Working Paper Series 738
(ISSN 2788-0443)

Effect of Exchange-Traded Funds Arbitrage Transactions on their Underlying Holdings

Gregory Boadu-Sebbe

CERGE-EI
Prague, October 2022

ISBN 978-80-7343-545-5 (Univerzita Karlova, Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-654-3 (Národohospodářský ústav AV ČR, v. v. i.)

Effect of Exchange-Traded Funds Arbitrage Transactions on their Underlying Holdings

Gregory Boadu-Sebbe

CERGE-EI*

September 20, 2022

Abstract

A critical aspect of trading Exchange-Traded Funds (ETFs) is the arbitrage trading strategy taken by authorized participants (APs) to keep ETF prices in line with their net asset values (NAVs). ETF arbitrage trading is a strategy that exploits the discrepancies between an ETF price and the value of the ETF's underlying assets. In this study, I quantitatively examine the effect of ETF arbitrage on the underlying assets of an ETF. I develop a dynamic state-space model that jointly estimates the price dynamics of an ETF and its underlying assets by explicitly incorporating the ETF arbitrage. The model is estimated individually for the Dow Jones Industrial Average ETF (DIA) and the VanEck Vectors Semiconductor ETF (SMH). The empirical results show that ETF liquidity shocks propagate to the underlying assets via the ETF arbitrage mechanism. These ETF liquidity shocks add a permanent layer of transitory volatility to the underlying asset prices. I find that a unit of liquidity shock to DIA brings a range of 0.1% to 0.93% of extra volatility to the underlying assets of DIA. Similarly, a unit of liquidity shock to SMH adds a range of 0.33% to 0.95% of additional volatility to the underlying assets. In addition, I show that it takes APs longer to correct deviations between the ETF price and its NAV. It takes approximately 4 and 10 minutes for APs to perform the ETF arbitrage for DIA and SMH, respectively. Finally, the findings suggest that an ETF arbitrage transaction speeds up the price discovery process in the ETF markets. There are approximately 74% and 67% variations in the premiums of DIA and SMH due to price discovery, respectively.

JEL codes: B26, C32, C58, E44, G12

Keywords: Exchange-Traded Funds, Underlying assets, ETF arbitrage mechanism, Liquidity shocks, Net asset value (NAV), Price discovery process

*CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences, Politickych veznu 7, 111 21 Prague, Czech Republic.

1. Introduction

Over the last two decades, Exchange Traded Funds (ETFs) have grown in the global financial markets. There are significant benefits to investing in ETFs over traditional index funds such as open-end mutual and hedge funds. The benefits for ETF investors include increased and continuous access to intraday liquidity, low trading costs, lower tax burden, and access to a broader set of markets (Ben-David et al., 2017, 2018). Generally, an ETF operates with an arbitrage mechanism designed to keep its market value in line with its net asset value (NAV) when a deviation between the ETF price and its NAV occurs. ETF arbitrage trading is a strategy that exploits the discrepancies between an ETF price and the value of the ETF's underlying assets.

However, one may wonder whether the ETF arbitrage mechanism has consequences for the underlying assets in the ETFs' portfolios. The liquidity of an ETF attracts high-frequency demand, which is likely to affect the prices of the underlying assets since the ETF arbitrage mechanism links the ETF and its underlying assets (Agarwal et al., 2018; Dannhauser, 2017; Evans et al., 2018). If there is a demand shock to an ETF, the underlying assets may be exposed to a new layer of demand shock, making them more volatile.

In this paper, I quantitatively examine the effect of ETF arbitrage transactions on the underlying assets of an ETF. In particular, I measure the intensity of ETF liquidity shocks on the prices of underlying assets of an ETF. I seek to determine how much excess volatility propagates to the underlying assets when there is a liquidity shock to the ETF market. In addition, I investigate whether the ETF arbitrage mechanism is a medium through which ETF liquidity shocks propagate to the underlying asset markets. I also examine the contribution of the ETF arbitrage mechanism in the price discovery process of an ETF, i.e., whether ETF arbitrage trading speeds up or slows down the price discovery process in the ETF market.

I develop a dynamic state-space model that jointly estimates the price dynamics of an ETF and the underlying assets of the ETF. The model is based on three main structures: ETF market structure, underlying asset market structures, and ETF arbitrage component linking the ETF to each underlying asset. The model defines the ETF arbitrage mechanism as the difference between the ETF price and its net asset value (NAV)[1]. The ETF arbitrage transaction is an essential component of the model performed by authorized participants (APs) to keep the ETF price in line with its NAV. For instance, if there is a high demand for ETF shares, this increases the price of the ETF and deviates upwards from its NAV. To correct such deviations, an AP creates more ETF shares by trading the 'basket' of the underlying assets with the ETF firm to obtain more of the ETF shares. The AP introduces the new ETF shares on the exchange market. The new shares increase the supply of the ETF, which puts pressure on the ETF price to decrease until it falls in line with the NAV. The underlying assets are likely to inherit the demand shock from the ETF through the trading strategy by the AP.

