# ETFs, Arbitrage, and Contagion

Itzhak Ben-David

*Fisher College of Business, The Ohio State University*

Francesco Franzoni

*Swiss Finance Institute and the University of Lugano*

Rabih Moussawi

*Wharton Research Data Services, The Wharton School, University of Pennsylvania*

May 2012

**Abstract**

We study arbitrage activity between Exchange Traded Funds (ETFs)—an asset class that has gained paramount importance in recent years—and their underlying securities. We show that shocks to ETF prices are passed down to the underlying securities via the arbitrage between the ETF and the assets it tracks. As a result, the presence of ETFs increases the volatility of the underlying securities. Also, we find evidence consistent with the conjecture that ETFs contributed to shock propagation between the futures market and the equity market during the Flash Crash on May 6, 2010. Overall, our results suggest that arbitrage activity may induce contagion and that High Frequency Trading adds noise to market prices and can pose a threat to market stability.

---

We thank George Aragon, Vincent Fardeau, Thierry Foucault, Robin Greenwood, Augustin Landier, Alberto J. Plazzi, Scott Richardson, Dimitri Vayanos, and participants at seminars at SAC Capital Advisors, University of Lugano, University of Verona, and at the 4th Paris Hedge Funds Conference for helpful comments and suggestions.

# 1 Introduction

With $1.4 trillion of assets under management globally (June 2011), Exchange Traded Funds (ETFs) are rising steadily among the big players in the asset management industry. This asset class is also capturing an increasing share of transactions in financial markets. For example, in August 2010, ETFs (along with other exchange traded products) represented about 40% of all trading volume in U.S. markets (Blackrock 2011). This explosive growth has attracted regulators' attention with regard to the hidden risks to which ETF investors are exposed and the threat that ETFs pose to market stability.[1] So far, however, no systematic attention has been devoted to understanding the role of ETFs in shock propagation.

In this paper, we investigate whether the arbitrage activity that takes place between ETFs and their underlying securities can create a new channel of shock transmission. In other words, we study whether the presence of ETFs can foster contagion in financial markets. Our question is inspired by a growing literature arguing that institutional trading can significantly affect the first and second moments of asset returns when arbitrage is limited (see Gromb and Vayanos 2010 for a survey). When traders are subject to limits of arbitrage, a liquidity shock to one asset can be propagated to another asset. Contagion can occur via a number of different channels including portfolio rebalancing by risk averse arbitrageurs (e.g., Greenwood 2005), wealth effects (e.g., Kyle and Xiong 2001), and liquidity spillovers (e.g., Cespa and Foucault 2012). Our empirical analysis studies the role of arbitrageurs in propagating liquidity shocks between two assets with the same fundamental value, such as the ETF and the basket of its components.

In an efficient market, the price of an ETF should equal the price of its underlying portfolio as the two assets have the same fundamental value. The fact that new shares of ETFs can be created and redeemed almost continuously facilitates arbitrage so that, on average, the ETF price cannot diverge consistently and substantially from its net asset value (NAV). However, the popularity of ETFs among retail and institutional investors for speculative and hedging purposes makes them increasingly exposed to non-fundamental demand shocks. If

[1] In more detail, the risks to investors from ETFs relate to their potential illiquidity, which manifested during the Flash Crash of May 6, 2010, when 65% of cancelled trades were for ETFs. Also important, regulators have taken into consideration the potential for counterparty risk which seems to be operating in case of both synthetic replication (as the swap counterparty may fail to deliver the index return) and physical replication (as the basket securities are often loaned out). Also, concerns have been expressed that a run on ETFs may endanger the stability of the financial system (Ramaswamy 2011).

arbitrage is limited these shocks can propagate from the ETF market to the underlying securities. As an example of the effect that we conjecture, consider a large sell order for an ETF originating from an institutional trader for liquidity purposes. As in Greenwood's (2005) model, risk averse arbitrageurs will buy the ETF and hedge this position by selling the underlying portfolio. The selling activity can lead to downward price pressure on the latter asset. As a result, the initial liquidity shock is propagated to the NAV, which falls without a fundamental reason. In this sequence of events, arbitrageurs' activity induces contagion of liquidity shocks. The alternative hypothesis to this conjecture is that arbitrage is perfect, in which case a non-fundamental shock to the ETF would be fully absorbed with no significant price impact on the NAV.

Establishing this channel of contagion is especially relevant in the current financial environment. The explosion of new financial products has introduced arbitrage relations between the newly created assets and existing securities. If arbitrage activity creates contagion, one can expect that liquidity shocks that originate in the markets for the new products could be transmitted to the related assets. Ultimately, this channel can cause an increase in non-fundamental volatility in financial markets. This seems like an unintended consequence of arbitrage and a yet-unexplored outcome of financial innovation.

Our study has three parts. In the first part, we present evidence of limits of arbitrage between ETFs and their underlying securities. In particular, we show that the discrepancy between the ETF price and the NAV, which we label 'ETF mispricing,' increases when arbitrageurs' capital becomes scarcer and when trading costs increase. In the time series, we show that mispricing is stronger following periods of high volatility, which is consistent with the results in Nagel (2011) on the positive link between market volatility and the profitability of liquidity provision. Also, ETF mispricing is greater following periods of poor stock market returns and poor returns for the financial sector. In line with Hameed, Kang, and Viswanthan (2011), these results suggest that mispricing is larger following times in which arbitrageurs are more constrained. Finally, we compute a profitability measure of ETF-arbitrage and show that following arbitrageurs' losses the aggregate mispricing widens. In the cross-section, we find that mispricing is larger for ETFs with high bid-ask spread and following arbitrageurs' losses in that ETF.

The second part of the paper has the main results on the impact of ETF arbitrage on the underlying assets. We find evidence that arbitrage trades facilitate the propagation of liquidity shocks from the ETFs to the underlying securities.[2] We begin with the observation that the underlying assets move in the direction of the lagged mispricing, while the ETF moves in the opposite direction of the mispricing. To distinguish the arbitrage trading channel from a price discovery channel (e.g., information hits the ETF first and only later the underlying stocks), we pursue additional procedures. First, we use a Vector Auto-Regression (VAR) analysis to show that a shock to ETF mispricing leads to a move of the underlying assets and a later reversal, which is consistent with a temporary frictional shock rather than information-based change in prices. Second, we identify non-fundamental shocks by singling out days in which order imbalance in the ETF by far outweighs the order imbalance in the underlying securities, consistent with demand shocks hitting only the ETF market.

Next, we explore the arbitrage channel of shock transmission. We find that mispricing generates two types of arbitrage activities between ETFs and the underlying assets. Specifically, we document that ETF shares are created as a response to lagged mispricing. Also, we find evidence for buying and selling pressure in the ETFs and underlying assets in the direction that closes the mispricing. Furthermore, we show that stock returns react differentially to ETF mispricing in accordance to theories about demand by arbitrageurs; e.g., returns of large and high beta stocks and of stocks with low idiosyncratic volatility react more strongly to ETF mispricing. These findings provide evidence for the cross-market arbitrage channel versus other, non-trading based, theories of contagion.

Another implication of shock propagation from ETFs to the underlying securities is that the volatility of the underlying securities is expected to increase when ETF ownership increases. The underlying securities are exposed to their own fundamental and non-fundamental shocks. Once ETF ownership increases, they also inherit the non-fundamental shocks from the ETF market. Consistent with this conjecture, we show the average volatility of individual stocks increases following the increase in ETF ownership. We estimate that median holding of ETFs in late 2010 caused daily stock volatility to increase by 13 basis points, a 3.4% increase. For the

[2] Other channels of shock propagation that are not explored in this paper include illiquidity contagion, as proposed in Cespa and Foucault (2012). In their model, market makers in one asset class (e.g., the underlying securities) extract signals from prices in a second market (e.g., ETFs). Hence, shocks in one market lead to price movements in the other market, even without cross-market arbitrage.

$90^{th}$ percentile of ETF ownership, the increase in daily volatility due to ETF ownership was 24 basis points, a 6.3% increase. The effect is more pronounced in small stocks, where arbitrage trading activity is expected to have greater price impact due to reduced liquidity. Furthermore, we predict that the volatility of the underlying securities should increase as a result of ETF introduction. Indeed, the volatility of the underlying securities increases following the introduction of ETFs on the S&P 500. We take these findings as further evidence that ETFs operate as a conduit of shocks to the underlying securities.

In the third and last part of the paper, we provide novel evidence suggesting that the contagion mechanism that we describe was at work during the Flash Crash of May 6, 2010. On that day, the S&P 500 declined dramatically in value as a result of a negative demand shock originated in the S&P 500 E-mini futures market (see the CFTC and SEC 2010 preliminary and final reports). The anecdotal evidence reports cross-market arbitrage between the futures and the ETFs tracking the index. After the decline of the futures prices, cross-market arbitrageurs sold index-tracking ETFs and bought futures, driving down the ETF prices. We conjecture and find consistent evidence that arbitrage between the ETFs and the underlying stocks contributed to propagate the initial shock to the spot market for stocks. During the downward move in the market, the ETF discount is a significant predictor of the negative return on the S&P 500 in the following second, controlling for returns of the futures contract.

The main message of the study is that arbitrage activity between ETFs and the underlying assets has the potential to propagate liquidity shocks. These findings imply that ETFs increase the risk of contagion across asset classes, and especially so for less liquid securities. More broadly, our results potentially extend to all situations in which assets are linked by arbitrage relations. For this reason, the paper raises the topical question on the extent to which the development of derivative markets has caused an increase in volatility. The results are also relevant in the current debate on the impact of High Frequency Trading (HFT) on market stability. As ETF arbitrage is the turf of high frequency traders, the evidence in the paper suggests that HFT impacts the volatility of asset prices and, in extreme cases such as the Flash Crash, can pose a serious threat to market stability.

A growing literature studies the impact of institutional trades on returns. On the theory side, asset pricing models have been developed that explicitly incorporate the impact of institutions on asset prices (e.g., Basak and Pavlova 2011, Vayanos and Wooley 2011). Empirically, there is mounting evidence on the effect of institutional investors on expected returns (Shleifer 1986, Greenwood 2005, Barberis, Shleifer, and Wurgler 2005, Coval and Stafford 2007, and Wurgler 2011 for a survey) and correlations (Anton and Polk 2011, Cella, Ellul, and Giannetti 2011, Chang and Hong 2011, Greenwood and Thesmar 2011, Lou 2011, and Jotikasthira, Lundblad, and Ramadorai 2012). Most closely related to our work, Greenwood (2005) develops a model for the price impact of arbitrageurs' response to a liquidity shock when arbitrage is limited and finds consistent evidence. Also related in terms of showing limited arbitrage, Hau, Massa, and Peress (2010) find that a demand shock following from a global stock index redefinition impacts both the prices of the stocks in the index and the currencies of the countries in which these stocks trade.

A few studies have dealt with potentially destabilizing effects of ETFs. Trainor (2010) investigates whether the daily rebalancing of leveraged ETFs can increase stock volatility and gives a negative answer. Bradley and Litan (2010) have voiced concerns that ETFs may drain the liquidity of already illiquid stocks and commodities, especially if a short squeeze occurs and ETF sponsors rush to create new ETF shares.

The paper proceeds as follows. Section 2 provides background information about arbitrage activity in ETFs and develops our hypothesis. Section 3 describes the data used in the study. Section 4 studies ETF mispricing and relates it to the limits of arbitrage. Section 5 shows evidence that arbitrage activity in ETFs can propagate shocks in the stock market. Section 6 focuses on the role of ETFs in shock propagation during the Flash Crash. Section 7 concludes.

## 2 ETF Arbitrage and Hypotheses Development

### 2.1 Mechanics of Arbitrage

Exchange Traded Funds are investment companies that typically focus on one asset class, industry, or geographical area. Most ETFs track an index, very much like passive index funds. ETFs were first introduced in the late 1980s, but became more popular with the issuance in January 1993 of the SPDR (known as "Spider," or Standard & Poor's Depository Receipts), which is an ETF that tracks the S&P 500 (which we label "SPY" from its ticker). In 1995, another SPDR, the S&P MidCap 400 Index (MDY) was introduced, and since then the number exploded to more than 1,000 ETFs by the end of 2011. Other popular ETFs are the DIA which tracks the Dow Jones Industrials Average and QQQQ which tracks the Nasdaq-100. Since 2008, the Securities and Exchange Commission (SEC) allows actively-managed ETFs.

Similar to closed-end funds, retail and institutional investors can trade ETF shares in the secondary market.[3] However, unlike closed-end funds, new ETFs shares can be created and redeemed. Since the price of ETF shares is determined by the demand and supply in the secondary market, it may diverge from the value of the underlying securities (the NAV). Some institutional investors (called "authorized participants," AP), which are typically market makers or specialists, can trade bundles of ETF shares (called "creation units," typically 50,000 shares) with the ETF sponsor. An AP can create new ETF shares by transferring the securities underlying the ETF to the ETF sponsor. Symmetrically, the AP can redeem ETF shares and receive the underlying securities in exchange. For some funds creations and redemptions of ETF shares can also happen in cash.[4]

To illustrate the arbitrage process, we focus on the two cases of (i) ETF premium (the price of the ETF exceeds the NAV) and (ii) ETF discount (the ETF price is below the NAV). In the case of an ETF premium, APs have an incentive to buy the underlying securities, submit them to the ETF sponsor, and ask for newly created ETF shares in exchange. Then, the AP sells the new supply of ETF shares on the secondary market. This process generates a decline in the ETF price and, potentially, an increase in the NAV, reducing the premium. In the case of an ETF discount, APs buy ETF units in the market and redeem them for the basket of underlying securities from the ETF sponsor. Then, the APs can sell the securities in the market.[5] This generates positive price pressure for the ETF and, possibly, negative pressure on the NAV, which reduces the discount.

Arbitrage can be undertaken by market participants who are not APs. Since both the underlying securities and ETFs are traded, investors can buy the inexpensive asset and short sell

[3] Barnhart and Rosenstein (2010) examine the effects of ETF introductions on the discount of closed-end funds and conclude that market participants treat ETFs as substitutes to closed-end funds.
[4] Creation and redemption for cash is especially common in ETFs on foreign assets, and where assets are illiquid, e.g., fixed income ETFs.
[5] See http://ftalphaville.ft.com/blog/2009/03/12/53509/the-curious-case-of-etf-nav-deviations/ for a description of trading strategies by APs.

the more expensive one.[6] For example, in case of an ETF premium, traders buy the underlying securities and short sell the ETF. They hold the positions until prices converge, at which point they cover their long and short positions to realize the arbitrage profit. Conversely, in the case of ETF discount, traders buy the ETF, and short sell the individual securities. ETF prices can also be arbitraged against other ETFs (see Marshall, Nguyen, and Visaltanachoti 2010), or against futures contracts (see Richie, Daigler, and Gleason 2008).[7] The latter case is relevant in our discussion of the Flash Crash, where we argue that the price drop in the E-mini futures on the S&P 500 was propagated to the ETFs on the same index via cross-market arbitrage. Given the fleeting nature of profit opportunities in this line of business, ETF arbitrage is carried out mostly at high frequencies by hedge funds doing statistical arbitrage.[8]

To be precise, although these trading strategies involve claims on the same cash flows, they are not *sensu stricto* arbitrages as they are not risk free. In particular, market frictions might induce noise into the process. For example, execution may not be immediate, or shares may not be available for short selling, or mispricing can persist for longer than expected. In the remainder of the paper, we will talk about ETF arbitrage implying the broader definition of 'risky arbitrage.'

## 2.2 Hypothesis Development

We conjecture that the arbitrage between ETFs and the securities in their baskets can propagate a liquidity shock from the ETF market to the prices of these securities. To exemplify the transmission mechanism that we have in mind, let us start from a situation in which the ETF price and the NAV are aligned at the level of the fundamental value of the underlying securities, as in Figure 1a. Then, we imagine a non-fundamental shock, such as an exogenous increase in demand, hitting the ETF market. This could happen, for example, if some large institution receives inflows and scales up its existing ETF allocation. This puts positive pressure on the ETF

[6] See http://www.indexuniverse.com/publications/journalofindexes/joi-articles/4036-the-etf-index-pricing-relationship.html for a description of trading strategies that eliminate mispricing between ETFs and their underlying securities.
[7] See http://seekingalpha.com/article/68064-arbitrage-opportunities-with-oil-etfs for a discussion of a trading strategy to exploit a mispricing between oil ETFs and oil futures.
[8] See, e.g., http://ftalphaville.ft.com/blog/2009/07/30/64451/statistical-arbitrage-and-the-big-retail-etf-con/ and http://ftalphaville.ft.com/blog/2011/06/06/584876/manufacturing-arbitrage-with-etfs/.

price (Figure 1b). At this point, cross-market arbitrageurs step in betting on the re-establishment of equilibrium between the ETF and the NAV. Arbitrageurs go long the ETF and short the securities in the ETF basket to hedge their position. As in Greenwood's (2005) model, we postulate that arbitrageurs have limited risk bearing capacity. In this framework, not only does the arbitrage trade impact the price of the ETF, but also it puts upward pressure on the prices of the basket components, as in Figure 1c. Eventually, liquidity flows back into both markets and prices revert back to the initial situation (Figure 1d).

The alternative hypothesis to our conjecture is that arbitrage is not limited, in which case the arbitrage trades do not move the NAV. This could happen, for example, if in this market operated liquidity providers who have full information about the fundamental value, or unlimited risk bearing capacity. As a necessary condition to separate our conjecture from this alternative hypothesis, we need to show that a shock to the ETF price is followed by movements in the NAV. A shock that only occurs in the ETF market widens the gap between the ETF price and the NAV, which we label ETF mispricing. So, in the empirical analysis, we test whether the ETF mispricing predicts subsequent movements in the NAV in the direction that closes this gap.

We argue that this is just a necessary condition to prove the propagation of liquidity shocks by arbitrage. It is not sufficient because similar predictability would emerge in case the initial shock was a fundamental shock. This situation is illustrated in Figure 2. The initial equilibrium (Figure 2a) is perturbed by a shock to the fundamental value of the ETF components (Figure 2b). If the ETF market is more liquid, it is possible that price discovery takes place in this market. So, the ETF price moves first (Figure 2c) and the prices of the underlying securities move with a delay (Figure 2d). Given this alternative, we need to provide further evidence that the predictability of the NAV by mispricing follows, at least in part, from an initial non-fundamental shock. We exploit two elements to accomplish this task. First, we look for demand pressure in the ETF that is not matched by comparable demand pressure in the underlying securities. Second, we test for the prediction that the

We also wish to separate our arbitrage-based story from explanations in which shock propagation occurs without cross-market arbitrageurs. For example, Cespa and Foucault (2012) have a model in which dealers in one market learn about fundamental value from the realization of prices in a related market. This mechanism can generate contagion even without the type of relative-value trading that we have in mind. So, to buttress our interpretation, we need to provide evidence that mispricing actually generates trading volume in the direction that is consistent with the re-establishment of the no-arbitrage relation. All this analysis is carried out in Section 5.

