+1)
ETF Flow	**-0.012***	**-0.006****	**0.007**	**-0.003**	**-0.010*****	**-0.037*****
	(-1.78)	**(-2.69)**	**(0.72)**	**(-0.67)**	**(-3.44)**	**(-3.70)**
Sales Growth	0.075	-0.035	0.066*	-0.015	-0.020	-0.198***
	(1.66)	(-1.33)	(1.90)	(-0.64)	(-1.50)	(-3.35)
Profit Margin	0.135**	-0.296*	0.079*	-0.079**	-0.037	0.176***
	(2.30)	(-1.84)	(2.02)	(-2.23)	(-1.55)	(3.38)
Asset Turnover	0.017**	0.001	0.006**	0.002	0.010***	0.002
	(2.57)	(1.01)	(2.94)	(0.64)	(3.42)	(0.35)
Analyst LT Forecast	0.197*	0.056***	0.361***	0.073	0.142**	-0.156
	(2.08)	(3.19)	(3.77)	(1.16)	(2.66)	(-0.84)
F_Score	0.028***	0.006	0.015**	0.003	0.018***	0.029***
	(7.41)	(0.89)	(2.60)	(1.43)	(7.04)	(3.82)
Beta_Market	0.053	-0.030	-0.157**	0.001	0.035	-0.208
	(1.46)	(-1.07)	(-2.29)	(0.04)	(1.51)	(-1.56)
Beta_HML	-0.013	-0.023	-0.094	-0.025**	-0.034**	0.042
	(-0.56)	(-0.91)	(-0.98)	(-2.16)	(-2.25)	(0.53)
Beta_SMB	0.005	0.062	0.053	-0.003	-0.002	0.137
	(0.11)	(1.22)	(0.60)	(-0.11)	(-0.07)	(1.09)
Beta_RMW	0.026	-0.016	0.065	0.068***	0.046**	-0.080
	(1.32)	(-1.36)	(1.22)	(3.89)	(2.83)	(-1.01)
Beta_CMA	0.017	0.024*	-0.037	-0.022	-0.018	0.041
	(0.91)	(1.86)	(-0.57)	(-1.46)	(-1.01)	(0.49)
Size (t-1)	-0.005	0.005	0.058***	-0.099***	-0.166***	-0.051*
	(-0.18)	(0.33)	(3.36)	(-6.38)	(-10.44)	(-1.77)
Book-to-Price (t-1)	-0.179***	-0.071**	-0.037**	-0.042***	-0.067***	0.173***
	(-5.76)	(-2.50)	(-2.15)	(-3.31)	(-5.00)	(2.65)
Ret(-3,-1)	0.225***	0.057	-0.012	0.011	0.054***	-0.329***
	(6.54)	(1.41)	(-0.38)	(0.68)	(3.39)	(-5.11)
Ret(-6,-4)	0.219***	0.001	0.176***	0.085***	0.107***	-0.608***
	(5.08)	(0.03)	(3.99)	(4.36)	(4.51)	(-5.20)

Ret(-9,-7)	0.004	0.063***	0.221***	-0.019	-0.006	0.080
	(0.22)	(4.26)	(4.60)	(-0.81)	(-0.45)	(1.47)
Ret(-12,-10)	0.025	0.038**	0.252***	0.008	-0.022	-0.041
	(1.13)	(2.57)	(5.28)	(0.36)	(-1.32)	(-0.79)
Observations	29,131	29,131	29,131	29,131	29,131	29,131
Fixed Effects	Firm&Year	Firm&Year	Firm&Year	Firm&Year	Firm&Year	Firm&Year
Adj. R-squared	0.309	0.243	0.719	0.891	0.870	0.111

Table 7 ETF Flows and Investors' Attention

This table reports the results from estimating the following regression equation: $Attention_{i,t} = b_0 + b_1 ETF\ Flow_{i,t} + b_2 Controls + \mu_t + \delta_i + \varepsilon_{i,t}$. I use the two variables to proxy for attention: (1) *#Questions* is the total number of questions being asked at conference calls during the year (2) *Question Length* is the total number of the words in questions being asked at conference calls during the year. I standardize *#Questions, Question Length*, and ETF flows by subtracting the sample mean and dividing by the sample standard deviation. Firm and year fixed effects are included. Standard errors are double clustered at the firm and year level. *t*-statistics are presented in parentheses. ***, **, and * represent significance at the 0.01, 0.05 and 0.10 levels, respectively.