To capture information about liquidity, the model incorporates bid and ask quotes for

[1] Intuitively, the market value of an ETF and its NAV is based on the unobserved efficient value of the fund. However, the ETF NAV is the value of each share portion of its underlying assets and cash at the end of the trading day. The NAV of an ETF depends on each of the efficient values of its underlying assets. Therefore, a deviation between an ETF price and its NAV equals the difference between the ETF price and the weighted sum of efficient values of the underlying assets.

the ETF and each underlying asset. I then decompose the ask and bid quotes for the ETF and the underlying asset markets into efficient prices and transitory components. The efficient prices are assumed to follow a random walk. The transitory components are assumed to have autocorrelation with themselves and each other. The model allows for permanent and transitory shocks from the ETF and the underlying asset markets. In this way, one can measure the impact and intensity of different shocks in the ETF market on the underlying assets and vice versa. The permanent (informational) and transitory conditional volatility follow EGARCH models.

The purpose of jointly modeling the dynamics of an ETF and its underlying assets and linking them via ETF arbitrage is 1) to be able to determine the channel that propagates shocks from the ETF market to the underlying assets. Hence, one can measure any additional volatility brought to the underlying asset prices via the channel; 2) to determine the speed APs may take to perform the ETF arbitrage transaction when there is a deviation between an ETF price and its NAV; 3) to decompose the ETF arbitrage component into transitory and price discovery components. In this way, one can measure how ETF arbitrage transactions contribute to the price discovery process in the ETF market.

Therefore, the main contribution of the dynamic model is that it can measure the intensity of any liquidity shock to an ETF via the ETF arbitrage mechanism on the underlying assets. In addition, the model can determine whether the ETF arbitrage trading speeds up or slows down the price discovery process in an ETF market.

The model implementation uses tick data for the Dow Jones Industrial Average ETF (DIA) and VanEck Vectors Semiconductor ETF (SMH) listed on NYSE. DIA and SMH have 30 and 25 underlying assets in their respective portfolios. The datasets span the first quarter of 2018 with milliseconds sample frequency.

The empirical results show that liquidity shocks to both DIA and SMH propagate to their respective underlying assets. This result is obtained from the variance of the deviations between the ETF price and its NAV. These deviations contemporaneously correlate with their underlying assets' returns, where a fraction of the volatility of liquidity shocks to DIA and SMH ETFs find their way into the underlying assets' returns. This finding shows that the ETF arbitrage (deviations between the ETF price and its NAV) is the channel that propagates the liquidity shocks to the underlying assets of DIA and SMH.

I find that if there is a liquidity shock to DIA or SMH, the prices of their underlying assets become more volatile in the next period. Thus, if there is one unit of liquidity shock to DIA ETF, then the ETF arbitrage mechanism propagates a range of 0.1% to 0.93% of additional volatility to DIA's underlying assets in the next period. Similarly, a unit of liquidity shock to SMH adds a range of 0.33% to 0.94% of extra volatility to SMH's underlying assets in the next period. I find that the underlying assets with higher weights in their respective ETF portfolios receive more volatility than the less weighted assets.

Furthermore, the results show that APs may take longer to correct deviations between an ETF price and its NAV. In the case of the Dow Jones Industrial Average ETF (DIA) and VanEck Vectors Semiconductor ETF (SMH), the average time an AP may correct any parity between the prices of the ETFs and their NAVs is approximately 4 and 10 minutes, respectively.

Finally, the findings of this study suggest that an ETF arbitrage transaction speeds up the price discovery process in an ETF market. The estimates of the price discovery components for DIA and SMH ETFs are 74% and 67%, respectively. Intuitively, these estimates make sense because approximately 74% and 67% of the variations in the premiums (ETF arbitrage) of DIA and SMH are due to price discovery.

An important question is whether ETFs are beneficial or harmful to the financial markets. One cannot take a stand or provide a straightforward answer to this question. However, based on the findings of this study, one could point out periods in which ETFs may be harmful or beneficial. During standard times like the sampled period this study considers, a high liquidity shock to ETFs motivates investors to place large short-term directional bets on the 'basket' of the underlying assets. The high liquidity shock increases the underlying assets' volatility and co-movement. That can be problematic in times of financial crisis, particularly for ETFs with illiquid underlying assets and trade OTC. During periods of market stress, ETFs' prices may deviate from their underlying asset values, which may destabilize financial institutions that depend on ETF shares for managing their liquidity. However, the deviation between the prices of ETFs and the values of their underlying assets may decrease any repercussions of market stress on the underlying assets. Hence, if financial institutions are frequenctly exposed to the risk associated with the underlying assets, the deviation between the prices of ETFs and the values of their underlying assets may decrease instead of increase systematic risk.