# 3 Data

## 3.1 Data Sources

We use CRSP, Compustat, and OptionMetrics to identify ETFs traded on the major US exchanges and to extract returns, prices, and shares outstanding. To identify ETFs, we first draw information from CRSP on all the 1,261 securities that have the historical share code of 73, which defines exclusively ETFs in the CRSP universe. We then merge these data with the ETFs that we could extract from Compustat XpressFeed price and OptionMetrics data, where we screen all US traded securities that can be identified as ETFs using the security type variables.[9] Compustat shares outstanding data are sparse before 2000, so we fill the gaps in the daily shares outstanding data using OptionMetrics total shares outstanding figures, which are available from 1996. OptionMetrics is then used to complement the ETF series and extract daily-level shares outstanding. Total shares outstanding allow us to compute the daily market capitalization of each ETF.[10]

Net Asset Value (NAV), as well as fund styles (objectives) and other characteristics are extracted from Lipper and Morningstar databases. This information starts being available in September 1998. We compute ETF mispricing as the difference between the ETF share price and the NAV of ETF portfolio at day close. Mispricing is expressed as a fraction of ETF price. Daily NAV returns are computed from daily NAVs. Since some ETFs are traded until 4:15pm (Engle and Sarkar 2006) while the major U.S. stock markets close at 4:00pm, we calculate the mispricing using 4:00 pm ETF prices drawn from TAQ, as the last trade in the ETF at or before

[9] Note that CRSP-Compustat Merged product does not have correct links between ETF securities in CRSP and Compustat universes. For this reason, we use historical CUSIP and ticker information to map securities in CRSP, Compustat, and OptionMetrics databases.

[10] We use short sale data from Compustat. We notice that short selling of ETFs is prevalent by hedge funds and other sophisticated investors as part of their hedging and market timing bets (see http://www.marketwatch.com/story/short-interest-in-etfs-down for example, when the iShares Lehman 20+ Year Treasury Bond Fund (TLT) had a whopping 235 percent of shares outstanding in short interest as of October 2004. The Short interest ratio for TLT was 15,669,711, while the total shares outstanding for this ETF were 4,000,000). Note that "ETFs, unlike regular shares, are exempt from the up-tick rule, so some investors use them for long/short and hedging strategies."

4:00 pm. Furthermore, starting with Table 5, we restrict our analysis to U.S. equity funds for which we are certain that the underlying stocks are traded in parallel to the trading of ETFs in the U.S.

Thomson-Reuters Mutual Fund holdings database allows us to construct the ETF holdings for each stock at the end of every month. ETFs are subject to Investment Company Act reporting requirements and, similar to mutual funds, they have to disclose their portfolio holdings at the end of each fiscal quarter. We use these data to align ETF ownerships every month using the most recently reported holdings. Then, for every stock, we sum the total ownership by various ETFs to construct our ETF holdings measure.

In our analysis of the SPDR ETF on the S&P 500 (SPY) on May 6, 2010, we construct our intraday return measures using TAQ data. We compute the volume weighted average price every second using the price and size for every trade that shows up in TAQ within each second. We then compute the NAV returns by aggregating the returns of the underlying stocks using their weights in the ETF portfolios as disclosed in the prior month-end reports. S&P 500 index intraday returns are constructed using the market capitalization of each constituent as weights. Order imbalances are computed for the individual ETFs and underlying stocks, after classifying trades into buyer- or seller-initiated transactions following Lee and Ready (1991) algorithm. The intraday prices on May 6, 2010, of the E-mini S&P 500 futures are obtained directly from the Chicago Mercantile Exchange (CME).

### 3.2 Descriptive Statistics

Our final sample consists of 1,146 distinct ETFs, with 1,065,832 daily observations with complete data from September 2, 1998 to March 31, 2011. Figure 1 illustrates the growth of ETFs over our sample period. At the start, the sample contains 20 ETFs, while at the end there are 986 ETFs with complete data. Table 1, Panel A, gives information on the growth of the assets in the ETF sector, showing that the average assets under management (AUM) in U.S. ETFs have grown from $9 billion in 29 ETFs during 1998 to over $1 trillion in 986 ETFs in March 2011. ETF growth in terms of assets and number of ETFs has picked up sharply after 2004. Panel B of Table 1, breaks down the ETFs in March 2011 by their Lipper objective code (for categories with more than $1 billion of AUM). The largest category by AUM contains the ETFs that track the S&P 500 with $95.6 billion in AUM and four ETFs, among which is the SPY that we study in the Flash-Crash analysis. The last column shows the fund objectives that have been included in the Equity ETF group in some of the regressions. From this group, we also exclude leveraged or short equity ETFs with the purpose of focusing on plain-vanilla equity ETFs.

Table 2 reports summary statistics for the variables that are used in the regressions. We defer a description of these variables until we use them in the analysis.

## 4 ETF Mispricing and the Limits of Arbitrage

In this section, we wish to provide the background for testing the hypothesis that arbitrage between the ETFs and the underlying basket can propagate shocks. A necessary condition for our argument is that ETF mispricing exists which would be evidence of limits of arbitrage. To this purpose, we quantify the extent of ETF mispricing and its relation to various measures of the limits of arbitrage in terms of the scarcity of speculators' capital and trading costs.

We note that previous studies document ETF mispricing. Engle and Sarkar (2006), Marshall, Nguyen, and Visaltanachoti (2010), and Petajisto (2011) report that ETFs exhibit mispricing with respect to the underlying securities and that these discrepancies can be exploited in profitable trading strategies. Our evidence complements the results from these other studies in showing that mispricing widens when limits of arbitrage become more binding.

### 4.1 Time Series of ETF Mispricing

In Figure 2a, we plot the daily percentage mispricing for the SPY, the ETF tracking the S&P 500. The mispricing is defined as the ETF price minus the NAV divided by the ETF price. All these variables are measured at the day close. The SPY is the largest equity ETF, with a market capitalization of $91.0 billion in December 2010. The figure shows that the average mispricing shrank over time. This was possibly the result of the ETF market becoming more liquid, which reduced transaction costs for ETF arbitrage. There are multiple episodes in which mispricing was sizeable. In particular, mispricing is larger during periods of market stress such the summer of 2007, and the fall of 2008, around the Lehman events. As an example, mispricing was 1% on October 22, 2008, and it was -1.2% on October 27, 2008. Note, further, that at times of high mispricing, the deviations from the NAV are both positive and negative, suggesting that the sign of the mispricing is less interesting than the magnitude of the mispricing as an indicator of limits of arbitrage. Overall, based on this graphical inspection, deviations from fundamental prices appear to be related to the overall liquidity in the market, which suggests a twofold interpretation. First, low market liquidity limits the profitability of ETF arbitrage due to the high transaction costs (see also Figure 2d). Second, low market liquidity can be a symptom of low funding liquidity (Brunnermeier and Pedersen 2009). In turn, a drop in funding liquidity implies that a reduced amount of capital is committed to ETF arbitrage allowing for a larger mispricing to persist.

In Figure 2b, we explore the evolution in the dispersion of mispricing for our entire sample of ETFs.[11] The chosen measure of dispersion is the interquartile range of mispricing across the ETFs. Consistent with figure 2a, the dispersion of mispricing has a general downward trend, yet ETF mispricing increases across the board during periods of market stress (e.g., late 2002, summer 2007, early 2008, fall 2008, May 2010 (Flash Crash)).

Another interesting measure of mispricing is the net mispricing. We define net mispricing as the difference between the absolute value of the percentage mispricing and the percentage bid-ask spread for the ETF at the day close. This variable approximates the extent to which arbitraging the mispricing for a given ETF-day is profitable after transaction costs. In Figure 2c we report the fraction of ETFs with positive net mispricing in a given day. The figure shows that as the ETF industry expands the fraction of mispriced ETFs increases. A likely explanation for this time-series relation is that bid-ask spreads shrink as the market becomes more familiar with ETFs and competition increases. As a consequence, a greater fraction of ETFs displays an absolute value of mispricing lying outside the bid-ask spread. Figure 2d confirms this conjecture. End-of-day spreads of ETFs decrease over time, but at times of market stress they increase. Especially, the bid-ask spreads increased dramatically during the crisis of 2008 paralleling the increase in mispricing observed in Figure 2b. Intuitively, as liquidity dried up, the bid-ask spread

[11] To gauge the evolution of the magnitude of mispricing over time for the cross-section of ETFs, we deem that the dispersion of the cross-sectional distribution is a more meaningful statistic than, say, the mean or the median. Because mispricing can be positive and negative, the latter statistics could provide the false impression that mispricing is low, when indeed for some funds it is very large and positive and for others it is very large and negative.

enlarged and arbitrageurs found it less profitable to trade on ETF mispricing, which widened as well. Incidentally, we note from Figure 2d that large drops of the ETF bid-ask spread occurred around August 2000 and February 2001. This is possibly the result of the decimalization of quotes on the Amex, where most ETFs were trading at the time (see Chen, Chou, and Chung 2008).

To obtain more systematic evidence on the determinants of mispricing we turn to regression analysis. In Table 3, we run time series regressions at the daily frequency where the dependent variables are summary measures of the daily ETF mispricing. The right-hand side variables are chosen to proxy for times where arbitrage capital is more likely to be scarce. Following Hameed, Kang, and Viswanthan (2011), we use the stock market (value-weighted index) return in the prior five days to approximate the change in capital constraints in the market making sector. For the same purpose, we consider the prior-five-day return for the financial sector portfolio, which includes broker-dealers and excludes commercial banks (from Prof. Ken French's forty-nine industry portfolios). Based on Nagel's (2011) results that times of high VIX are related to a decrease in the supply of liquidity, we include the average level of the VIX in the prior five days. Following Boyson, Stahel, and Stulz (2010) we include the TED spread through its average over the prior five trading days.[12] Finally, we construct a measure of arbitrageurs' profits as follows. For each ETF, we compute the daily return from a strategy going long (short) in the cheaper (more expensive) between the ETF and the basket of underlying securities, based on prior day closing price and NAV. To have an aggregate measure, the ETF-level variable is averaged across ETFs on each date. Finally, as for the other time-series, we compute its average over the prior five trading days. This variable is meant to measure the availability of trading capital to the arbitrageurs that are directly involved in ETF arbitrage. In all regressions we also include the prior-five-day average of the dependent variable, as mispricing displays high time dependence.

We consider different samples: Columns (1) to (3) present regressions using the entire sample of ETFs, Columns (4) to (6) exclude the peak of the financial crisis (the second half of 2008), which was arguably a 'special' time, and Columns (7) to (9) include only equity ETFs. We also study a sample that is limited to observations post-2000 (as transaction costs were

[12] The TED spread reflects the difference between the interbank interest rate and on short-term U.S. government debt.

substantially higher in the early years, as suggested by Figure 2d), which yields similar results and is excluded for brevity.

In Panel A of Table 3, the dependent variable is the interquartile range of ETF mispricing, which is plotted in Figure 2b. Consistent with a tightening of capital constraints on market makers and arbitrageurs, the estimates show an increase in the dispersion of mispricing following periods of low stock market returns. Even more convincing on the capital constraints channel, we find that mispricing increases following low past returns for the financial sector, controlling for the return on the stock market. Excluding the financial crisis (Columns (4) to (6)), we identify separate significant effects for the stock market and financial sector returns, and the two variables are jointly significant. In general, the dispersion of mispricing increases with the VIX index. The TED spread also has the expected sign and it makes an independent contribution from the VIX. Finally, the negative and significant sign on the proxy for arbitrage profits suggests that after arbitrageurs make money in this market, the magnitude of the of mispricing shrinks. This evidence suggests that arbitrageurs' ability to correct mispricing depends on their funding liquidity, consistent with the predictions in Brunnermeier and Pedersen (2009).

To corroborate our results, in Table 3, Panel B, we consider an alternative dependent variable, the fraction of ETFs with positive net mispricing, which is plotted in Figure 2c. The panel shows that following periods of low financial sector returns the fraction of ETFs with positive net mispricing increases. The result is even stronger than in Panel A, as it holds also when the financial crisis is not in the sample. Interestingly, the regressions show that the fraction of ETFs with positive net mispricing decreases as the VIX index increases. In unreported analysis, we find that this effect takes place because bid-ask spreads expand at periods of high VIX (see also Figure 2d). The TED spread and, especially, the proxy for arbitrage profits seem to have a significant impact on this measure of mispricing.

Overall, the results in this section present evidence that is consistent with the idea that ETF mispricing is larger at times in which arbitrageurs scale back their involvement in the market, either because they are losing capital or because of increased uncertainty.

### 4.2 Limits of Arbitrage in the Cross-Section

To provide additional evidence on the relation between ETF mispricing and limits to arbitrage, we exploit the cross-section of ETFs. Specifically, in Table 4 we regress the absolute value of mispricing on cross sectional determinants. Again, our focus is the magnitude of the mispricing more than its direction. Our sample is a panel of daily ETFs between 1998 and 2010. Time and fund fixed effects are included and standard errors are clustered at the fund level.

The first result that we examine is the relation between mispricing and arbitrageurs' past performance. This relation is negative: following weeks of losses, ETF mispricing is larger. The result is statistically significant ($t$ = 4.5) when all ETFs are considered, however it loses its statistical significance when equity ETFs are considered in isolation ($t$ = 1.4). This result is consistent with the idea of limits to arbitrage: arbitrageurs have limited resources, and following losses they become more constrained in their trades, leading to greater mispricing.

Consistent with trading costs imposing a limit to arbitrage, the mispricing is bigger for ETFs when the ETF bid-ask spread increases. Instead, holding costs, as measured by return volatility, do not seem to impact mispricing. In all these regressions, we have ETFs fixed effects. Hence, these estimates capture the incremental effect relative to the average impact of trading and holding costs on mispricing.

To summarize the results from the time-series and the cross-sectional analyses, mispricing appears to be significantly larger when arbitrageurs capital is limited, aggregate uncertainty increases, and trading costs increase. This evidence, relating mispricing to the limits of arbitrage, indirectly suggests that arbitrage is taking place between the ETFs and the underlying securities.

## 5 ETFs and Shock Propagation

After showing that the arbitrage between ETFs and their underlying securities is limited, we can test whether in this context arbitrageurs' trade can propagate shocks from the ETF market to the underlying securities. In this part of the analysis, our focus is restricted to ETFs that trade in U.S. equity securities.

### 5.1 Base Specification: The Relation between ETF Mispricing and Returns

#### 5.1.1 ETF Mispricing and NAV Returns

The conjecture that we explore in this paper is that arbitrage activity propagates non-fundamental shocks across markets. ETFs are an ideal candidate to shed light on this hypothesis because of the tight arbitrage relation that links them to the underlying securities. A liquidity shock occurring in the ETF market can cause a deviation of the ETF price from the NAV. Then, arbitrage activity may induce price pressure in the market for the underlying securities in the same direction as the initial shock in the ETF market.

The first step in building this argument is to show that the underlying securities' prices move in the same direction as the ETF mispricing. Using daily data after 2000 for equity ETFs, in Columns (1) to (4) of Table 5, Panel A, we regress the day-$t$ return on the NAV onto the mispricing in day $t-1$ and other controls. Date fixed effects are always included and standard errors are clustered at the date level. Columns (1) and (2) show that, whether or not we control for fund fixed effects, the NAV return moves significantly in the same direction as the mispricing. This is consistent with the conjecture that arbitrage activity transmits a shock in the ETF market to the market of the underlying securities. The transmission occurs when there is a discrepancy between the ETF price and the NAV. As for the economic magnitude, for example, in Column (2) a one-standard deviation increase of mispricing in the previous day (0.619%) is associated with a 10 bps increase in the daily return of the NAV. Given that the daily expected return for the average stock is of the order of magnitude of a few bps, the magnitude seems sizeable.[13]

There is another possible interpretation of this result. Price discovery may be taking place in the ETF first and the underlying securities' prices may be following with some delay. For example, upon the arrival of news, investors may be trading on this information in the ETF market because it is less expensive than trading in the basket of the underlying assets. In this case, we would observe a temporary mispricing which is then closed as the NAV catches up with a delay. To account for this channel, in Columns (3) and (4), we include the ETF return in day $t-1$. This variable controls for the lead-lag relationship induced by early price discovery in the ETF

[13] If we take an equity premium of about 6% annually and 250 trading days in a year, this corresponds to a daily equity premium of 2.4 bps.

market, to the extent that this effect plays out within the daily lag. Furthermore, to confound our identification, the mispricing on day $t - 1$ may originate from a shock to the NAV on day $t - 1$. Then, the predictability that we observe may result from the reversal of this shock on day $t$ when the shock is absorbed. This would not be the evidence for contagion that we are after. To filter this effect out, we also control for the NAV return on day $t - 1$. Once these effects have been controlled for, the coefficient on mispricing arguably captures the impact of mispricing arbitrage on the next day's NAV. The relevant slope on mispricing in Columns (3) and (4) remains statistically significant. Quite intuitively, the magnitude declines by about 15% as we are filtering out the component of mispricing that results from day $t - 1$ movements both in the ETF price and the NAV. So, the residual component of mispricing is the mispricing that has accumulated in days prior to $t - 1$.

### 5.1.2 ETF Mispricing and ETF Returns

Arbitrage activity in response to mispricing is predicted to generate an ETF price movement of the opposite sign of the mispricing. So, to corroborate the conclusion that the estimated positive relation between mispricing and subsequent NAV returns is due to arbitrage, we run a regression of ETF return onto prior day mispricing. Columns (5) to (8) of Table 5, Panel A, replicate the set of explanatory variables from the previous models. The negative and significant slope on mispricing is consistent with the movement expected if arbitrage activity is taking place. It is interesting to compare the magnitude of the coefficients, e.g., between Columns (4) and (8) of Table 5. The coefficients on the mispricing variable have opposite signs and the magnitude is larger for the ETF price regressions (0.140 vs. -0.385), suggesting that given a mispricing in day $t - 1$, both the NAV and ETF move to close the mispricing on day $t$, with the ETF price moving faster. This evidence is consistent with the NAV being more closely tied to fundamental, while ETF prices are more sensitive to liquidity shocks. This could also happen because the ETF is more liquid than the underlying securities so that much of the trading and price movements occur in ETF market.

As the effect of non-fundamental shocks to the ETF price is to generate mispricing, the results in Table 5, Panel A, are consistent with the transmission of shocks from the ETF market to the prices of the underlying securities via arbitrage activity.

## 5.2 Identification of Non-Fundamental Shocks

Given the results in Table 5, Panel A, one cannot completely rule out the possibility that the initial shock originates from fundamental news that are impounded in the ETF price first, while the NAV follows with a delay. This would not be the evidence of contagion that we wish to produce. To be able to establish contagion, we need to identify non-fundamental shocks originating in the ETF market. To this purpose we proceed in two ways. First, we show that the effect of mispricing on the NAV is quickly absorbed in the next days, consistent with a liquidity shock. Second, we calculate the order imbalance of ETFs and of the underlying securities and we propose that large differences in these measures identify demand shocks in the ETF market.