VARIABLES	(1) #Questions	(2) Question Length
ETF Flow	**-0.050****	**-0.042***
	(-2.16)	**(-1.87)**
Sales Growth	0.017	-0.004
	(0.49)	(-0.51)
Profit Margin	0.288***	-0.040
	(3.01)	(-0.55)
Asset Turnover	-0.003	0.005
	(-0.45)	(0.81)
Analyst LT Forecast	1.711***	0.143
	(3.45)	(1.00)
F_Score	-0.027***	-0.008**
	(-3.07)	(-2.29)
Beta_Market	-0.521**	-0.015
	(-2.16)	(-0.22)
Beta_HML	0.274**	-0.015
	(2.89)	(-0.50)
Beta_SMB	0.086	0.060
	(0.46)	(1.22)
Beta_RMW	0.224**	-0.003
	(2.55)	(-0.12)
Beta_CMA	0.057	-0.012
	(0.56)	(-0.32)
Size (t-1)	-0.321***	0.155***
	(-4.58)	(4.57)
Book-to-Price (t-1)	-0.226**	-0.043**
	(-2.57)	(-2.28)
Ret(-3,-1)	-0.108	-0.004
	(-1.15)	(-0.08)
Ret(-6,-4)	-0.336**	-0.104*
	(-2.35)	(-2.00)
Ret(-9,-7)	-0.229**	-0.002
	(-2.57)	(-0.06)
Ret(-12,-10)	-0.418**	0.028
	(-2.13)	(0.58)
Observations	11,958	11,958
Fixed Effects	Firm&Year	Firm&Year
Adj. R-squared	0.375	0.513

Table 8 Cross-Sectional Analyses

Table 8 reports the results for cross-sectional analyses. Column (1) and column (2) compare stocks with more and fewer shareholders. Column (3) and column (4) compare stocks with high or low ownership by active investors, where active investors are non-index mutual funds with at least 60% active shares (Cremers and Petajusto 2009). Column (5) and column (6) compare stocks with high and low idiosyncratic volatility. Column (7) and column (8) compare stocks with high and low ownership by lendable ETFs. The breakpoints for each test are based on the sample median. I standardize V/P, relative growth, and ETF flows by subtracting the sample mean and dividing by the sample standard deviation in the entire sample. Firm and year fixed effects are included. Standard errors are double clustered at the firm and year level. *t*-statistics are presented in parentheses. ***, **, and * represent significance at the 0.01, 0.05 and 0.10 levels, respectively.

	Few SH=1 if the number of shareholders is below median		Low Active =1 if ownership by active investors is below median		High IVOL=1 if idiosyncratic return volatility is above median		Low Lend=1 if ownership by lendable ETFs is below median	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	V/P	Relative Growth	V/P	Relative Growth	V/P	Relative Growth	V/P	Relative Growth
ETF Flow*Few SH	**-0.046****	**-0.015***						
	(-2.55)	**(-1.72)**						
ETF Flow*Low Active			**-0.048*****	**-0.021***				
			(-3.04)	**(-1.69)**				
ETF Flow*High IVOL					**-0.034****	**-0.023****		
					(-2.41)	**(-2.33)**		
ETF Flow*Low Lend							**-0.023****	**-0.025****
							(-2.11)	**(-2.43)**
Few SH	0.008	-0.027						
	(0.14)	(-1.44)						
Low Active			-0.011	-0.026*				
			(-0.99)	(-1.73)				
High IVOL					-0.067***	-0.001		
					(-2.67)	(-0.03)		
Low Lend							-0.043	-0.067***
							(-1.63)	(-3.87)
ETF Flow	-0.032**	-0.042***	-0.043***	-0.037***	-0.040***	-0.049***	-0.067***	-0.048***
	(-2.27)	(-3.60)	(-3.70)	(-3.07)	(-3.61)	(-3.23)	(-3.77)	(-3.66)
Observations	25,418	23,534	27,327	24,795	26,213	24,065	27,327	24,795
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Firm&Year	Firm&Year	Firm&Year	Firm&Year	Firm&Year	Firm&Year	Firm&Year	Firm&Year
Adj. R-squared	0.708	0.235	0.716	0.236	0.716	0.243	0.708	0.250