This study relates to recent analyses of the effect of trading ETFs on their underlying assets. Ben-David et al. (2018) argue that the ETF arbitrage transaction mechanism is a channel through which liquidity shocks from the ETF markets propagate to the underlying assets[2]. The authors further argue that the ETF liquidity shocks might add extra non-fundamental volatility to underlying asset prices. My study investigates the arguments by Agarwal et al. (2018), where I quantitatively show that ETF liquidity shocks propagate to the underlying assets through the ETF arbitrage mechanism. In addition, my model helps to quantify the intensity of any liquidity shocks from an ETF market on the underlying assets. Thus, one can measure the extra non-fundamental volatility that the ETF liquidity shocks, via the ETF arbitrage, add to the prices of the underlying assets.

Contrary to Ben-David et al. (2018), Box et al. (2019), who analyze the ask and bid quotes on the 1-minute frequency of ETFs, find that an ETF arbitrage transaction is not responsible for the propagation of liquidity shocks to the underlying assets. They argue that parity between ETF prices and their NAVs are restored via the adjustments in the ask and bid quotes of the underlying assets. However, the model considered by Box et al. (2019) does not capture the ETF arbitrage mechanism and the price dynamics of the underlying assets of an ETF. Perhaps this may have contribute to their findings of

[2] Agarwal et al. (2018) show that ETFs create demand pressure on the prices of their underlying assets. The authors argue that the noise or shocks from the high demand for ETF shares may affect the prices of the underlying assets. Furthermore, related studies focus on ETF discounts and premiums. These studies include Engle and Sarkar (2006), Hilliard (2014), Lin and Chou (2006), and Charteris et al. (2014). A related strand of literature focuses on the co-movement relationship between ETFs and their underlying assets. The general conclusion from this literature is that there are co-movements between ETFs and their underlying assets, see e.g. Da and Shive (2013), Madhavan and Sobczyk (2016), Israeli et al. (2017), Nam (2017), Dannhauser (2017), Agarwal et al. (2018), Evans et al. (2018), and Broman and Shum (2018).

no-arbitrage pressures on the underlying asset prices.

The remainder of this paper is structured as follows: Section 2 presents an overview of ETF trading and pricing[3]. Section 3 develops a dynamic state-space model that combines an ETF, price dynamics of the underlying assets, and the ETF arbitrage transaction mechanism. This section also describes the data and provides descriptive statistics. Section 4 presents the empirical results and Section 5 concludes.

2. Background of ETF Trading

2.1 Creation/Redemption Process

ETFs are investment firms whose shares are traded continuously during intraday trading hours. They track specified indices and are traded like stocks on organized exchanges. To some extent, ETFs behave like closed-end mutual funds, i.e., ETF prices can deviate from their NAVs. Retailers and institutional investors mostly trade ETF shares. In contrast to closed-end mutual funds, new ETF shares can be created, and existing shares can be redeemed. As documented in the literature, the price of an ETF can deviate from its NAV, for example, refer to Pontiff (1996). Contracts have been signed between institutions acting as authorized participants (APs) and ETF firms to trade ETF shares and, for that matter, to perform the ETF arbitrage. The primary role of an AP is to create or redeem ETF shares when the need arises.

Creation process: When an ETF price deviates upwards from its NAV (ETF trading at a premium), there is a high demand for the EFT shares, which increases the price of the shares. To keep the ETF price in line with its NAV, an AP creates more shares of the ETF by assembling additional underlying assets with proper weighting and, in exchange, obtains shares of the new ETF from the ETF firm. The AP then introduces the new ETF shares on the exchange market to increase the supply of the ETF shares. In this way, pressure is exerted on the ETF price to move downwards, simultaneously leading to an increase in the NAV.

Redemption process: Similarly, if an ETF price deviates downwards from its NAV (ETF trading at a discount), APs redeem the ETF shares from the exchange market since the price of the shares is lower than the expected price. The APs then exchange these redeemed shares for the 'basket' of the underlying assets from the ETF firm. APs can then sell the assets in the underlying asset markets. The later process of ETF redemption pressures the ETF price upwards, which may exert potential negative pressure on the NAV.

2.2 Trading Channels

This study argues that the deviation between an ETF price and its NAV echoes a mixture of essential information and noise. Based on the creation/redemption process, the model in this study depends on two main trading channels, namely, ETF arbitrage and informational trading channels.