### 5.2.1 Vector Auto-Regression Analysis

In Table 5, Panel B, we report the results from a Vector Auto-Regression (VAR) analysis for the mispricing and the NAV return with five lags. The ETF-day-level data used for the analysis of Table 5, Panel A, are averaged at each date. The average is value-weighted to replicate the performance of a portfolio that invests in each ETF according to its market capitalization. Given that mispricing is defined as the difference between the ETF price and the NAV, once a sufficient number of lags for the NAV return is included, a shock to the ETF return is mapped into a shock to the mispricing. For this reason, we leave out the ETF return from the VAR. Column (1) has the equation for the mispricing. We notice that the mispricing displays strong autocorrelation, which is significant at least up to the fifth lag.[14] In the same specification, the lags of the NAV return have a positive and significant impact on current mispricing, which is consistent with a positive feedback from the prior period NAV to the current ETF price. In Column (2), the dependent variable is the NAV return. As in Panel A of Table 5, we find a positive and significant relation between lagged mispricing and the NAV return. The larger magnitude in this specification probably originates from averaging the data, which reduces the noise. The additional information that we gather from this analysis is the fact that the second and third lags of the mispricing are negatively and significantly related to the current NAV return

[14] We omit higher order lags for parsimony and because they do not change the qualitative conclusions of the analysis.

with a magnitude that counterbalances the effect of the first lag. In other words, the effect of the initial shock from the ETF price to the NAV starts reverting from the second day.

In Figure 5 we report the Impulse Response Function (Figure 5a) and Cumulative Impulse Response Function (Figure 5b) for the NAV return after a one-unit shock to the mispricing, along with two-standard error bands. The graphs replicate the inference from the regression coefficients. A shock to mispricing on day 0 has a positive impact on the NAV return in day 1, which is counterbalanced by the negative effects on days 2 and 3. In particular, from Figure 5b, the cumulative impact on the NAV is already non-significantly different from zero by day 2. The evidence of reversal of the initial effect validates our conjecture that non-fundamental shocks are propagated from the ETF market to the prices of the underlying securities. If the shocks that are propagated were only due to fundamental news, one should not expect the NAV to revert to the initial level.

### 5.2.2 Discrepancies in Buy-Sell Order Imbalance

In the next set of tests, we identify non-fundamental demand shocks that hit only the ETF market by comparing the buy-sell order imbalance (OI) between this market and the market for the underlying securities. OI captures buying pressure (if OI is positive) or selling pressure (if OI is negative). OI is computed using the Lee and Ready (1991) algorithm, which uses intraday data (from TAQ) to classify transactions as buyer or seller initiated according to whether a trade price is above or below the corresponding quote mid-point. For each ETF we calculate OI daily. For the underlying stocks we value-weight daily OI using the weights of the stocks in the ETF and compute an OI measure for the whole portfolio.

We make the assumption that large OI in the ETF market that is not matched by large OI in the underlying securities identifies a demand shock that only concerns the ETF market and, for this reason, it is a non-fundamental shock. To exemplify, our strategy is meant to capture situations such as a large trade in the ETF market by an institution that needs to return liquidity to its investors, or to invest newly collected capital, in a similar spirit to Coval and Stafford (2007).

We define an indicator variable for large positive OI days for a given ETF as a realization of OI that is more than one standard deviation from the mean of OI for that ETF. The large negative OI indicator is also defined for the days on which the realization of the ETF-level OI is more than one standard deviation below the mean. Also, in the definition of these indicators, we require that the OI for the portfolio of underlying securities is within one standard deviation from its mean. When we define these indicators for day $t - 1$, we measure the OI of the ETF on day $t - 1$, while we measure the OI for the underlying stocks over three alternative intervals: day $t - 1$, days $t - 1$ to day $t + 4$, days $t - 6$ to day $t + 4$. We consider larger-than-one-day windows in measuring the OI for the underlying securities to allow for the potential non-synchronous manifestation of the demand shocks in the two markets. We assume that if large demand is present in the ETF market and not in the underlying securities market, then it has to come from a non-fundamental demand shock.

In Panel A of Table 6 the OI for the underlying securities is measured on $t - 1$. First, we use the large OI dummies as instruments for lagged mispricing in a two-stage least squares (2SLS) framework. The goal is to identify the component of mispricing that originates from non-fundamental shocks in the ETF market. The first stage (Column (1)) confirms our interpretation for the large OI dummies. Mispricing is positively (negatively) and significantly related to the large positive (negative) OI indicator. The effects appear symmetrical. In the second stage (Columns (2) and (4)) these indicators are used as instruments for lagged mispricing. The 2SLS analysis identifies the same predictability as in the OLS regressions of Table 5, Panel A. Same as above, we conclude that the component of mispricing that originates from non-fundamental shocks in the ETF market induces propagation of these shocks to the prices of the underlying stocks.

Then, we examine the relation between lagged ETF mispricing and current returns on the NAV and the ETF in a simple OLS regression, as in Table 5, Panel A. However, here we restrict the sample to the ETF-day observations for which the large OI indicators, either positive or negative, equal one. In Columns (2) and (4), we find the same predictability as for the whole sample. This finding suggests that the propagation of shocks from the ETF to the NAV is occurring also when these shocks likely originate from non-fundamental demand.

Tables 6, Panels B and C, provide robustness results. In particular, in these panels we vary the definition of the large OI indicators by measuring the OI of the underlying securities between $t - 1$ and $t + 4$ (Panel B) and between $t - 1$ and $t + 6$ (Panel C). The results in both panels resemble those in the Panel A thus supporting our conclusions.

### 5.3 Evidence of Arbitrage Activity from Share Creation, Order Imbalance, and Hedging Demand

#### 5.3.1 Share Creation as Response to ETF Mispricing

The next step in building our argument is to provide evidence that non-fundamental shock propagation occurs, at least partly, as a result of arbitrage activity between the ETF and the NAV. An alternative view to our conjecture is that separate groups of investors operate in the two markets and the investors in one market make inference about fundamental value from observing price realizations in the other market, as in the Cespa and Foucault's (2012) model. In this alternative framework there is no need of cross-market arbitrage for the propagation of shocks, as dealers in each market adjust prices on basis of the signals coming from the other market. While the two stories may very well coexist, to validate our conjecture we need to provide evidence consistent with arbitrage trading. We do this in three ways. First, we look at ETF share creation. Then, we study order imbalance. Finally, we test some cross-sectional implications of the arbitrage conjecture.

As explained above, the role of Authorized Participants (APs) is to avoid large discrepancies between the ETF price and its NAV. This process involves creation or redemption of ETF shares. In particular, if the ETF price trades above the NAV, the APs buy the underlying assets in the market and convert them into to ETF shares. On the other hand, if the ETF price trades at a discount relative to the NAV, the APs buy ETF shares in the market and redeem them for the underlying basket. Petajisto (2011) shows that on average share creation and redemption take place on about 21% of all trading days. The distribution being very skewed, the median is lower at 11%. Conditioning on the days when shares outstanding change, the amount of these trades is large, as the mean transaction accounts for 21% of fund's assets and the median for 5%.

Here, we wish to show that these changes in shares outstanding are linked to APs' arbitrage motive. To this purpose, we regress the percentage change in ETF shares as on lagged mispricing and, in some specifications, we control for the lagged returns of the ETF and NAV, to parallel one of the specifications in Tables 5 and 6. As in these tables, we restrict the focus to ETFs on US equity. The results are presented in Table 7. The regressions show that the number of ETF shares increases following days of increasing of ETF mispricing. The results are statistically significant. The economic magnitude is not large as, from the specification in Column (1), a one-standard deviation change in mispricing induces a roughly 3 bps change in shares outstanding on the next day. This magnitude likely reflects Petajisto's (2011) finding that shares are created and redeemed in a discontinuous fashion. Still, we can assert that the arbitrage motive is significantly tied to the process of share creation and redemption.

### 5.3.2 Order Imbalance in ETFs and Underlying Assets and ETF Mispricing

We note that we AP's actions are only one of the drivers of ETF arbitrage. Another important channel is the trading by hedge funds and other arbitrageurs. These institutions are involved at a higher frequency in the exploitation of ETF mispricing. Their activity does not entail changes in shares outstanding, which is why we cannot find an immediate trace of their actions. However, we should find a trace of their transactions in order imbalance. An important prediction of the arbitrage relation between ETFs and the underlying stocks is that positive ETF mispricing causes consequent selling pressure on the ETF and buying pressure on the underlying stocks, vice versa for negative mispricing.

We test this idea in Table 8 where we examine the impact of day-*t* ETF mispricing on the buy-sell order imbalance of ETFs and of their underlying portfolio in day $t + 1$. For the OI of ETFs, we expect a negative relationship: when ETF mispricing is high at the end of the day, i.e., ETFs are more expensive than the underlying securities, next-day OI in the ETF should be negative—arbitrageurs apply selling pressure on ETF shares. Vice versa, a positive mispricing should induce buying pressure, i.e. a positive OI, on the underlying stocks. In Columns (1) to (3) the dependent variable is next day's ETF OI. The explanatory variable of interest is the current ETF mispricing. Because daily OI is highly autocorrelated (see Chordia and Subrahmanyam 2004), as well as the daily mispricing, we include ten daily lags of each variable in the regression. Also, given this high degree of autocorrelation of the dependent variable, besides the usual standard errors clustered by date (Column (1)), we also report standard errors clustered by ETF (Column (2)), and standard errors computed with two-way clustering by both date and ETF (Column (3)), as suggested by Petersen (2009). Consistent with our conjecture, the regression shows that there is selling pressure on the ETF following an increase in ETF mispricing. We test whether there is buying pressure on the underlying securities following ETF mispricing in Columns (4) to (6). Here, we regress the next day's OI for the value-weighted portfolio of underlying stocks on current ETF mispricing, while controlling for current OI and the ten lags of these variables. The results show that, consistent with arbitrage activity, there is significant buying pressure at the underlying stocks' level following an increase in ETF mispricing.

### 5.3.3 Stock Characteristics and Sensitivity to ETF Mispricing

Finally, we focus on the stock-level implications of arbitrage activity. Similar to Greenwood (2005), we exploit the heterogeneity of hedging demands that a given liquidity shock in the ETF market generates for the underlying stocks. In particular, focusing on the Spider ETF on the S&P 500 equity index (symbol SPY), we conjecture that a given level of mispricing should generate larger return reaction in the stocks that make a larger hedging contribution in arbitrageurs' portfolios. This contribution is positively related to a stock's weight in the index, that is, its market capitalization. It is also potentially positively related to a stock's beta, as arbitrageurs who do optimized replication of the index select a subset of stocks with the highest correlation with the index. For a similar reason, stocks with high idiosyncratic volatility are less desirable to arbitrageurs and are likely to receive less hedging demand (also see, Pontiff 2006).

In Table 9, we regress next-day return for stocks in the S&P 500 on current mispricing on the SPY interacted with stock characteristics. Column (1) interacts ETF mispricing and logged stock market capitalization. The coefficient on the interaction is positive and statistically significant, consistent with the prediction on hedging demand being larger for stocks that have higher weight in the index. Columns (2) and (3) confirm the sign of the predictions for beta and idiosyncratic volatility, although the statistical significance is weak. Finally, in Column (4), which has all the three characteristics together, the statistical significance of the beta interaction rises to the 10% level and the significance the idiosyncratic volatility interaction increases to the 5% level.

Overall, our results show that current ETF mispricing generates share creation, trading activity, and stock-level return reactions that are consistent with the exploitation of this profit opportunities by arbitrageurs.

### 5.4 ETF Ownership and the Effect on Volatility

#### 5.4.1 Changes in Volatility and Stock-Level ETF Ownership

If non-fundamental shocks to ETF prices are passed down to the securities that compose the ETF basket, we should expect ETF ownership to increase stock volatility *ceteris paribus*. For this to happen, it has to be the case that arbitrage activity takes place between the ETF and the underlying assets. The results in Section 3 reveal that the intensity of arbitrage activity is time-varying as a function of limits to arbitrage. When arbitrage is constrained ETF mispricing is larger across the board.

Based on this intuition, we develop a test of the effect of ETF ownership on stock volatility, using the interquartile range of mispricing in a given time period as an inverse proxy for the intensity of arbitrage activity (see Figure 2b and Table 3). In Panel B of Table 10 we look at the effect of a change in ETF ownership in month $t$ on the change in stock volatility between month $t$ and month $t + 1$. Again, the idea is that an increase in ETF ownership should bring about an increased exposure of the underlying stocks' prices to non-fundamental shocks. The results are in Panel B of Table 10. Stock fixed effects are included in all specifications along with a control for a change in total institutional ownership. Standard errors are clustered at stock level and stock fixed effects are added to some of the specifications. From Columns (1) and (2), we note that an increase in ETF ownership of the stock raises the stock daily volatility in the following month, which is consistent with shock transmissions from the ETF to the underlying stocks. In terms of magnitude (from Column (1)), a 1% increase in the ETF weight raises daily volatility by 3 bps. Hence, for the stock with the median ETF ownership in December 2010 (4.3% ETF ownership), the daily volatility has increased over time as a consequence of ETF ownership by roughly 13 bps[15]—which amounts to 3.4% of daily stock volatility.[16] For the stock at the 90th percentile of ETF ownership in December 2010 (ETF ownership of 7.9%), the cumulative increase in volatility is approximately 24 bps, or 6.3% of daily volatility.

Naturally, one would expect smaller stocks to be more sensitive to shock transmission from the ETFs due to their lower liquidity. In Columns (3) and (4) of Panel A, Table 10, we add an interaction between a small stock indicator (capitalization the below the CRSP median in the month) and the change in ETF weight. As expected, the regressions show that the magnitude of the increase in volatility is significantly larger for smaller stocks. Actually it appears that the entire effect of the change in ETF weight plays out among these smaller stocks, as the baseline effect is statistically insignificant.

### 5.4.2 Changes in Volatility and New ETF Holdings

In the third test of the effect of ETF ownership on stock volatility, we focus on ETFs that begin or stop holding a stock. Under our hypothesis, an increase in the number of ETFs that own the stock should increase stock volatility because of the increased exposure to the non-fundamental shocks coming from the ETF market. The opposite happens if ETFs stop holding the stock. The number of ETFs holding the stock is drawn from the ETF investment company filings with the SEC, and which are available in Thomson-Reuters Mutual Fund Ownership database.

In Columns (1) and (2) of Panel B, Table 10, we test this conjecture. The dependent variable is the same as in Panel B, the change in volatility between month $t$ and month $t+1$. The number of ETFs is measured at month $t$. We include as controls the change in total institutional ownership (as reported in the institutional 13F filings), logged market capitalization, volatility in month $t$, turnover in month $t$, and the number of ETFs that hold the stock in month $t$, as our focus is on the change (positive or negative) in this number. Standard errors are clustered at the stock level. Consistent with our conjecture, the regressions show that monthly volatility increases when additional ETFs start holding the stock, and it decreases when ETFs stop including the stock in their basket, holding constant the total number of ETFs that own the stock. Coverage by

---

[15] $4.3 \times 3$ bps = 12.9 bps.
[16] From Table 2, Panel C: Mean daily volatility is 3.8%.

one additional ETF increases the daily volatility in the next month by 0.016% to 0.019%. A withdrawal of an ETF decreases the daily volatility in the next month by 0.038% to 0.047%.

If this volatility effects occurs via the arbitrage activity induced by ETF mispricing, we should observe increased trading in the stock as new ETFs cover the stock. In Columns (3) and (4) we test this conjecture. The dependent variable is the change in turnover between months $t$ and $t + 1$. The explanatory variables are the same as in Columns (1) and (2). The results are consistent with our prior, as the change in turnover is significantly related to positive and negative changes in ETFs covering the stock. As more ETFs cover the stock, turnover increases, keeping constant the total number of ETFs owning the stock. Stock turnover decreases when ETFs stop holding the stock.

### 5.4.3 The Introduction of S&P 500 ETFs

To provide additional evidence about the effect of ETFs on stock volatility, we explore stock volatility around the introduction of new ETFs. We focus on two massive ETFs: IVV and VOO—both track the S&P 500. We use difference-in-difference methodology: we examine the differential effects on volatility of stocks in the S&P 500 relative to other stocks and relative to their own past volatility. We isolate the month before and the month after the introduction of the ETFs. Our dependent variable is the stocks' daily volatility calculated either in the month prior to the introduction of the ETFs or in the month following the introduction.

Table 11 performs the analysis. Columns (1) to (3) focus on the introduction of IVV and Columns (4) to (6) examine the introduction of VOO. Columns (7) to (9) present results for the combined sample. The variable of interest is the interaction between the indicator of whether the observation in question is from the post-introduction month and the indicator to whether the stock is included in the index. The results show that daily volatility of stocks that are included in the index is higher by about 0.5% for the introduction of the IVV, or around 0.2% for the introduction of the VOO.

Overall, the results in this section suggest that ETFs have a significant impact on the prices of the stocks in their basket. This effect results from arbitrage activity which propagates shocks in the ETF price to the prices of the underlying securities. As a result ETF ownership increases volatility of the underlying stocks.

## 6 Evidence from the Flash Crash

The events in the U.S. stock market in May 6, 2010 (the "Flash Crash") drew the attention of the media and of regulators to ETFs. On that day, the S&P 500 plunged nearly 6% within minutes and recovered by the end of the day (see Figure 4a). According to CFTC and SEC report (2010), which summarized their findings about the Flash Crash, the price decline began in the futures market, when a large institutional investor sold S&P 500 E-mini futures contracts at an increasing rate, which as a consequence led to a liquidity dryup in the futures market. At the present time, a full account on how the liquidity problem in the futures market led to a crash in the equity market is still missing.[17]

In this section we test whether arbitrage trading on ETFs contributed to transmitting the shock from the futures market to the equity market. The idea is that ETFs tracking the S&P 500 were arbitraged against two types of assets: the futures contracts (S&P 500 E-minis),[18] and the basket of underlying stocks (the S&P 500). The liquidity shock hit initially the futures contract. Consistent with the anecdotal evidence (see the CFTC and SEC 2010 preliminary and final reports), we conjecture that the arbitrage relation between the futures market and the ETF market led the ETFs to decline as well. Then, an arbitrage relation between ETFs and the underlying stocks led to the transmission of the liquidity shock to the equity market.

We begin by eye-balling the S&P 500 index (the NAV), the S&P 500 E-mini futures, and the SPY (the largest ETF on the S&P 500) in Figures 4b and 4c in the time period leading to trough of the three series, which occurred at about 14:45:45. The figure shows that the E-mini was leading the decline in price, then the ETF followed, and the NAV moved last. In most of the seconds in the two charts, during the way down, the NAV is located above the ETF. This suggests an explanation in which the futures price decline induced arbitrageurs to sell ETFs and

[17] Among traded securities, ETFs were among the ones that declined the most. The prevailing explanation among industry practitioners (e.g., Borkovec, Domowitz, Serbin, and Yegerman 2010) for this fact is that market makers for ETF pulled out of the market after suffering severe losses. As a result market liquidity dried up, leading to further decline in prices.

[18] Richie, Daigler, and Gleason (2008) describe the process in which the arbitrage between S&P 500 futures and the SPY ETF takes place.

by futures. Then, the ETF traded at a discount relative to the NAV, which made it profitable to buy ETFs and sell the basket of underlying securities, causing part of the decline in the S&P 500.