Table 9 ETF Flow and Firm Share Issuances and Disclosures

This table reports the results for firm behavior in response to ETF flows. Panel A reports the results from estimating the following regressions: $SEO_{i,t+1}$ or $Adjusted\ Repurchases_{i,t+1}$ or $Insider\ Sales_{i,t+1} = b_0 + b_1 ETF\ Flow_{i,t} + b_2 Controls + \mu_t + \delta_i + \varepsilon_{i,t}$. $SEO_{i,t+1}$ equals one if the firm has an SEO during the subsequent year, and zero otherwise. $Adjusted\ Repurchases_{i,t+1}$ equals the total amount for share repurchases after removing the effect of share issuances and the potential effect of repurchases for exercised stock options, scaled by beginning-year total assets. $Insider\ Sales_{i,t+1}$ is the ratio of shares sold by all insiders to the sum of shares sold and purchased by all insiders during the subsequent year. Panel B reports the results from estimating the following regression: $\#Guidance_{i,t+1}$ or $\#WalkUp_{i,t+1}$ or $\#WalkDown_{i,t+1} = b_0 + b_1 ETF\_Flow_{i,t} + b_2 Controls + \mu_t + \delta_i + \varepsilon_{i,t}$. $\#Guidance_{i,t+1}$ is the number of long-horizon management earnings forecasts during the subsequent year. $\#WalkUp_{i,t+1}$ and $\#WalkDown_{i,t+1}$ are the numbers of long-horizon, walk-up and walk-down earnings forecasts issued during subsequent year, respectively. Firm and year fixed effects are included. Standard errors are double clustered at the firm and year level. *t*-statistics are presented in parentheses. ***, **, and * represent significance at the 0.01, 0.05 and 0.10 levels, respectively.

Panel A ETF Flows, Share Issuances, Repurchases, and Insider Sales

	(1)	(2)	(3)
VARIABLES	SEO(t+1)	Adjusted Repurchases (t+1)	Insider Sales(t+1)
ETF Flow	**0.005***	**-0.029***	**0.034*****
	(1.85)	**(-1.89)**	**(2.99)**
Sales Growth	-0.000	-0.026	0.002**
	(-0.30)	(-1.10)	(2.27)
Profit Margin	-0.030**	0.256**	0.043
	(-2.71)	(2.47)	(0.93)
Asset Turnover	-0.001	0.019**	0.000
	(-1.09)	(2.21)	(0.05)
Analyst LT Forecast	-0.022	-0.592*	0.006
	(-0.70)	(-2.10)	(0.05)
F_Score	-0.000	0.061***	0.006
	(-0.22)	(7.26)	(0.94)
Beta_Market	0.012	-0.086	0.113
	(0.78)	(-0.85)	(1.74)
Beta_HML	0.008	0.041	0.109***
	(1.24)	(0.77)	(3.27)
Beta_SMB	-0.003	0.427***	0.107*
	(-0.16)	(3.56)	(1.88)
Beta_RMW	0.000	0.033	-0.030
	(0.02)	(0.47)	(-0.80)
Beta_CMA	-0.006	0.303**	-0.004
	(-0.97)	(2.56)	(-0.17)
Size (t-1)	-0.024***	0.361***	0.074***
	(-4.58)	(9.95)	(5.99)
Book-to-Price (t-1)	-0.011*	0.064**	-0.176***
	(-1.76)	(2.35)	(-4.41)

Ret(-3,-1)	0.018**	-0.005	0.174***
	(2.47)	(-0.10)	(4.57)
Ret(-6,-4)	0.020**	-0.095*	0.238***
	(2.32)	(-1.84)	(7.46)
Ret(-9,-7)	0.006	0.098*	0.237***
	(0.77)	(1.89)	(5.24)
Ret(-12,-10)	-0.011	0.179**	0.148***
	(-1.36)	(2.68)	(6.71)
Observations	29,590	25,704	29,590
Fixed Effects	Firm&Year	Firm&Year	Firm&Year
Adj. R-squared	0.315	0.314	0.724