To test this relation more formally, we turn to a time-series regression framework, using one-second level data for the period between 14:30 and 14:45. In Table 12, Column (1), we regress the returns on the S&P 500 index on the SPY mispricing in the previous second. The positive coefficient suggests that the S&P 500 declined more strongly following seconds in which the mispricing was negative, i.e., the S&P 500 was above the SPY. The magnitude of the coefficient can be interpreted as follows: a one-standard deviation decrease in the SPY mispricing (i.e., the SPY is lower than S&P 500 index) is associated with a 0.6 bps decline in the S&P 500 in the following second.

Two potential non-mutually exclusive explanations can cause this relation. The first one is the arbitrage relation we have discussed so far: market participants buy the ETF and short sell the NAV. The second explanation is based on price discovery: market participants observe the prices of the futures contract and of the ETF, and use them as guidelines for the true valuation of the S&P 500 (Cespa and Foucault 2011). To disentangle the two stories, we control for the lagged returns of the S&P 500, the lagged returns of the SPY, and the lagged returns of the e-mini S&P 500 futures contract (Column (2)). The regression shows that the magnitude and the significance of the SPY mispricing remain intact even when these variables are included. There is also the possibility that the shock was transmitted from S&P 500 future directly to S&P 500 stocks, without passing through ETFs. We test this possibility in Column (3), where we introduce the one-second lagged mispricing of the E-mini. Indeed, our results show that this variable explains much of the variation of the S&P 500 returns (adjusted-$R^2$ increased from 0.102 in Column (2) to 0.223 in Column (3)). Following the introduction of this variable, the SPY mispricing variable is still statistically significant, however, has lower magnitude. A one-standard deviation decrease in the SPY mispricing is associated with a 0.25 bps decline in the S&P 500 in the following second.

To summarize, the results in this section are consistent with the idea that the arbitrage relation between ETFs and the underlying securities, and between ETFs and the futures market, contributed to the propagation of the Flash Crash from the futures market to the equity market.

# 7 Conclusion

The paper shows that arbitrage activity can lead to the propagation of non-fundamental shocks across assets that are tied by an arbitrage relation. We present several pieces of evidence on this mechanism in the ETF market. First, we show that arbitrage activity is taking place between ETFs and their underlying securities. Second, we show that coverage of stocks by ETFs is associated with increased volatility and turnover, especially in small stocks. Third, we present evidence from the Flash Crash demonstrating that ETFs served as a conduit for shock transmission from the futures market to the equity market.

Our results provide support for the theories of limits of arbitrage. Arbitrage does not only adjust prices of mispriced securities, but it can also move the price of securities that are correctly priced. Thus, the large amount of capital that is employed in arbitrage trading strategies does not necessarily improve the efficiency of prices if arbitrage is limited.

Our findings should be of interest to regulators. The evidence in the paper suggests that ETFs, a relatively new instrument that grew tremendously in the last few years, may increase the risk of contagion in financial markets by transmitting non-fundamental shocks. Our study of the Flash Crash of May 6, 2010, is a notable example in this direction. Furthermore, our conclusions bear on the current debate on the impact of high-frequency trading (HFT) on market stability. As much of ETF arbitrage is carried out at high frequencies, the evidence in the paper seems to suggest that HFT adds to the non-fundamental volatility of asset prices, at the very least. In more extreme situations, such as the Flash Crash, HFT can be highly destabilizing as it propagates shocks across markets at very high speed.

## References


Amihud, Yakov, 2002, Illiquidity and Stock Returns: Cross-Section and Time-Series Effects, *Journal of Financial Markets* 5(1), 31-56.

Anton, Miguel, and Christopher Polk, 2010, Connected Stocks, Working Paper, London School of Economics.

Barnhart, Scott W., and Stuart Rosenstein, 2010, Exchange-Traded Fund Introductions and Closed-End Fund Discounts and Volume, *Financial Review* 45(4), 973-994.

Barberis, Nicholas, Andrei Shleifer, and Jeffrey Wurgler, 2005, Comovement, *Journal of Financial Economics* 75(2), 283-317.

Basak, Suleyman, and Anna Pavlova, 2010, Asset Prices and Institutional Investors, Working Paper, London Business School.

Blackrock, 2011, ETF Landscape – Industry Review, 2011 – H1.

Borkovec, Milan, Ian Domowitz, Vitaly Serbin, and Henry Yegerman, 2010, Liquidity and Price Discovery in Exchange-Traded Funds, Investment Technology Group Report.

Boyson, Nicole M., Christof W. Stahel, and René M. Stulz, 2010, Hedge Fund Contagion and Liquidity Shocks, *Journal of Finance* 65(5), 1789-1816.

Bradley, Harold, and Robert E. Litan, 2010, Choking the Recovery: Why New Growth Firms Aren't Going Public and Unrecognized Risks of Future Market Disruptions, Kauffman Foundation.

Brunnermeier, Markus K., and Lasse H. Pedersen, 2009, Market Liquidity and Funding Liquidity, *Review of Financial Studies* 22, 2201-2238.

Cella, Cristina, Andrew Ellul, and Mariassunta Giannetti, 2011, Investors' Horizons and the Amplification of Market Shocks, Working Paper, Stockholm School of Economics and Indiana University.

Cespa, Giovanni, and Thierry Foucault, 2012, Illiquidity Contagion and Liquidity Crashes, Working paper, HEC Paris.

Chang, Yen-Cheng and Harrison G. Hong, 2011, Rules and Regression Discontinuities in Asset Markets, Working Paper, Princeton University.

Chen, Wei-Peng, Robin K. Chou, and Huimin Chung, 2008, Decimalization and the ETFs and Futures Pricing Efficiency, *Journal of Futures Markets* 29(2), 157-178.

Chordia, Tarun and Avinadhar Subrahmanyam, 2004, Order imbalance and individual stock returns: Theory and evidence, *Journal of Financial Economics* 72, 485-518.

Commodity Futures Trading Commission and Securities & Exchange Commission, 2010, Findings regarding the market events of May 6, 2010, www.sec.gov/news/studies/2010/marketevents-report.pdf

Coval, Joshua, and Erik Stafford, 2007, Asset Fire Sales (and Purchases) in Equity Markets, *Journal of Financial Economics* 86(2), 479-512.

Crawford, Steve S., James Hansen and Richard Price, 2010, CRSP Portfolio Methodology and the Effect on Excess Returns, Working Paper, Rice University.

Engle, Robert F., and Debojyoti Sarkar, 2006, Premiums-Disounts and Exchange-Traded Funds, *ETFs and Indexing* 1, 35-53.

Greenwood, Robin, Short- and Long-term Demand Curves for Stocks: Theory and Evidence on the Dynamics of Arbitrage. *Journal of Financial Economics* 75(3), 607-649.

Greenwood, Robin, and David Thesmar, 2011, Stock Price Fragility, *Journal of Financial Economics* 102(3), 471-490.

Gromb, Denis, and Dimitri Vayanos, 2010, Limits of Arbitrage: The State of the Theory, *Annual Review of Financial Economics* 2, 251–275.

Hameed, Allaudeen, Wenjin Kang, and S. Viswanathan, 2010, Stock Market Declines and Liquidity, *Journal of Finance* 65(1), 257-293.

Hamm, Sophia J.W., 2010, The Effect of ETFs on Stock Liquidity, Working Paper, The Ohio State University.

Hau, Harald, Massimo Massa, and Joel Peress, 2010, Do Demands For Currencies Slope Down? Evidence from the MSCI Global Index Change, *Review of Financial Studies* 23(4), 1681-1717.

Jiang, Guohua, Charles M. C. Lee, and Yi Zhang, 2005, Information Uncertainty and Expected Returns, *Review of Accounting Studies* 10, 185-221.

Jotikasthira, C., C. Lundblad, and T. Ramadorai, 2012, Asset Fire Sales and Purchases and the International Transmission of Financial Shocks, *Journal of Finance*, forthcoming.

Kodres, Laura E., and Matthew Pritsker, 2002, A Rational Expectations Model of Financial Contagion, *Journal of Finance*, 57, 769-799.

Kyle, Albert S., and Wei Xiong, Contagion as a Wealth Effect, *Journal of Finance* 56(4), 1401-1440.

Lee, Charles M. C., and Mark J. Ready, 1991, Inferring trade direction from intraday data, *Journal of Finance* 46(2), 733-746.

Lou, Dong, 2011, A Flow-Based Explanation of Returns Predictability, Working Paper, London School of Economics.

Madhavan, Anath, 2012, Exchange-Traded Funds, Market Structure and the Flash Crash, Working Paper, BlackRock Inc.

Marshall, Ben R., Nhut H. Nguyen, and Nuttawat Visaltanachoti, 2010, ETF Arbitrage, Working Paper, Massey University.

Nagel, Stefan, 2010, Evaporating Liquidity, Working Paper, Stanford University.

Newey, Whitney K., and Kenneth D. West, 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55(3), 703-708.

Pasquariello, Paolo, 2007, Imperfect Competition, Information Heterogeneity, and Financial Contagion, *Review of Financial Studies*, 20, 391-426.

Petajisto, Antti, 2011, Inefficiencies in the Pricing of Exchange-Traded Funds, Working paper, NYU Stern School of Business.

Petersen, Mitchell, 2009, Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches, *Review of Financial Studies* 22, 435-480.

Pontiff, Jeffrey, 2006, Costly arbitrage and the myth of idiosyncratic risk, *Journal of Accounting and Economics*, 42, 35-52.

Ramaswamy, Srichander, 2011, Market Structures and Systemic Risks of Exchange-Traded Funds, Bank of International Settlements Working Paper No 343.

Richie, Nivine, Robert Daigler, and Kimberly C. Gleason, 2007, The Limits to Stock Index Arbitrage: Examining S&P 500 Futures and SPDRs, *Journal of Futures Markets* 28(12), 1182-1205.

Shleifer, Andrei, 1986, Do Demand Curves for Stocks Slope Down? *Journal of Finance* 41, 579-590.

Trainor, William J. Jr., 2010, Do Leveraged ETFs Increase Volatility? *Technology and Investment* 1(3), 215-220.

Vayanos, Dimitri, and Paul Woolley, 2011, Fund Flows and Asset Prices: A Baseline Model, Working Paper, London School of Economics.

Wurgler, Jeffrey, 2011, On the Economic Consequences of Index-Linked Investing, in Challenges to Business in the Twenty-First Century, eds. Gerald Rosenfeld, Jay W. Lorsch, and Rakesh Khurana, American Academy of Arts and Sciences, Cambridge, MA..

## Appendix: List of Variables

| ETF variables | Description | Data Sources |
|---|---|---|
| ETF Return | ETF Closing Price and ETF distributions made during the period, divided by ETF closing price in the previous period | CRSP, Compustat, OptionMetrics |
| NAV Return | Change in the Net Asset Value of ETF portfolio securities. NAV is computed as the fair market value of all ETF security holdings, divided by ETF shares outstanding | CRSP Mutual Fund Database, Lipper |
| ETF Mispricing | Difference between ETF Price and ETF NAV. Positive (Negative) ETF mispricing is referred to as ETF Premium (Discount). | CRSP, Compustat, OptionMetrics, CRSP MFDB and Lipper |
| NAV Volatility | Standard deviation of the NAV return | CRSP Mutual Fund Database, Lipper |
| ETF relative bid-ask spread | Difference between closing ask and closing ask, relative to closing midpoint | CRSP, Compustat, OptionMetrics |
| Equity ETF | Identifying ETFs with the majority of portfolio in equity securities using Lipper (CRSP MFDB) and Morningstar investment objective codes. Non-Equity ETFs include Bond, commodities, derivatives (e.g. short bias, leveraged, etc.) and other asset classes. | CRSP Mutual Fund Database, Lipper, Morningstar |
| ETF Turnover | ETF Trading Volume during the period, scaled by period end ETF shares outstanding | CRSP, Compustat, OptionMetrics |
| ETF AUM | ETF market value calculated as day end shares outstanding multiplied by closing ETF price | CRSP, Compustat, OptionMetrics |
| ETF Short Interest Ratio in the past 30 days | End of month and mid-month short interest shares (adjusted) scaled by day end shares outstanding | Compustat |

| Cross Sectional Measures | Description | Data Sources |
|---|---|---|
| Daily interquartile range | interquartile range of mispricing across all ETFs in each time period used as an inverse proxy for the intensity of arbitrage activity | CRSP, Compustat, OptionMetrics, CRSP MFDB and Lipper |
| Daily fraction of ETFs with positive net mispricing | Number of ETFs with ETF price above the NAV, scaled by the total number of ETFs. Fraction > 0.5 is when most ETFs exhibit premiums possibly due to positive demand shocks | CRSP, Compustat, OptionMetrics, CRSP MFDB and Lipper |

## Appendix: List of Variables (Cont.)

| Stock Level Variables | Description | Data Sources |
|---|---|---|
| Daily volatility within the month (%) | Standard deviation of daily returns during the month | CRSP |
| Turnover | Period Volume scaled by period-end shares outstanding, after adjusting both volume and shares outstanding to splits and similar events. | CRSP |
| ETF weight in the stock (%) | Total shares owned by ETF scaled by total shares outstanding, for each common stock. ETF holdings are extracted from their most recent holdings reports (N-CSR, N-CSRS, and N-Qs) that they are required to file pursuant to the Investment Company Act of 1940, and which are collected by Thomson-Reuters Mutual Fund Ownership Database | Thomson-Reuters Mutual Fund Ownership Data |
| Total institutional ownership (%) | Total shares owned by institutions divided by stock shares outstanding. | Thomson-Reuters 13F Data |
| # ETFs first reporting to hold the stock | Using ETF mutual fund holdings report to determine the number of new ETFs that started reporting during that month and that they hold this stock. | Thomson-Reuters Mutual Fund Ownership Data |
| # ETFs last reporting to hold the stock | The number of ETFs that own this stock and that will never report their holdings afterwards. Conditional analysis on those two variables allows a better identification, by focusing on the increase in weights that coincide with inception of new ETFs that will hold the stock (and vice versa for stocks with decreasing ETF weights because of closing ETFs). | Thomson-Reuters Mutual Fund Ownership Data |
| # ETFs reporting to hold the stock | The breadth of ownership by ETF which is the number of ETFs that reported their holdings in this stock, in the most recent ETF mutual fund ownership filings. | Thomson-Reuters Mutual Fund Ownership Data |

## Appendix: List of Variables (Cont.)

| **Intraday Variables** | **Description** | **Data Sources** |
|---|---|---|
| S&P500 Return | Using TAQ and CME trade data for individual ETFs, common stocks, and E-minis, volume weighted average prices are constructed at the second intervals using all valid trades in each second. Intraday returns are then computed each second as the price in second t divided by the price in second t-1, minus one. If there are no trades in a particular second, the return is set to zero. S&P 500 returns are computed by averaging the returns of individual components each second, using as weights, the market value of S&P 500 components in day -1 | TAQ |
| SPY Return | | TAQ |
| E-Mini Return | | CME |
| S&P500 Stocks Average Order Imbalance | After computing the second-level buy sell imbalance as fraction of stock market value for each stock, a weighted average order imbalance is aggregated across all S&P500 components, similar to intraday return computation. | TAQ |
| SPY Average Short Volume | Using ARCA RegSho data, short volume are aggregated each second and then divided by total shares outstanding. | ARCA |

## Table 1. ETF Sample Description

The table presents the distribution of ETFs in our sample. Panel A has the number of ETFs at year-end and the average monthly total assets under management (AUM, in $billion) of ETFs over the year. Panel B presents summary statistics on AUM (in $billion), the number of funds, and a value-weighted expense ratio by objective code as of end of March 2011 (for funds for which the objective code is not missing). The last column of Panel B shows whether the fund is included in the equity funds' sample.

**Panel A: ETF Statistics, by Year**

| Year | # ETFs | AUM ($bn) |
|---|---|---|
| 1998 | 29 | 9 |
| 1999 | 32 | 16 |
| 2000 | 92 | 36 |
| 2001 | 118 | 59 |
| 2002 | 126 | 99 |
| 2003 | 136 | 124 |
| 2004 | 170 | 181 |
| 2005 | 223 | 258 |
| 2006 | 373 | 361 |
| 2007 | 633 | 507 |
| 2008 | 747 | 564 |
| 2009 | 822 | 607 |
| 2010 | 948 | 834 |
| 2011 | 986 | 1,019 |

## Table 1. ETF Sample Description (Cont.)

### Panel B: ETF Statistics, by Objective Code

| Fund Objective Code | AUM ($bn) | # Funds | VW Expense Ratio | Equity ETF |
|---|---|---|---|---|
| S&P 500 index objective funds | 95.6 | 4 | 0.09% | Yes |
| Growth funds | 82.6 | 94 | 0.21% | Yes |
| Emerging markets funds | 70.9 | 49 | 0.61% | Yes |
| Gold oriented funds | 57.6 | 24 | 0.44% | No |
| International funds | 53.5 | 38 | 0.35% | Yes |
| Small-cap funds | 36.7 | 30 | 0.21% | Yes |
| Mid-cap funds | 28.8 | 32 | 0.23% | Yes |
| Intermediate investment grade debt funds | 24.6 | 8 | 0.18% | No |
| Treasury inflation protected securities | 21.2 | 5 | 0.20% | No |
| Dedicated short bias funds | 20.4 | 97 | 0.94% | No |
| Corporate debt funds BBB-rated | 18.7 | 8 | 0.21% | No |
| Growth and income funds | 17.9 | 19 | 0.11% | Yes |
| Commodities funds | 16.7 | 64 | 0.78% | No |
| Latin American funds | 15.0 | 13 | 0.62% | Yes |
| China region funds | 14.4 | 19 | 0.73% | No |
| Pacific ex Japan funds | 13.4 | 14 | 0.56% | No |
| Financial services funds | 13.2 | 26 | 0.40% | Yes |
| Natural resources funds | 12.3 | 25 | 0.40% | Yes |
| Real estate funds | 12.0 | 15 | 0.32% | Yes |
| Short investment grade debt funds | 11.2 | 4 | 0.16% | No |
| Equity income funds | 10.3 | 13 | 0.38% | Yes |
| High current yield funds | 9.7 | 3 | 0.46% | Yes |
| Science & technology funds | 9.2 | 32 | 0.37% | Yes |
| Short U.S. treasury funds | 8.8 | 4 | 0.15% | No |
| European region funds | 8.2 | 25 | 0.47% | Yes |
| Health/biotechnology funds | 7.6 | 22 | 0.39% | Yes |
| General U.S. treasury funds | 7.5 | 14 | 0.15% | No |
| Basic materials funds | 5.9 | 19 | 0.39% | No |
| Currency funds | 5.7 | 32 | 0.47% | No |
| Japanese funds | 5.5 | 9 | 0.55% | Yes |
| Industrials funds | 5.2 | 22 | 0.37% | Yes |
| Ultra-short obligations funds | 5.2 | 3 | 0.15% | No |
| Consumer goods funds | 5.0 | 15 | 0.31% | Yes |
| Utility funds | 4.8 | 16 | 0.32% | Yes |
| Global natural resources funds | 4.2 | 17 | 0.55% | Yes |
| Diversified leverage funds | 3.8 | 14 | 0.95% | No |
| Specialty/miscellaneous funds | 3.7 | 18 | 0.56% | No |
| General municipal debt funds | 3.5 | 6 | 0.24% | No |
| Consumer services funds | 3.5 | 16 | 0.34% | Yes |
| Global funds | 3.0 | 13 | 0.39% | Yes |
| International income funds | 2.2 | 4 | 0.50% | No |
| Short municipal debt funds | 2.0 | 6 | 0.22% | No |
| Emerging markets debt funds | 2.0 | 2 | 0.57% | No |
| Global financial services funds | 2.0 | 7 | 0.65% | Yes |
| U.S. mortgage funds | 1.9 | 3 | 0.25% | No |
| International real estate funds | 1.6 | 7 | 0.58% | Yes |
| Pacific region funds | 1.4 | 4 | 0.16% | No |
| Telecommunication funds | 1.3 | 11 | 0.49% | Yes |
| International small-cap funds | 1.0 | 3 | 0.59% | Yes |
| Total or Average | 772.3 | 948 | 0.40% | |

## Table 2. Summary Statistics

The table presents summary statistics about the variables used in the regressions. Panel A shows summary statistics of ETF data aggregated at a daily level. Panel B shows summary statistics about the dataset that is at the ETF-day level. Panel C presents summary statistics for data at the stock-month level. Panel D presents second-level data used at the Flash Crash analysis.

### Panel A: Time-series, ETF-level, analysis

**ALL ETFs**

| | N | Mean | S.D. | Min | Median | Max |
|---|---|---|---|---|---|---|
| Daily interquartile range | 3104 | 0.00504 | 0.00371 | 0.00129 | 0.00402 | 0.0351 |
| Daily fraction of ETFs with positive net mispricing | 3104 | 0.325 | 0.163 | 0 | 0.362 | 0.723 |
| Past week stock market returns | 4509 | 0.00195 | 0.0255 | -0.186 | 0.00358 | 0.2 |
| Past week financial sector returns | 4509 | 0.00309 | 0.043 | -0.272 | 0.00457 | 0.373 |
| Past week average VIX | 4509 | 0.207 | 0.0857 | 0.0968 | 0.196 | 0.729 |

| CORRELATIONS | | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|---|
| Daily interquartile range | (1) | 1 | | | | |
| Daily fraction of ETFs with positive net mispricing | (2) | -0.388 | 1 | | | |
| Past week stock market returns | (3) | -0.1353 | -0.0308 | 1 | | |
| Past week financial sector returns | (4) | -0.117 | -0.0636 | 0.8865 | 1 | |
| Past week average VIX | (5) | 0.6188 | -0.2025 | -0.1367 | -0.0978 | 1 |

**EQUITY ETFs**

| | N | Mean | S.D. | Min | Median | Max |
|---|---|---|---|---|---|---|
| Daily interquartile range | 3104 | 0.00438 | 0.00323 | 0.00116 | 0.00348 | 0.0313 |
| Daily fraction of ETFs with positive net mispricing | 3104 | 0.298 | 0.154 | 0 | 0.322 | 0.728 |
| Past week stock market returns | 3099 | 0.00146 | 0.0284 | -0.186 | 0.00286 | 0.2 |
| Past week financial sector returns | 3099 | 0.00248 | 0.049 | -0.272 | 0.00372 | 0.373 |
| Past week average VIX | 3099 | 0.226 | 0.0918 | 0.1 | 0.218 | 0.729 |

| CORRELATIONS | | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|---|
| Daily interquartile range | (1) | 1 | | | | |
| Daily fraction of ETFs with positive net mispricing | (2) | -0.3938 | 1 | | | |
| Past week stock market returns | (3) | -0.1402 | -0.0503 | 1 | | |
| Past week financial sector returns | (4) | -0.1139 | -0.082 | 0.8849 | 1 | |
| Past week average VIX | (5) | 0.6026 | -0.1866 | -0.1299 | -0.0895 | 1 |

# Table 2. Summary Statistics (Cont.)

## Panel B: ETF-day level analysis

### All ETFs

| | N | Mean | S.D. | Min | Median | Max |
|---|---|---|---|---|---|---|
| ETF Ret | 1029590 | 0.000242 | 0.0187 | -0.0641 | 0.000553 | 0.0634 |
| NAV Ret | 1029590 | 0.000158 | 0.0182 | -0.0634 | 0.000429 | 0.0627 |
| abs(ETF mispricing) | 1029592 | 0.0041 | 0.00645 | 1.37E-08 | 0.00169 | 0.0409 |
| ETF mispricing | 1029592 | 0.000618 | 0.00679 | -0.0275 | 0.000264 | 0.0274 |
| Past week volatility(NAV) | 1029592 | 0.0152 | 0.0134 | 0.000527 | 0.0113 | 0.0773 |
| Past week EFT return | 1029592 | 0.00104 | 0.0391 | -0.132 | 0.00244 | 0.123 |
| ETF turnover | 1029592 | 0.0487 | 0.12 | 0 | 0.0108 | 0.824 |
| ETF relative bid-ask spread | 1029592 | 0.00491 | 0.0101 | 0.000126 | 0.00183 | 0.0723 |
| Number of times ETF shares changed in past 30 days | 1029592 | 3.92 | 5.62 | 0 | 1.07 | 25 |
| Average short interest in past 30 days | 1029592 | 0.0911 | 0.212 | 0.000152 | 0.0204 | 1.46 |
| Δ ETF Shares (%) | 1029590 | 0.164 | 1.5 | -5.17 | 0 | 10.3 |

| CORRELATIONS | | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ETF Ret | (1) | 1 | | | | | | | | | |
| NAV Ret | (2) | 0.882 | 1 | | | | | | | | |
| abs(ETF mispricing) | (3) | -0.0115 | -0.0225 | 1 | | | | | | | |
| ETF mispricing | (4) | 0.1677 | -0.0858 | 0.0686 | 1 | | | | | | |
| Past week volatility(NAV) | (5) | -0.0096 | -0.011 | 0.171 | -0.0508 | 1 | | | | | |
| Past week EFT return | (6) | -0.0318 | 0.0057 | -0.0655 | 0.0491 | -0.0899 | 1 | | | | |
| ETF turnover | (7) | -0.0062 | -0.0057 | -0.0339 | -0.0131 | 0.3271 | -0.0233 | 1 | | | |
| ETF relative bid-ask spread | (8) | -0.0157 | -0.019 | 0.3634 | 0.0132 | 0.1998 | -0.0461 | -0.0208 | 1 | | |
| Number of times ETF shares changed in past 30 days | (9) | -0.001 | -0.0004 | -0.1116 | 0.0081 | 0.0713 | -0.0056 | 0.2732 | -0.1559 | 1 | |
| Average short interest in past 30 days | (10) | -0.0031 | -0.0028 | -0.0817 | -0.0117 | 0.0494 | -0.0056 | 0.4289 | -0.0238 | 0.1926 | 1 |
| Δ ETF Shares (%) | (11) | -0.0048 | -0.0037 | 0.0075 | 0.037 | 0.009 | 0.0012 | 0.0359 | -0.0085 | 0.0459 | 0.0095 |

### Equity ETFs after 2000

| | N | Mean | S.D. | Min | Median | Max |
|---|---|---|---|---|---|---|
| ETF Ret | 709430 | 0.000276 | 0.018 | -0.0641 | 0.000773 | 0.0634 |
| NAV Ret | 709430 | 0.000191 | 0.0177 | -0.0634 | 0.000704 | 0.0627 |
| abs(ETF mispricing) | 709430 | 0.00361 | 0.00593 | 1.52E-08 | 0.00145 | 0.0405 |
| ETF mispricing | 709430 | 0.00036 | 0.00619 | -0.0274 | 0.000132 | 0.0271 |
| Past week volatility(NAV) | 709430 | 0.015 | 0.0123 | 0.000543 | 0.0115 | 0.0774 |
| Past week EFT return | 709430 | 0.00135 | 0.0376 | -0.132 | 0.00341 | 0.123 |
| ETF turnover | 709430 | 0.0383 | 0.1 | 0 | 0.00914 | 0.824 |
| ETF relative bid-ask spread | 709430 | 0.00466 | 0.00963 | 0.000126 | 0.00182 | 0.0723 |
| Number of times ETF shares changed in past 30 days | 709430 | 3.87 | 5.68 | 0 | 1 | 25 |
| Average short interest in past 30 days | 709430 | 0.104 | 0.234 | 0.000152 | 0.0205 | 1.46 |
| Δ ETF Shares (%) | 709430 | 0.142 | 1.44 | -5.17 | 0 | 10.3 |

| CORRELATIONS | | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ETF Ret | (1) | 1 | | | | | | | | | |
| NAV Ret | (2) | 0.89 | 1 | | | | | | | | |
| abs(ETF mispricing) | (3) | -0.0175 | -0.0263 | 1 | | | | | | | |
| ETF mispricing | (4) | 0.1438 | -0.1111 | 0.0281 | 1 | | | | | | |
| Past week volatility(NAV) | (5) | -0.0091 | -0.0125 | 0.2209 | -0.0485 | 1 | | | | | |
| Past week EFT return | (6) | -0.0323 | -0.0018 | -0.0898 | 0.0457 | -0.1043 | 1 | | | | |
| ETF turnover | (7) | -0.0067 | -0.0069 | -0.0374 | -0.0064 | 0.2027 | -0.0274 | 1 | | | |
| ETF relative bid-ask spread | (8) | -0.0189 | -0.0223 | 0.3726 | 0.027 | 0.2511 | -0.0621 | -0.0019 | 1 | | |
| Number of times ETF shares changed in past 30 days | (9) | -0.0012 | -0.0008 | -0.1433 | 0.007 | 0.0485 | -0.0028 | 0.2615 | -0.1408 | 1 | |
| Average short interest in past 30 days | (10) | -0.0025 | -0.0022 | -0.088 | -0.0084 | 0.0358 | -0.0045 | 0.5625 | -0.0301 | 0.2169 | 1 |
| Δ ETF Shares (%) | (11) | -0.0034 | -0.0028 | 0.0032 | 0.0326 | 0.0022 | 0.0132 | 0.0267 | -0.0045 | 0.0417 | 0.0155 |

# Table 2. Summary Statistics (Cont.)

## Panel C: Stock-month level analysis

| | N | Mean | S.D. | Min | Median | Max |
|---|---|---|---|---|---|---|
| Daily volatility within the month (%) | 545838 | 3.8 | 2.92 | 0.564 | 2.91 | 16.4 |
| Monthly change in daily volatility | 543456 | -0.00278 | 2.14 | -7.09 | -0.0529 | 8.28 |
| Turnover (1000x#shares traded/#shares outstanding) | 547405 | 0.292 | 2.9 | 0 | 0.0742 | 883 |
| Monthly change in turnover | 536522 | 30.5 | 2486 | -249058 | -0.419 | 855212 |
| ETF weight in the stock (%) | 421903 | 2.46 | 2.02 | 2.42E-06 | 1.87 | 9.03 |
| Monthly change in ETF weight | 410980 | 0.0413 | 0.278 | -0.998 | 0.000562 | 1.24 |
| Total institutional ownership (%) | 556285 | 43.7 | 32.5 | 0 | 41.6 | 110 |
| Monthly change in institutional ownership | 545740 | 0.177 | 2.56 | -10.1 | 0 | 12.5 |
| # ETFs first reporting to hold the stock | 421903 | 0.54 | 1.58 | 0 | 0 | 21 |
| # ETFs last reporting to hold the stock | 421903 | 0.113 | 0.405 | 0 | 0 | 7 |
| # ETFs reporting to hold the stock | 421903 | 14.2 | 13.7 | 1 | 11 | 87 |
| log(market capitalization/1000) | 547405 | 19.4 | 2.12 | 11.7 | 19.3 | 27.1 |
| Interquintile mispricing of ETFs in the month | 559469 | 0.00455 | 0.00288 | 0.00159 | 0.0036 | 0.0163 |
| log(volume) | 547526 | 16.8 | 2.41 | 3 | 17 | 25.4 |

| CORRELATIONS | | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Daily volatility within the month (%) | (1) | 1 | | | | | | | | | | | | |
| Monthly change in daily volatility | (2) | -0.3662 | 1 | | | | | | | | | | | |
| Turnover (1000x#shares traded/#shares outstanding) | (3) | 0.2084 | -0.0255 | 1 | | | | | | | | | | |
| Monthly change in turnover | (4) | 0.0891 | -0.0227 | 0.7086 | 1 | | | | | | | | | |
| ETF weight in the stock (%) | (5) | -0.1275 | 0.0039 | -0.0058 | -0.0033 | 1 | | | | | | | | |
| Monthly change in ETF weight | (6) | 0.0115 | -0.0108 | -0.0173 | -0.0023 | 0.1198 | 1 | | | | | | | |
| Total institutional ownership (%) | (7) | -0.2331 | 0.0063 | -0.0424 | -0.0058 | 0.5248 | 0.061 | 1 | | | | | | |
| Monthly change in institutional ownership | (8) | -0.0676 | 0.0128 | -0.033 | -0.0125 | -0.0431 | 0.0043 | 0.0527 | 1 | | | | | |
| # ETFs first reporting to hold the stock | (9) | -0.0162 | 0.007 | -0.0152 | -0.0019 | 0.0614 | 0.0735 | 0.1704 | 0.004 | 1 | | | | |
| # ETFs last reporting to hold the stock | (10) | -0.0672 | -0.0117 | -0.0031 | -0.0032 | 0.3655 | -0.0019 | 0.133 | -0.0035 | -0.0551 | 1 | | | |
| # ETFs reporting to hold the stock | (11) | -0.183 | 0.0059 | -0.0247 | -0.0056 | 0.7037 | 0.065 | 0.529 | -0.0381 | 0.2619 | 0.3808 | 1 | | |
| log(market capitalization/1000) | (12) | -0.375 | -0.0132 | -0.102 | -0.0271 | 0.2412 | 0.0301 | 0.5359 | 0.0279 | 0.2765 | 0.1399 | 0.6141 | 1 | |
| Interquintile mispricing of ETFs in the month | (13) | 0.3781 | -0.0381 | 0.0211 | 0.0094 | -0.1447 | 0.0578 | -0.1039 | -0.0485 | 0.0313 | -0.061 | -0.1176 | -0.0789 | 1 |
| log(volume) | (14) | -0.0105 | -0.0678 | 0.0786 | 0.025 | 0.3164 | 0.0361 | 0.565 | 0.0045 | 0.2386 | 0.1531 | 0.5873 | 0.7591 | -0.0345 |

## Table 2. Summary Statistics (Cont.)

**Panel D: Intraday (May 6, 2010) second-level analysis**

**Whole sample**

| | N | Mean | S.D. | Min | Median | Max |
|---|---|---|---|---|---|---|
| Return S&P500 | 1800 | -0.000008 | 0.000467 | -0.003680 | -0.000014 | 0.003550 |
| SPY mispricing | 1800 | 0.003420 | 0.007720 | -0.009510 | 0.000141 | 0.032400 |
| Return Emini | 1794 | -0.000007 | 0.000510 | -0.006410 | -0.000015 | 0.006270 |
| Return SPY | 1800 | -0.000009 | 0.002420 | -0.025000 | -0.000017 | 0.025100 |
| S&P500 Order Imbalance | 1801 | -0.012300 | 0.049700 | -0.340000 | -0.005900 | 0.251000 |
| SPY average short volume (t, t+5) | 1801 | 0.001600 | 0.001430 | 0.000013 | 0.001170 | 0.011000 |

| CORRELATIONS | | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|---|
| Return S&P500 | (1) | 1 | | | | |
| SPY mispricing | (2) | 0.131 | 1 | | | |
| Return Emini | (3) | 0.238 | 0.087 | 1 | | |
| Return SPY | (4) | 0.077 | -0.031 | 0.121 | 1 | |
| S&P500 Order Imbalance | (5) | 0.342 | 0.151 | 0.337 | 0.076 | 1 |
| SPY average short volume (t, t+5) | (6) | -0.022 | 0.218 | -0.079 | 0.007 | -0.053 |

**Before trough**

| | N | Mean | S.D. | Min | Median | Max |
|---|---|---|---|---|---|---|
| Return S&P500 | 945 | -0.000062 | 0.000209 | -0.001360 | -0.000023 | 0.000713 |
| SPY mispricing | 945 | -0.000132 | 0.001040 | -0.009510 | 0.000080 | 0.005260 |
| Return Emini | 939 | -0.000065 | 0.000453 | -0.006410 | -0.000032 | 0.006270 |
| Return SPY | 945 | -0.000068 | 0.000683 | -0.005590 | -0.000031 | 0.007620 |
| S&P500 Order Imbalance | 946 | -0.023100 | 0.057900 | -0.340000 | -0.015200 | 0.251000 |
| SPY average short volume (t, t+5) | 946 | 0.001480 | 0.001180 | 0.000013 | 0.001170 | 0.009160 |

| CORRELATIONS | | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|---|
| Return S&P500 | (1) | 1 | | | | |
| SPY mispricing | (2) | 0.353 | 1 | | | |
| Return Emini | (3) | 0.608 | 0.098 | 1 | | |
| Return SPY | (4) | 0.319 | -0.139 | 0.391 | 1 | |
| S&P500 Order Imbalance | (5) | 0.667 | 0.125 | 0.411 | 0.257 | 1 |
| SPY average short volume (t, t+5) | (6) | -0.280 | -0.315 | -0.157 | -0.111 | -0.158 |

**After trough**

| | N | Mean | S.D. | Min | Median | Max |
|---|---|---|---|---|---|---|
| Return S&P500 | 855 | 0.000051 | 0.000636 | -0.003680 | 0.000039 | 0.003550 |
| SPY mispricing | 855 | 0.007350 | 0.009750 | -0.006050 | 0.002140 | 0.032400 |
| Return Emini | 855 | 0.000057 | 0.000559 | -0.003010 | 0.000020 | 0.002980 |
| Return SPY | 855 | 0.000057 | 0.003440 | -0.025000 | 0.000026 | 0.025100 |
| S&P500 Order Imbalance | 855 | -0.000281 | 0.034800 | -0.223000 | 0.002870 | 0.140000 |
| SPY average short volume (t, t+5) | 855 | 0.001750 | 0.001660 | 0.000016 | 0.001150 | 0.011000 |

| CORRELATIONS | | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|---|
| Return S&P500 | (1) | 1 | | | | |
| SPY mispricing | (2) | 0.077 | 1 | | | |
| Return Emini | (3) | 0.139 | 0.036 | 1 | | |
| Return SPY | (4) | 0.059 | -0.048 | 0.092 | 1 | |
| S&P500 Order Imbalance | (5) | 0.296 | 0.074 | 0.237 | 0.057 | 1 |
| SPY average short volume (t, t+5) | (6) | 0.027 | 0.271 | -0.050 | 0.023 | 0.012 |

## Table 3. Limits of Arbitrage in the Time Series

The table presents regressions using day-level data. Panel A regresses the interquartile range of ETF mispricing on time-series determinants. Panel B regresses the daily fraction of ETFs with net mispricing (i.e., NAV is outside the bid-ask bounds), on time-series determinants. All regressions are OLS regressions. Standard errors are clustered at the ETF level. *t*-statistics are presented in parentheses. ***, **, * represent statistical significance at the 1%, 5%, or 10% levels, respectively.