Panel B ETF Flows and Firm Disclosures

VARIABLES	(1) #Guidance(t+1)	(2) #Walkup(t+1)	(3) #Walkdown(t+1)
ETF Flow	**-0.012*****	**-0.013***	**-0.007**
	(-2.59)	**(-1.84)**	**(-0.97)**
Sales Growth	-0.000	-0.009	0.007
	(-1.13)	(-0.68)	(0.64)
Profit Margin	0.107***	0.028	0.276***
	(3.94)	(0.45)	(4.04)
Asset Turnover	0.003	0.001	0.003
	(1.15)	(0.17)	(1.04)
Analyst LT Forecast	0.293***	0.091	0.494***
	(3.16)	(0.57)	(2.99)
F_Score	0.012***	0.014	0.008
	(2.70)	(1.62)	(1.28)
Beta_Market	0.021	0.106	0.135*
	(0.56)	(1.24)	(1.77)
Beta_HML	0.014	-0.027	-0.006
	(0.52)	(-0.41)	(-0.14)
Beta_SMB	-0.077*	-0.116*	-0.057
	(-1.90)	(-1.87)	(-0.88)
Beta_RMW	-0.021	0.070	-0.129***
	(-0.85)	(1.43)	(-3.08)
Beta_CMA	0.002	-0.005	-0.016
	(0.13)	(-0.13)	(-0.67)
Size (t-1)	0.059***	-0.022	0.098***
	(4.15)	(-1.01)	(5.24)
Book-to-Price (t-1)	0.005	-0.032	-0.007
	(0.47)	(-0.76)	(-0.59)
Ret(-3,-1)	0.069***	0.110***	0.074**
	(3.92)	(3.96)	(2.56)
Ret(-6,-4)	0.094**	-0.127**	0.223***
	(2.05)	(-2.23)	(5.20)
Ret(-9,-7)	0.100***	0.082*	0.104***
	(4.77)	(2.01)	(3.27)
Ret(-12,-10)	0.055**	-0.027	0.131***
	(2.31)	(-0.70)	(3.94)
Observations	29,590	29,590	29,590
Fixed Effects	Firm&Year	Firm&Year	Firm&Year
Adj. R-squared	0.590	0.404	0.327

Table 10 Cross-section of Stock Returns

Table 10 reports the results for the cross-sectional return analysis. At the *beginning* of every calendar month, stocks are ranked in ascending order based on the last available 12-month ETF flows. The time-series averages of the weighted-average portfolio raw returns are reported in column (1) and Fama-French five-factor alphas are reported in column (2). Also reported are the zero-cost portfolio returns buying the top 20% stocks and selling short the bottom 20% stocks. *t*-statistics with Newey-West correction for autocorrelation are reported in parentheses. ***, **, and * represent significance at the 0.01, 0.05 and 0.10 levels, respectively.

	(1)	(2)
ETF Flow Rank	Raw Return	Fama French 5-factor Alpha
1-Low	1.392%	0.408%
2	1.275%	0.130%
3	1.297%	0.296%
4	0.927%	-0.140%
5-High	0.643%	-0.301%
High minus Low	-0.749%**	-0.709%**

Table 11 Mutual Fund Flows

This table reports the results from estimating the following regressions: $(\frac{V}{P})_{i,t} or\ g\_Relative_{i,t} = b_0 + b_1 ETF\ Flow_{i,t} + b_2 NonIndex\ MF\ Flow_{i,t} + b_3 Index\ MF\ Flow_{i,t} + b_4 Controls + \mu_t + \delta_i + \varepsilon_{i,t}$. Non-index and index mutual fund flows at the fund level are calculated as: $TNA_{j,\,t} - TNA_{j,t-1} * (1 + R_{j,t})$. I standardize the V/P, relative growth, ETF flows, non-index mutual fund flows, and index mutual fund flows by subtracting the sample mean and dividing by the sample standard deviation. Firm and year fixed effects are included. Standard errors are double clustered at the firm and year level. *t*-statistics are presented in parentheses. ***, **, and * represent significance at the 0.01, 0.05 and 0.10 levels, respectively.

	(1)	(2)
VARIABLES	V/P	Relative Growth
ETF Flow	**-0.072*** **	**-0.049*** **
	(-4.78)	**(-4.54)**
Non-index MF Flow	0.005	0.009
	(0.85)	(1.08)
Index MF Flow	-0.014*	-0.016*
	(-1.89)	(-1.80)
Observations	27,327	24,795
Control Variables	Yes	Yes
Fixed Effects	Firm&Year	Firm&Year
Adj. R-squared	0.720	0.244