### Panel A: Determinants of Time-Series Interquartile Range of Mispricing

| | Dependent variable: Interquartile range of ETF mispricing | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Entire sample | | | | Excluding crisis | | | | Equity ETFs | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Past week stock market returns | -0.015*** | -0.009*** | -0.008*** | -0.009*** | -0.008*** | -0.005* | -0.003 | -0.005* | -0.011*** | -0.013*** | -0.013*** | -0.012*** |
| | (-11.377) | (-3.348) | (-2.775) | (-3.171) | (-6.411) | (-1.678) | (-1.168) | (-1.778) | (-11.411) | (-6.682) | (-6.434) | (-6.293) |
| Past week financial sector returns | | -0.004** | -0.004** | -0.003** | | -0.002 | -0.003* | -0.002 | | 0.002 | 0.001 | 0.001 |
| | | (-2.203) | (-2.428) | (-2.000) | | (-1.528) | (-1.790) | (-1.242) | | (1.525) | (1.274) | (1.268) |
| Past week average VIX | | | 0.002*** | 0.004*** | | | 0.002*** | 0.003*** | | | 0.005*** | 0.005*** |
| | | | (3.329) | (5.691) | | | (3.815) | (4.017) | | | (9.869) | (10.134) |
| Past week average TED spread | | | 0.016* | 0.031*** | | | 0.009 | 0.012 | | | 0.022*** | 0.026*** |
| | | | (1.686) | (3.227) | | | (0.893) | (1.170) | | | (3.499) | (4.040) |
| Past week average arbitrage profits | | | | -0.037*** | | | | -0.011 | | | | -0.016** |
| | | | | (-5.333) | | | | (-1.601) | | | | (-2.338) |
| Past week average IQ range | 0.923*** | 0.924*** | 0.873*** | 0.901*** | 0.932*** | 0.933*** | 0.892*** | 0.899*** | 0.843*** | 0.842*** | 0.584*** | 0.625*** |
| | (82.400) | (82.472) | (52.726) | (51.726) | (82.495) | (82.523) | (59.147) | (56.090) | (56.274) | (56.102) | (21.366) | (19.239) |
| Constant | 0.000*** | 0.000*** | 0.000 | -0.000 | 0.000*** | 0.000*** | -0.000 | -0.000 | 0.000*** | 0.000*** | -0.000 | -0.000 |
| | (5.863) | (5.803) | (1.052) | (-0.727) | (5.094) | (5.065) | (-0.195) | (-0.350) | (9.224) | (9.285) | (-0.610) | (-1.343) |
| Observations | 3,099 | 3,099 | 3,099 | 3,098 | 2,950 | 2,950 | 2,950 | 2,949 | 3,099 | 3,099 | 3,099 | 3,098 |
| Adj. $R^2$ | 0.692 | 0.692 | 0.694 | 0.697 | 0.698 | 0.698 | 0.699 | 0.700 | 0.534 | 0.534 | 0.552 | 0.553 |

### Panel B: Determinants of the Daily Fraction of ETFs with Positive Net Mispricing

| | Dependent variable: Daily fraction of ETFs with positive net mispricing | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Entire sample | | | | Excluding crisis | | | | Equity ETFs | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Past week stock market returns | -0.200*** | -0.022 | -0.042 | -0.030 | -0.191*** | -0.064 | -0.078 | -0.085 | -0.220*** | -0.035 | -0.060 | -0.085 |
| | (-5.448) | (-0.281) | (-0.534) | (-0.383) | (-4.703) | (-0.743) | (-0.899) | (-0.981) | (-4.157) | (-0.308) | (-0.520) | (-0.743) |
| Past week financial sector returns | | -0.116** | -0.111** | -0.109** | | -0.084* | -0.080 | -0.071 | | -0.121* | -0.117* | -0.099 |
| | | (-2.548) | (-2.421) | (-2.392) | | (-1.663) | (-1.570) | (-1.404) | | (-1.835) | (-1.759) | (-1.500) |
| Past week average VIX | | | -0.024* | 0.056*** | | | -0.030* | 0.050** | | | -0.051*** | 0.006 |
| | | | (-1.734) | (2.882) | | | (-1.936) | (2.459) | | | (-2.585) | (0.229) |
| Past week average TED spread | | | -0.099 | 0.543** | | | -0.061 | 0.389 | | | 0.275 | 0.719* |
| | | | (-0.388) | (1.968) | | | (-0.189) | (1.190) | | | (0.741) | (1.830) |
| Past week average arbitrage profits | | | | -1.146*** | | | | -1.287*** | | | | -1.136*** |
| | | | | (-5.930) | | | | (-6.064) | | | | (-3.267) |
| Past week average IQ range | 0.977*** | 0.976*** | 0.973*** | 0.962*** | 0.979*** | 0.978*** | 0.973*** | 0.958*** | 0.917*** | 0.915*** | 0.910*** | 0.908*** |
| | (152.690) | (152.390) | (146.512) | (139.632) | (150.524) | (150.015) | (141.364) | (131.180) | (75.885) | (75.555) | (73.064) | (73.070) |
| Constant | 0.008*** | 0.009*** | 0.016*** | 0.014*** | 0.008*** | 0.008*** | 0.017*** | 0.018*** | 0.024*** | 0.024*** | 0.036*** | 0.031*** |
| | (3.540) | (3.689) | (3.908) | (3.431) | (3.344) | (3.455) | (3.557) | (3.927) | (6.362) | (6.483) | (6.256) | (5.258) |
| Observations | 3,099 | 3,099 | 3,099 | 3,098 | 2,950 | 2,950 | 2,950 | 2,949 | 3,099 | 3,099 | 3,099 | 3,098 |
| Adj. $R^2$ | 0.883 | 0.883 | 0.883 | 0.885 | 0.885 | 0.885 | 0.885 | 0.886 | 0.651 | 0.651 | 0.652 | 0.653 |

## Table 4. Limits of Arbitrage in the Cross Section

The table presents regressions using ETF-day-level data. The dependent variable is the absolute value of ETF mispricing. The independent variables include ETF-level determinants: past week arbitrage of ETF mispricing profits, ETF bid-ask spread at the end of the day, past week average abs(ETF mispricing), and the volatility of the daily ETF returns in the preceding month. Calendar day fixed effects and ETF fixed effects are included in all regressions. All regressions are OLS regressions. Standard errors are clustered at the ETF level. *t*-statistics are presented in parentheses. ***, **, * represent statistical significance at the 1%, 5%, or 10% levels, respectively.

| | Depedent variable: abs(ETF mispricing) | | | |
|---|---|---|---|---|
| Sample: | All ETFs | | Equity ETFs | |
| Sample period: | 1998-2010 | 2001-2010 | 1998-2010 | 2001-2010 |
| | (1) | (2) | (3) | (4) |
| Past week arbitrage profits | -0.007*** | -0.007*** | -0.005 | -0.005 |
| | (-4.575) | (-4.485) | (-1.354) | (-1.403) |
| ETF relative bid-ask spread | 0.118*** | 0.121*** | 0.162*** | 0.162*** |
| | (18.885) | (19.179) | (15.259) | (15.658) |
| Past month return volatility | 0.003 | 0.003 | -0.002 | -0.002 |
| | (0.618) | (0.524) | (-0.384) | (-0.265) |
| Past week average | 0.503*** | 0.503*** | 0.453*** | 0.450*** |
| abs(ETF mispricing) | (15.695) | (15.225) | (8.366) | (8.066) |
| Calendar day fixed effects | Yes | Yes | Yes | Yes |
| ETF fixed effects | Yes | Yes | Yes | Yes |
| Observations | 876,494 | 857,145 | 476,028 | 465,607 |
| Adj. $R^2$ | 0.459 | 0.462 | 0.385 | 0.390 |

## Table 5. ETF Mispricing and Subsequent NAV and ETF Returns

The table presents regressions using ETF-day-level data. In Panel A, Columns (1) through (4) regress the returns on NAV at time *t* on lagged determinants: ETF mispricing, NAV return, and ETF return. Columns (5) through (8) regress the returns on ETF at time *t* on lagged determinants: ETF mispricing, NAV return, and ETF return. Calendar day fixed effects are included in all regressions, and ETF fixed effects are included in Columns (2), (4), (6), and (8). Standard errors are clustered at the ETF level. Panel B presents a vector auto-regression (VAR) analysis of current mispricing and NAV return as a function of lagged mispricing and NAV return. All regressions are OLS regressions. *t*-statistics are presented in parentheses. ***, **, * represent statistical significance at the 1%, 5%, or 10% levels, respectively.

### Panel A: One-Day Relation between ETF Mispricing and NAV and ETF Returns

| | NAV Ret(t) | | | | ETF Ret(t) | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Mispricing(t-1) | 0.141*** | 0.164*** | 0.118*** | 0.140*** | -0.454*** | -0.539*** | -0.316*** | -0.385*** |
| | (10.998) | (11.014) | (10.126) | (10.159) | (-27.762) | (-30.757) | (-17.869) | (-19.006) |
| NAV Ret(t-1) | | | -0.071*** | -0.067*** | | | 0.185*** | 0.171*** |
| | | | (-3.952) | (-3.803) | | | (6.741) | (6.417) |
| ETF Ret(t-1) | | | 0.014 | 0.010 | | | -0.267*** | -0.253*** |
| | | | (1.385) | (0.954) | | | (-11.561) | (-11.046) |
| Calendar day fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| ETF fixed effects | No | Yes | No | Yes | No | Yes | No | Yes |
| Observations | 515,151 | 515,151 | 514,797 | 514,797 | 515,190 | 515,190 | 514,835 | 514,835 |
| Adj. $R^2$ | 0.004 | 0.005 | 0.008 | 0.008 | 0.037 | 0.044 | 0.068 | 0.071 |

### Table 5. ETF Mispricing and subsequent NAV and ETF Returns (Cont.)

**Panel B: VAR Analysis of ETF Mispricing and consequent NAV and ETF Returns**

| | Mispricing(t) | NAV Ret(t) |
|---|---|---|
| | (1) | (2) |
| Mispricing(t-1) | 0.332*** | 0.390*** |
| | (17.466) | (3.376) |
| Mispricing(t-2) | 0.170*** | -0.398*** |
| | (8.477) | (-3.275) |
| Mispricing(t-3) | 0.195*** | -0.196 |
| | (9.769) | (-1.619) |
| Mispricing(t-4) | 0.105*** | 0.059 |
| | (5.254) | (0.486) |
| Mispricing(t-5) | 0.110*** | 0.124 |
| | (5.765) | (1.074) |
| | | |
| NAV Ret(t-1) | 0.010*** | -0.073*** |
| | (3.206) | (-3.795) |
| NAV Ret(t-2) | 0.014*** | -0.059*** |
| | (4.293) | (-3.055) |
| NAV Ret(t-3) | 0.006** | 0.004 |
| | (1.992) | (0.200) |
| NAV Ret(t-4) | 0.001 | -0.017 |
| | (0.448) | (-0.867) |
| NAV Ret(t-5) | -0.001 | -0.038** |
| | (-0.381) | (-1.980) |
| | | |
| Constant | 0.000* | 0.000 |
| | (1.929) | (0.385) |
| | | |
| Observations | 2,752 | 2,752 |
| $R^2$ | 0.6593 | 0.0184 |

## Table 6. Identifying Liquidity Shocks with Order Imbalance

The table presents regressions using ETF-day-level data. The sample is all equity ETFs between 2001 and 2010. Columns (1), (2) and (4) present results from 2SLS regressions. Columns (3) and (5) present OLS regressions: Column (1) is the first stage regression and Columns (2) and (4) are second stage regressions. The dependent variable in Columns (2) and (3) is the return of the NAV. The dependent variable in Columns (4) and (5) is the return of the ETF. The dependent variable in Column (1) is the lagged ETF mispricing. The independent variables include: the lagged ETF mispricing and two dummies for large positive and negative order imbalance (OI) in the ETF which is not matched by large OI in the underlying stocks. Large OI is defined as a daily realization of ETF OI that is larger than one standard deviation from the mean in absolute value. Calendar day fixed effects are included in all regressions. Standard errors are clustered at the date level. *t*-statistics are presented in parentheses. ***, **, * represent statistical significance at the 1%, 5%, or 10% levels, respectively.

### Panel A: OI of NAV computed at *t* - 1

| | First Stage | Second Stage | OLS | Second Stage | OLS |
|---|---|---|---|---|---|
| Dependent variable: | Mispricing(t-1) | NAV Ret(t) | NAV Ret(t) | ETF Ret(t) | ETF Ret(t) |
| Sample: | All | All | OI ETF > 1 st dev | All | OI ETF > 1 st dev |
| | (1) | (2) | (3) | (4) | (5) |
| Mispricing (t -1) | | 0.342*** | 0.089*** | -0.205** | -0.438*** |
| | | (3.687) | (6.445) | (-2.257) | (-14.796) |
| NAV Ret(t-1) | | -0.012 | -0.057** | 0.204*** | 0.256*** |
| | | (-0.516) | (-2.143) | (4.646) | (6.746) |
| ETF Ret(t-1) | | -0.052** | 0.019 | -0.288*** | -0.285*** |
| | | (-2.048) | (1.264) | (-7.031) | (-8.596) |
| OI ETF-mean(OI ETF)>1 st dev | 0.001*** | | | | |
| & abs(OI NAV-mean OI NAV)<1 st dev | (17.284) | | | | |
| OI ETF-mean(OI ETF)<-1 st dev | -0.001*** | | | | |
| & abs(OI NAV-mean OI NAV)<1 st dev | (-14.913) | | | | |
| Calendar day fixed effects | Yes | Yes | Yes | Yes | Yes |
| Observations | 386,014 | 385,946 | 90,631 | 385,957 | 90,633 |
| Adj. $R^2$ | 0.004 | 0.001 | 0.004 | 0.071 | 0.123 |

### Panel B: OI of NAV computed on (*t* - 1, *t* + 4)

| | First Stage | Second Stage | OLS | Second Stage | OLS |
|---|---|---|---|---|---|
| Dependent variable: | Mispricing(t-1) | NAV Ret(t) | NAV Ret(t) | ETF Ret(t) | ETF Ret(t) |
| Sample: | All | All | OI ETF > 1 st dev | All | OI ETF > 1 st dev |
| | (1) | (2) | (3) | (4) | (5) |
| Mispricing (t -1) | | 0.293*** | 0.093*** | -0.225** | -0.456*** |
| | | (2.942) | (5.447) | (-2.326) | (-13.735) |
| NAV Ret(t-1) | | -0.022 | -0.079*** | 0.200*** | 0.207*** |
| | | (-0.918) | (-3.001) | (4.660) | (4.195) |
| ETF Ret(t-1) | | -0.041 | 0.022 | -0.284*** | -0.275*** |
| | | (-1.491) | (1.379) | (-6.919) | (-7.057) |
| OI ETF-mean(OI ETF)>1 st dev | 0.001*** | | | | |
| & abs(OI NAV-mean OI NAV)<1 st dev | (17.219) | | | | |
| OI ETF-mean(OI ETF)<-1 st dev | -0.001*** | | | | |
| & abs(OI NAV-mean OI NAV)<1 st dev | (-14.124) | | | | |
| Calendar day fixed effects | -0.000* | 0.000 | -0.000** | 0.000 | -0.000 |
| | (-1.816) | (0.370) | (-2.475) | (0.650) | (-0.923) |
| Observations | 386,026 | 385,958 | 86,666 | 385,969 | 86,666 |
| Adj. $R^2$ | 0.004 | 0.001 | 0.007 | 0.073 | 0.125 |

## Table 6. Identifying Liquidity Shocks with Order Imbalance (Cont.)

### Panel C: OI of NAV computed on (*t* - 6, *t* + 4)

| | First Stage | Second Stage | OLS | Second Stage | OLS |
|---|---|---|---|---|---|
| Dependent variable: | Mispricing(t-1) | NAV Ret(t) | NAV Ret(t) | ETF Ret(t) | ETF Ret(t) |
| Sample: | All | All | OI ETF > 1 st dev | All | OI ETF > 1 st dev |
| | (1) | (2) | (3) | (4) | (5) |
| Mispricing (t -1) | | 0.329*** | 0.093*** | -0.222** | -0.469*** |
| | | (3.327) | (5.746) | (-2.248) | (-13.984) |
| NAV Ret(t-1) | | -0.014 | -0.062** | 0.200*** | 0.206*** |
| | | (-0.622) | (-2.569) | (4.620) | (4.225) |
| ETF Ret(t-1) | | -0.049* | 0.009 | -0.284*** | -0.265*** |
| | | (-1.819) | (0.620) | (-6.844) | (-6.849) |
| OI ETF-mean(OI ETF)>1 st dev | 0.001*** | | | | |
| & abs(OI NAV-mean OI NAV)<1 st dev | (16.252) | | | | |
| OI ETF-mean(OI ETF)<-1 st dev | -0.001*** | | | | |
| & abs(OI NAV-mean OI NAV)<1 st dev | (-14.971) | | | | |
| Calendar day fixed effects | -0.000 | 0.000 | -0.000 | 0.000 | 0.000 |
| | (-1.126) | (0.440) | (-0.351) | (0.652) | (0.434) |
| Observations | 386,026 | 385,958 | 86,041 | 385,969 | 86,041 |
| Adj. $R^2$ | 0.004 | 0.001 | 0.005 | 0.072 | 0.127 |

## Table 7. Evidence for Arbitrage Activity: Change in ETF Shares and ETF Mispricing

The table presents regressions using ETF-day-level data. The sample is all equity ETFs between 2001 and 2010. The dependent variable is the daily rate of change in ETF shares (in %). The independent variables include: lagged NAV return, lagged ETF return, lagged ETF mispricing. All regressions are OLS regressions. Calendar day fixed effects are included in all regressions, and ETF fixed effects are included in Column (2). Standard errors are clustered at the ETF level. *t*-statistics are presented in parentheses. ***, **, * represent statistical significance at the 1%, 5%, or 10% levels, respectively.

| | Δ ETF Shares (%) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| ETF mispricing(t-1) | 0.048*** | 0.041*** | 0.057*** | 0.050*** |
| | (13.526) | (10.780) | (15.506) | (12.772) |
| NAV Ret(t-1) | | | 0.014*** | 0.012*** |
| | | | (4.763) | (4.429) |
| ETF Ret(t-1) | | | -0.017*** | -0.016*** |
| | | | (-6.718) | (-6.431) |
| Calendar day fixed effects | Yes | Yes | Yes | Yes |
| Fund fixed effects | No | Yes | No | Yes |
| Observations | 514,794 | 514,794 | 514,794 | 514,794 |
| Adj. $R^2$ | 0.000 | 0.006 | 0.000 | 0.006 |

## Table 8. ETF Mispricing and Order Imbalance

The table presents regressions exploring the relation between ETF mispricing and buy-sell order imbalance at the ETF and underlying stocks. The table focuses on buy-sell order imbalance at the ETF level (Columns (1) to (3)) and at the underlying assets (Columns (4) to (6)) on past ETF mispricing and past ETF order imbalance. All regressions are OLS regressions. *t*-statistics are presented in parentheses. ***, **, * represent statistical significance at the 1%, 5%, or 10% levels, respectively.

| Dependent variable: | Buy-sell order imbalance (t+1) of… | | | | | |
|---|---|---|---|---|---|---|
| | ETFs | | | Underlying assets | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| ETF mispricing (t) | -0.873** | -0.873*** | -0.873** | 0.167** | 0.167** | 0.167* |
| | (-2.498) | (-2.596) | (-2.102) | (2.384) | (2.551) | (1.789) |
| ETF order imbalance (t) | 0.135*** | 0.135*** | 0.135*** | 0.005*** | 0.005*** | 0.005*** |
| | (51.147) | (41.497) | (37.178) | (10.634) | (11.499) | (8.321) |
| ETF mispricing (1-10 lags) | Yes | Yes | Yes | Yes | Yes | Yes |
| ETF order imbalance (1-10 lags) | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 365,414 | 365,414 | 365,414 | 366,308 | 366,308 | 366,308 |
| Adj. $R^2$ | 0.130 | 0.130 | 0.130 | 0.030 | 0.030 | 0.030 |
| Error clustering: | Date | ETF | Date & ETF | ETF | Date | Date & ETF |

## Table 9. Predictions for S&P 500 Stocks

The table presents regressions of future stock returns on current ETF mispricing interacted with stock characteristics. The sample is restricted to stock-days that are included in the S&P 500 index. The dependent variable is stock returns (%). All regressions are OLS regressions. *t*-statistics are presented in parentheses. ***, **, * represent statistical significance at the 1%, 5%, or 10% levels, respectively.

| Dependent variable: | Stock Return (t+1) (%) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| ETF mispricing (t) × log(Market capitalization) | 0.180** | | | 0.128 |
| | (2.118) | | | (1.422) |
| log(Market capitalization) | -0.002*** | | | -0.003*** |
| | (-14.689) | | | (-18.944) |
| ETF mispricing (t) × Beta | | 0.252 | | 0.384* |
| | | (1.131) | | (1.684) |
| Beta | | -0.008*** | | -0.008*** |
| | | (-21.028) | | (-18.661) |
| ETF mispricing (t) × Idiosyncratic volatility | | | -3.846 | -6.049** |
| | | | (-1.442) | (-2.248) |
| Idiosyncratic volatility | | | -0.079*** | -0.056*** |
| | | | (-14.823) | (-10.132) |
| Calendar day fixed effects | Yes | Yes | Yes | Yes |
| Observations | 1,250,138 | 1,250,385 | 1,242,366 | 1,242,366 |
| Adj. $R^2$ | 0.118 | 0.119 | 0.119 | 0.120 |

## Table 10. ETF Mispricing, Arbitrage Activity, and Stock Volatility

The table presents regressions using stock-day-level data. Panel A presents regressions of changes in daily volatility at month *t*, on changes in ETF ownership, and interactions with stock size. Panel B presents regressions of changes in daily volatility at month *t*, and changes in monthly turnover on counter of ETFs starting covering the stock, counter of ETFs stopping covering the stock and stock characteristics. All regressions are OLS regressions. Calendar day fixed effects are included in all regressions, and ETF fixed effects are included in Columns (2), (4), (6), and (8). Standard errors are clustered at the stock level. *t*-statistics are presented in parentheses. ***, **, * represent statistical significance at the 1%, 5%, or 10% levels, respectively.

### Panel A: Effects of ETF Ownership on Volatility, per Stock Size

| | Change in volatility | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| I(small stock) × Change in ETF weight | | | 0.088*** | 0.088*** |
| | | | (4.160) | (4.128) |
| Change in ETF weight | 0.030** | 0.031*** | -0.012 | -0.011 |
| | (2.556) | (2.618) | (-0.946) | (-0.876) |
| I(small stock) | | | 0.012*** | 0.029*** |
| | | | (2.733) | (3.161) |
| Change in institutional ownership | 0.008*** | 0.008*** | 0.007*** | 0.007*** |
| | (6.555) | (6.818) | (5.535) | (5.475) |
| I(small stock) × Change in institutional ownership | | | 0.001 | 0.002 |
| | | | (0.394) | (0.793) |
| Stock fixed effects | No | Yes | No | Yes |
| Calendar day fixed effects | Yes | Yes | Yes | Yes |
| Observations | 431,807 | 431,807 | 431,792 | 431,792 |
| Adj. $R^2$ | 0.101 | 0.102 | 0.101 | 0.102 |
| Number of stocks | | 9,279 | | 9,279 |

**Table 10. ETF Mispricing, Arbitrage Activity, and Stock Volatility (Cont.)**

**Panel B: Volatility, Turnover, and Introduction/Exit of ETFs**

| | Monthly change in daily volatility (%) | | Monthly change in turnover | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| # ETFs first reporting to hold the stock | 0.016*** | 0.019*** | 0.001*** | 0.001*** |
| | (7.455) | (8.286) | (8.570) | (8.376) |
| # ETFs last reporting to hold the stock | -0.038*** | -0.047*** | -0.003*** | -0.003*** |
| | (-5.888) | (-6.342) | (-6.261) | (-4.904) |
| Change in institutional ownership | 0.001 | 0.001 | 0.000*** | 0.000*** |
| | (1.079) | (0.762) | (6.032) | (4.823) |
| log(market capitalization) | -0.184*** | -0.419*** | -0.005*** | -0.032*** |
| | (-39.894) | (-35.650) | (-26.476) | (-48.238) |
| lag(daily volatility) | -0.471*** | -0.617*** | -0.001*** | -0.003*** |
| | (-148.765) | (-183.612) | (-9.200) | (-15.275) |
| lag(turnover) | 1,315.918*** | 321.618*** | -137.460*** | -329.588*** |
| | (43.178) | (7.443) | (-52.180) | (-83.741) |
| # ETFs reporting to hold the stock | 0.003*** | 0.004*** | 0.000*** | 0.000** |
| | (4.380) | (4.248) | (7.464) | (2.411) |
| Calendar month fixed effects | Yes | Yes | Yes | Yes |
| Stock fixed effects | No | Yes | No | Yes |
| Observations | 428,205 | 428,205 | 424,989 | 424,989 |
| Adj. $R^2$ | 0.289 | 0.381 | 0.060 | 0.075 |
| Number of stocks | | 9,269 | | 9,234 |

## Table 11. Stock Volatility and the Introduction of New ETFs

The table presents diff-in-doff regressions using stock-level data around the introduction of new ETFs. For each stock, the regressions include 2 observations: one for the month prior to the introduction of ETFs on the S&P 500 index (IVV or VOO), and one for the month following the introduction. The All regressions are OLS regressions. *Post introduction* is an indicator variable for the month following the introduction of the ETF. *Stock in index* is an indicator variable to whether a stock is included in the index. Standard errors are clustered at the stock level. ***, **, * represent statistical significance at the 1%, 5%, or 10% levels, respectively.

| | Dependent variable: Daily volatility (%) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Sample: ±1 month around | Introduction of IVV | | | Introduction of VOO | | | Introduction of IVV, VOO | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Post introduction × Stock in index | 0.571*** | 0.475*** | 0.427*** | 0.176*** | 0.183*** | 0.227*** | 0.508*** | 0.359*** | 0.328*** |
| | (6.955) | (6.090) | (5.757) | (3.305) | (3.689) | (5.120) | (9.896) | (7.835) | (7.458) |
| Post introduction | -1.464*** | -1.365*** | -1.408*** | -0.404*** | -0.250*** | -0.398*** | -1.073*** | -3.226*** | -0.257*** |
| | (-36.036) | (-36.000) | (-37.876) | (-11.872) | (-7.622) | (-7.395) | (-36.755) | (-61.490) | (-7.220) |
| Stock in index | -2.566*** | -1.897*** | | -1.407*** | 0.157** | | -2.426*** | -1.267*** | |
| | (-24.854) | (-15.168) | | (-28.588) | (2.143) | | (-34.514) | (-15.494) | |
| Institutional ownership ratio | | -3.591*** | -0.799*** | | -0.787*** | -0.159 | | -2.640*** | -0.713*** |
| | | (-21.238) | (-2.617) | | (-8.950) | (-0.340) | | (-25.687) | (-2.684) |
| Market cap | | -0.852*** | -1.024*** | | -0.778*** | 0.218 | | -0.793*** | -0.737*** |
| | | (-26.927) | (-6.351) | | (-28.668) | (0.763) | | (-34.327) | (-5.329) |
| Turnover | | 1.180*** | 1.439*** | | 0.453*** | 1.326*** | | 0.911*** | 1.413*** |
| | | (44.311) | (24.670) | | (21.046) | (15.265) | | (48.957) | (29.037) |
| Month fixed effects | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |
| Stock fixed effects | No | No | Yes | No | No | Yes | No | No | Yes |
| Observations | 13,127 | 13,092 | 13,092 | 8,004 | 7,973 | 7,973 | 21,131 | 21,065 | 21,065 |
| Adj. $R^2$ | 0.069 | 0.324 | 0.301 | 0.061 | 0.352 | 0.255 | 0.060 | 0.406 | 0.291 |
| Number of stocks | | | 6,687 | | | 4,029 | | | 10,716 |

## Table 12. Flash Crash: S&P 500 Return and Order Imbalance

The table presents regressions using second-level data. Panel A presents regressions of second-level S&P 500 returns on May 6th, 2010, on lagged variables: SPY mispricing, S&P 500 return, SPY return, E-mini futures return, as well as cumulative returns. In Panel B, the independent variable is order imbalance (calculated using the Lee and Ready (1991) algorithm). In Panel C, the independent variable is average short selling volume in the following 5 seconds. All regressions are OLS regressions. Calendar day fixed effects are included in all regressions, and ETF fixed effects are included in Columns (2), (4), (6), and (8). *t*-statistics are presented in parentheses. ***, **, * represent statistical significance at the 1%, 5%, or 10% levels, respectively.

| Dependent variable: | Return S&P500 (t) | | |
|---|---|---|---|
| Sample: | Before trough 14:30:00 - 14:45:45 | | |
| | (1) | (2) | (3) |
| SPY mispricing (t-1) | 0.064*** | 0.055*** | 0.025* |
| | (10.396) | (4.710) | (1.812) |
| E-mini mispricing (t-1) | | | 0.073*** |
| | | | (4.887) |
| Cum. Ret. S&P500 (t-1, t-60) | | 0.082*** | 0.126*** |
| | | (5.516) | (7.344) |
| Cum. Ret. SPY (t-1, t-60) | | -0.011 | -0.004 |
| | | (-0.814) | (-0.327) |
| Cum. Ret. Emini (t-1, t-60) | | -0.058*** | -0.110*** |
| | | (-4.309) | (-6.813) |
| Cum. Ret. S&P500 (t-1, t-600) | | -0.000 | -0.007 |
| | | (-0.041) | (-1.172) |
| Cum. Ret. SPY (t-1, t-600) | | -0.003 | -0.002 |
| | | (-0.482) | (-0.349) |
| Cum. Ret. Emini (t-1, t-600) | | 0.007 | 0.011** |
| | | (1.187) | (1.983) |
| Constant | -0.000*** | 0.000 | 0.000 |
| | (-8.167) | (1.219) | (0.219) |
| Observations | 945 | 943 | 937 |
| Adj. $R^2$ | 0.102 | 0.189 | 0.223 |

**Figure 1: Non-Fundamental Shocks Are Propagated Via Arbitrage**

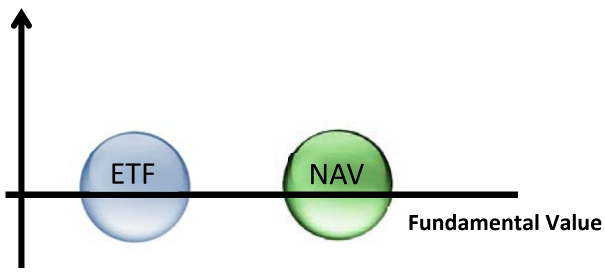


Figure 1a. Initial Equilibrium

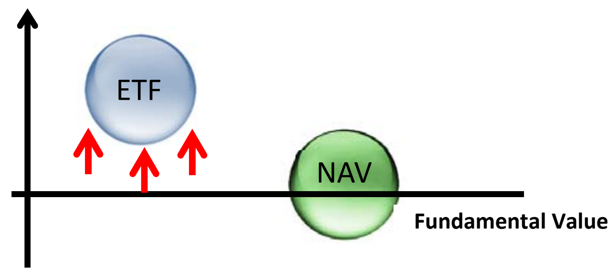


Figure 1b. Non-Fundamental Shock to ETF

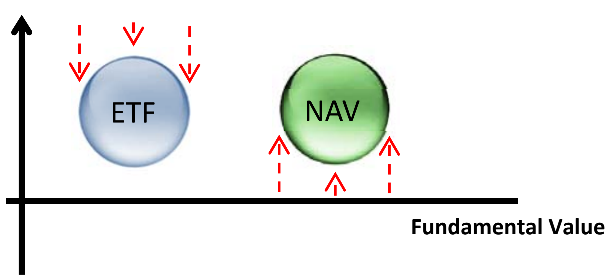


Figure 1c. Initial Outcome of Arbitrage: the non-fundamental shock is propagated to the NAV, the ETF price starts reverting to Fundamental Value

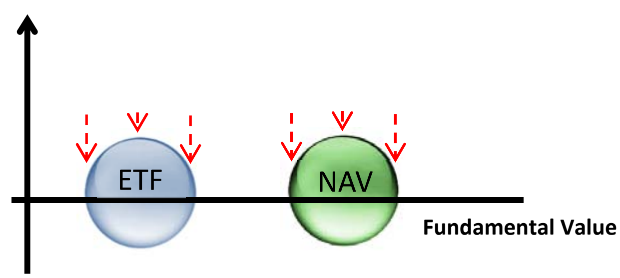


Figure 1d. Re-establishment of Equilibrium: after some time both the ETF price and the NAV revert to Fundamental Value

**Figure 2: Fundamental Shock with Price Discovery Occurring in the ETF Market: the ETF moves first, the NAV Follows with Some Delay**

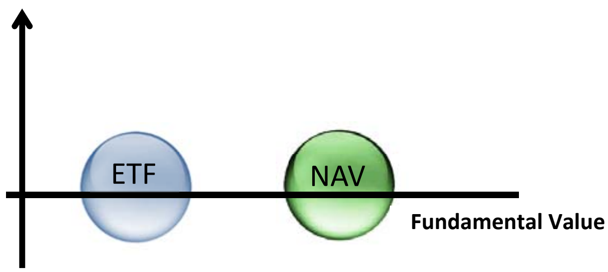


Figure 2a. Initial Equilibrium

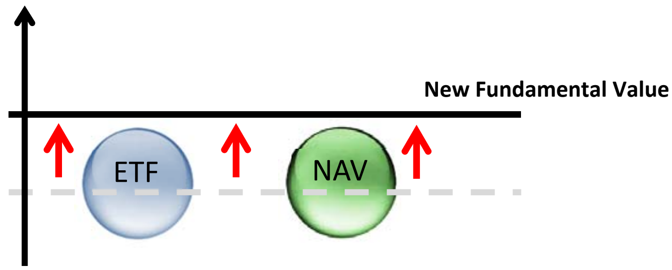


Figure 2b. Shock to Fundamental Value

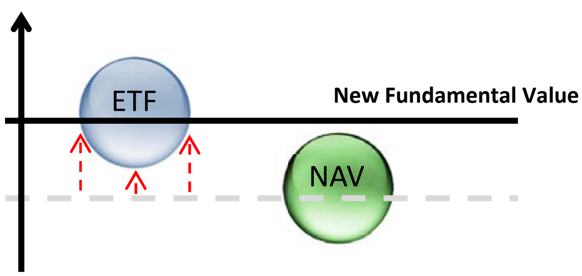


Figure 2c. The ETF price moves to the New Fundamental Value

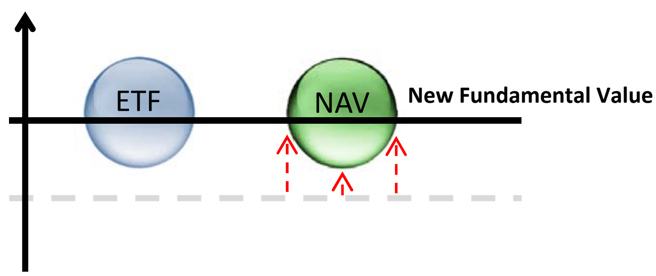


Figure 2d. The NAV catches up with a delay with the New Fundamental Value

**Figure 3. ETF Growth in the U.S.**

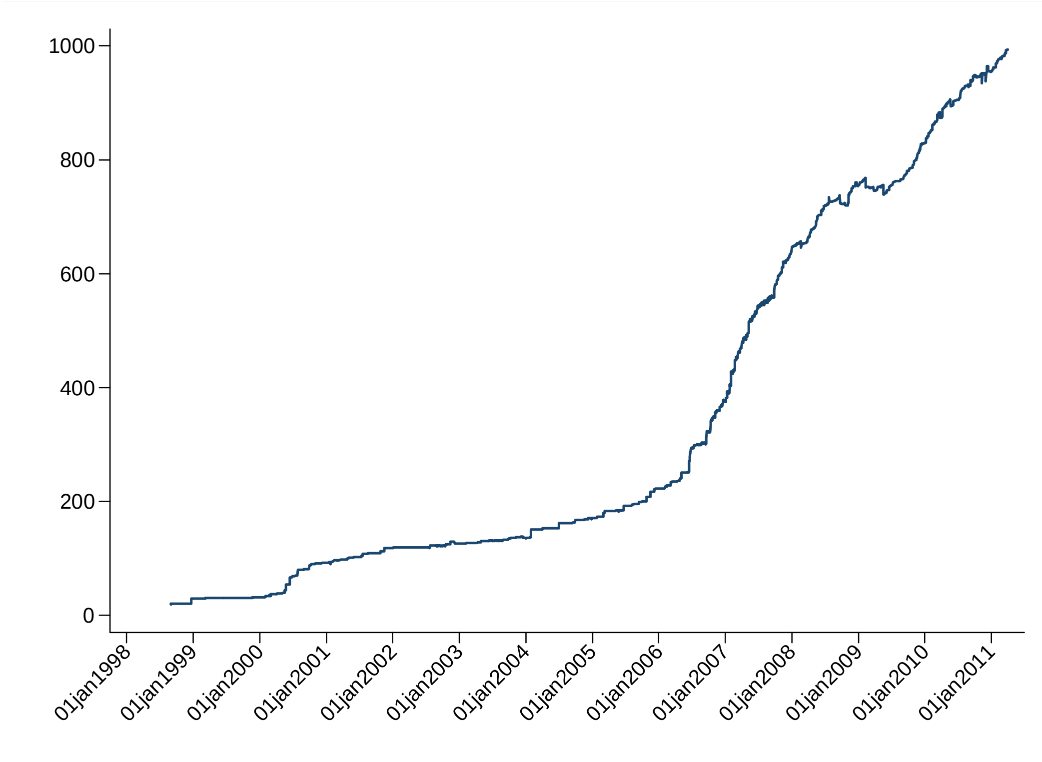

**Figure 4. Time Series of ETF Mispricing**

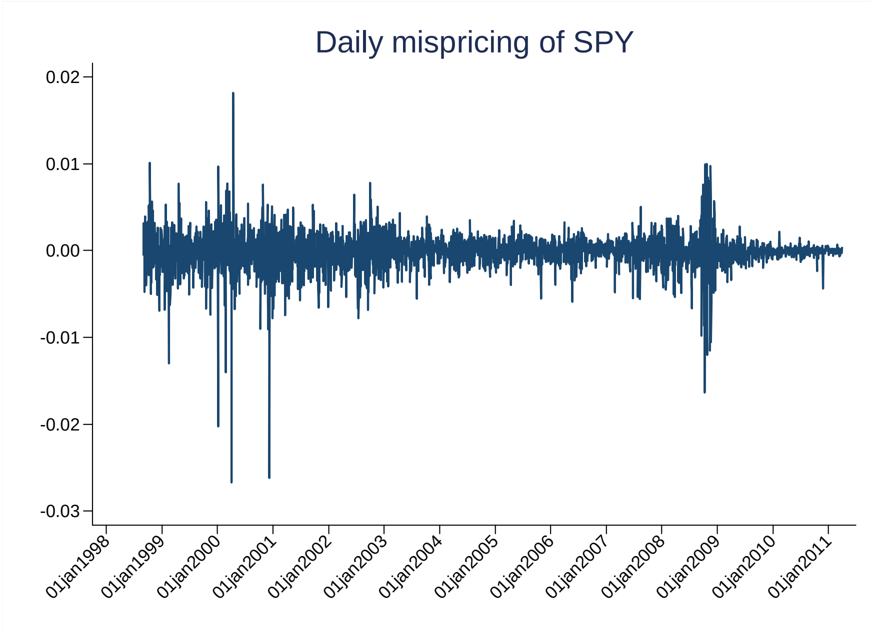


Figure 4a. Example of ETF mispricing: SPY

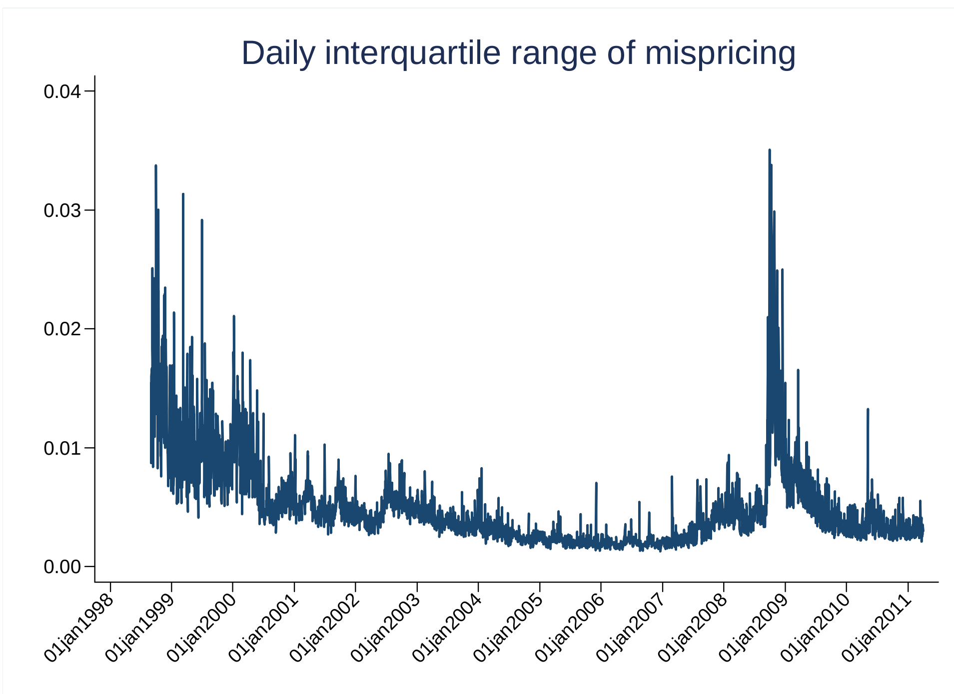


Figure 4b. Daily interquartile range of mispricing

**Figure 4. Time Series of ETF Mispricing (Cont.)**

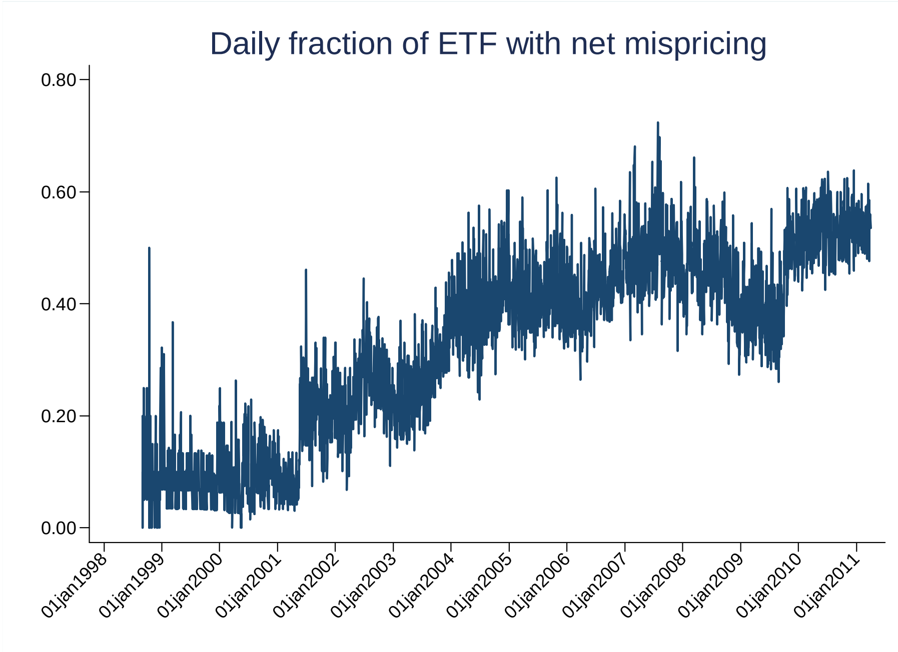


Figure 4c. Daily fraction of firms with positive net mispricing, which is the difference between the absolute value of mispricing and the bid-ask spread

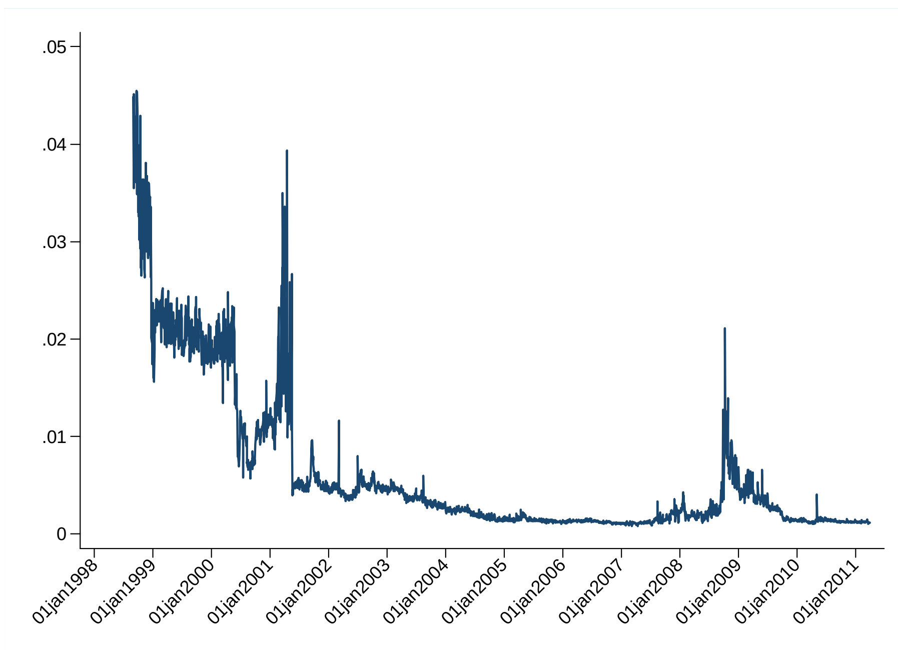


Figure 4d. Daily median bid-ask spread

**Figure 5. Impulse Response Functions of a Shock of Mispricing on Future NAV Returns**

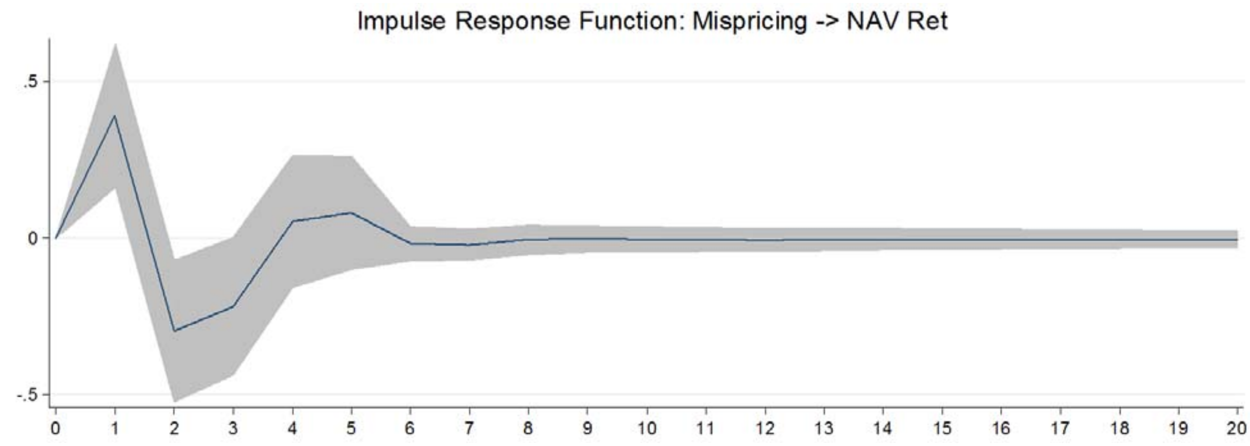


Figure 5a.

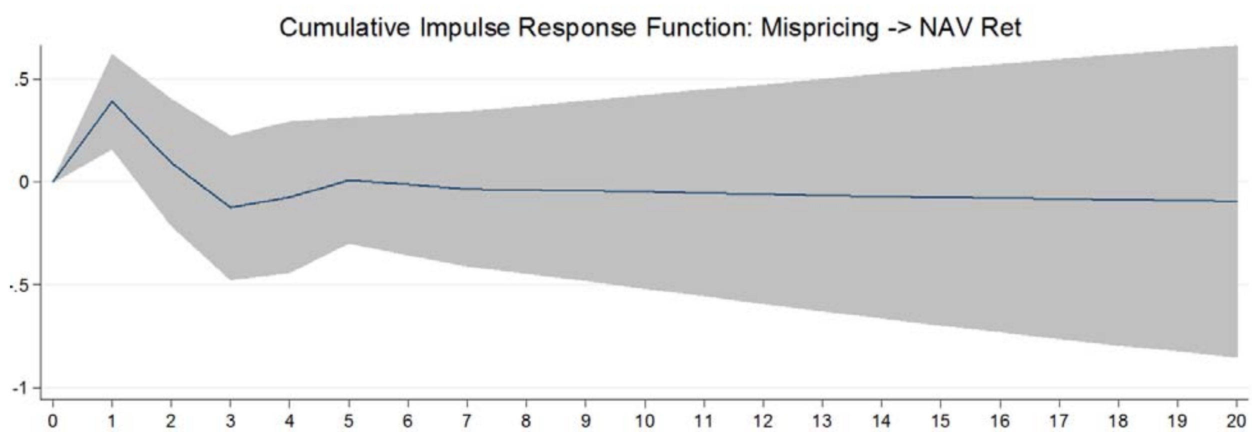


Figure 5b.

**Figure 6. S&P 500 E-mini futures, S&P 500 (NAV), and SPY (ETF) in the Flash Crash**

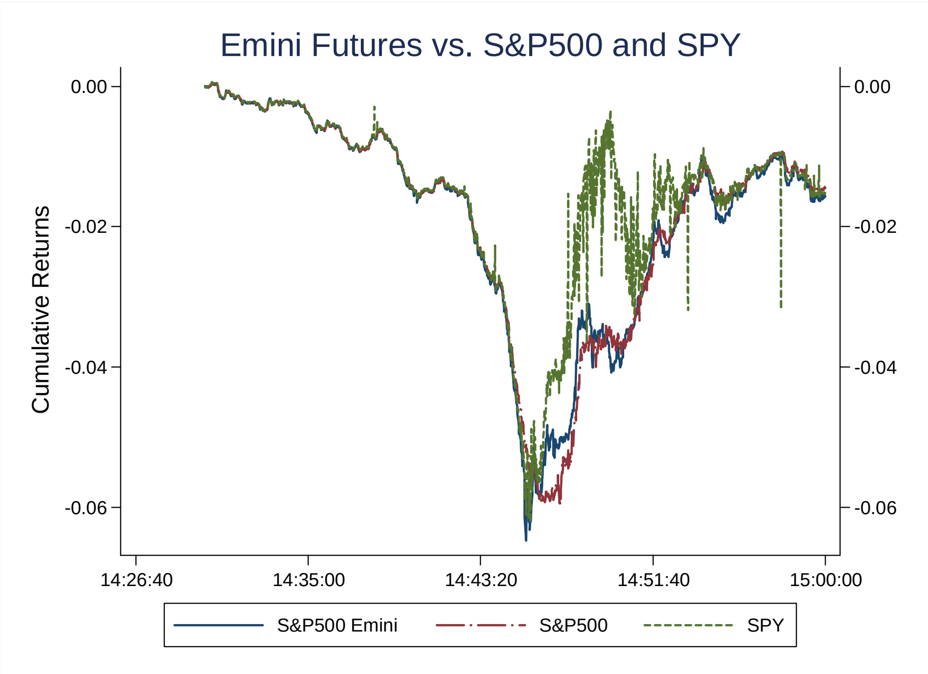


Figure 6a. S&P 500 E-mini futures, S&P 500 (NAV), and SPY (ETF) in May 6, 2010.

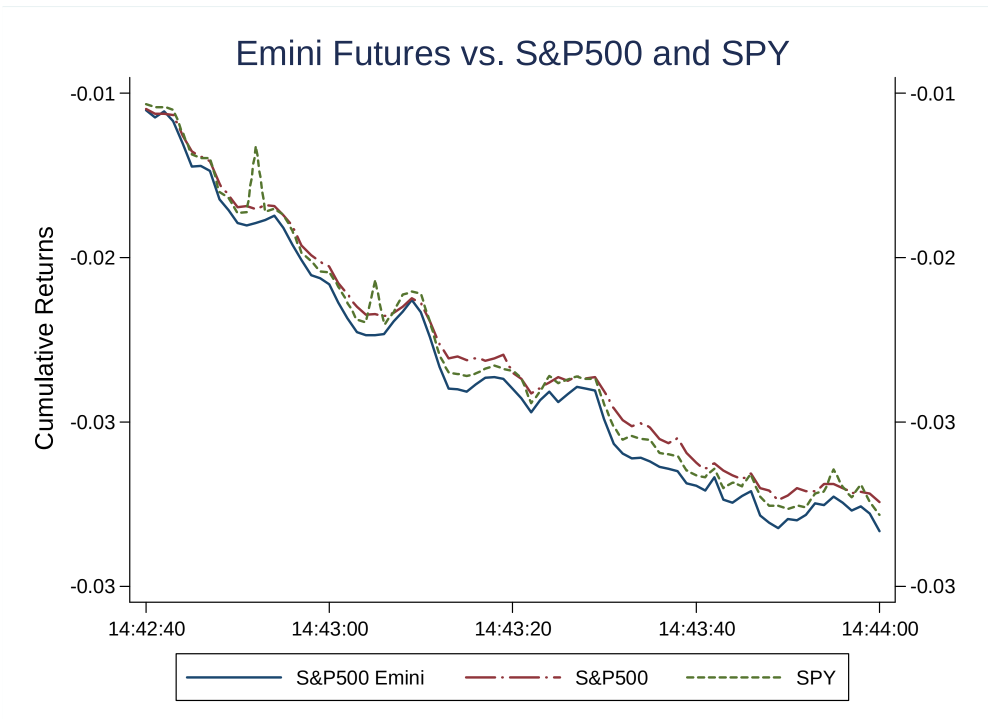


Figure 6b. S&P 500 E-mini futures, S&P 500 (NAV), and SPY (ETF) in May 6, 2010, 14:42:40 to 14:44:00.

**Figure 6. S&P 500 E-mini futures, S&P 500 (NAV), and SPY (ETF) in the Flash Crash (Cont.)**

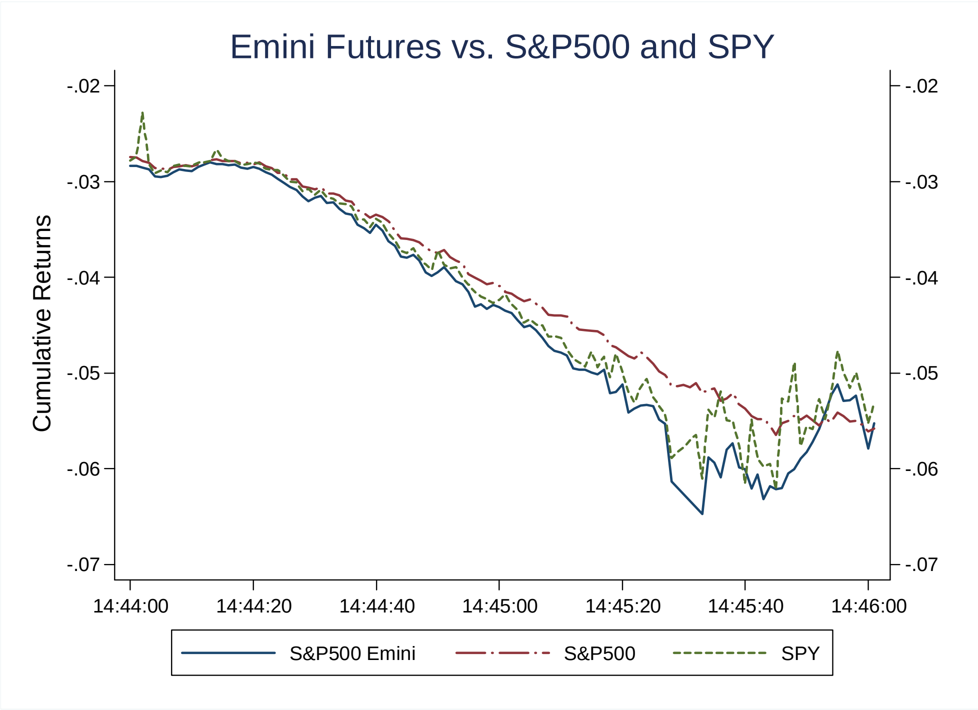


Figure 6c. S&P 500 E-mini futures, S&P 500 (NAV), and SPY (ETF) in May 6, 2010, 14:44:00 to 14:46:00.