ETF Arbitrage Trading: The creation and redemption procedures described above constitute the ETF arbitrage trading strategy which takes place continuously throughout the trading hours. For instance, when there is a high demand for the shares of SPY

3 For a detailed treatment of the growing literature on ETF trading and pricing, see Ben-David et al. (2017).

ETF, the price of SPY will increase and move above the NAV of SPY. An authorized participant (AP) acts to correct the deviation between the price of SPY and its NAV. Since SPY is trading at a premium, the AP will sell the SPY shares received during the creation process on the open market at a higher price. Afterwards, the AP will go into the underlying asset markets to buy their shares at lower prices. This process will help the AP to capture the spread between the cost of the underlying assets he bought for the ETF firm and the selling price of SPY shares [4].

Informational Trading: Informational trading is when investors initiate trading upon available information. Suppose there is an announcement of economic changes by a government, or there is a high demand for SPY ETF. After receiving or observing the news of the shock or changes in the SPY, traders of the underlying assets start to trade their assets in the direction of the initial shock, but not simultaneously. The traders' actions may cause the underlying assets to affect each other due to their interactions. For example, traders of Apple stock (AAPL) observing the news of a shock coming to the financial market will start to trade their assets in the direction of the shock. We soon observe an immediate adjustment to quotes in other stocks without actual trading in response to changes in AAPL.

3. Model

In this study, I develop a dynamic state-space model that combines the best ask/bid quotes of an ETF, the best ask/bid quotes of the ETF's underlying assets, and the ETF arbitrage transaction mechanism. The model also features transitory and informational volatility effects and it assumes that the price of an ETF and the prices of the underlying assets can be decomposed into permanent and transitory components[5]. The model builds on the microstructure price decomposition that separates quoted prices into long-run (efficient) value and transitory non-informative components. Furthermore, the model explicitly incorporates the ETF arbitrage mechanism based on the quoted prices of an ETF and the underlying assets.

Notation: I denote the long-run value of an ETF as v_t and m_t^i as the long-run value of an underlying asset i. The price of an ETF and the price of an underlying asset i are denoted by p_t^f and p_t^i, respectively. Denote a_t^f and b_t^f as the ask and bid quotes of an ETF and a_t^i and b_t^i as the ask and bid quotes for underlying asset i.

Now, suppose there is an ETF that invests in n risky assets, then the price of the ETF in seconds t may relate to the lagged prices of the n risky assets and vice versa. In addition, the price of the i^{th} asset may relate to the lagged price of the other n underlying assets. Hence, the ETF and i^{th} asset's transitory components in seconds

4 Nonetheless, APs can utilize arbitrage trading strategies to realize profits. Although APs can engage in the underlying asset markets, they mainly focus on the ETF market. They buy the less expensive assets and short-sell the expensive ones between the ETF and the 'basket' of underlying assets. APs can realize profits by holding the positions until prices converge, at which point they close their positions. In a strict sense, the uncertainty involved in this profit making does not qualify these trades as an arbitrage (Ben-David et al., 2018).

5 The model I develop resembles the models developed by Madhavan and Sobczyk (2016), Pascual and Veredas (2010), and Menkveld et al. (2007)

t may relate to the lagged transitory components of the other n assets and the ETF during the previous seconds or minute.

3.1 Model Description

3.1.1 Price Dynamics of the Underlying Assets of an ETF

For each underlying asset i, the following equations define its price dynamics: The long-run value of an asset i follows a random walk with drift;

$$m_t^i = \mu_{m^i} + m_{t-1}^i + \epsilon_{m^i,t}. \tag{1}$$

The long-run values of the underlying assets characterize informational arrivals in both the underlying markets and the ETF market. Next, I model the deviations between an underlying asset i price and its long-run value as:

$$\begin{aligned} p_t^i - m_t^i = \nu_{i,p}(p_{t-1}^i - m_{t-1}^i) + \lambda_{i,a}(a_{t-1}^i - p_{t-1}^i) + \lambda_{i,b}(p_{t-1}^i - b_{t-1}^i) + \\ \delta_i(p_{t-1}^f - v_{t-1}) + \epsilon_{p^i,t} \end{aligned} \tag{2}$$

where p_t^i is the mid-quote at time t representing the price of asset i. The price of an underlying asset i is a function of an underlying arbitrage component $(p_{t-1}^f - m_{t-1}^i)$, transitory components of the same underlying asset $(a_{t-1}^i - p_{t-1}^i,\ \ p_{t-1}^i - b_{t-1}^i)$, and the ETF arbitrage component $(p_{t-1}^f - v_{t-1})$. The transitory components for an underlying asset i are also modeled as follows:

$$\begin{aligned} a_t^i - p_t^i = \beta_i + \nu_{i,a}(p_{t-1}^i - m_{t-1}^i) + \sum_{i=1}^{n} \phi_{aa}^{ij}(a_{t-1}^i - p_{t-1}^i) + \sum_{i=1}^{n} \phi_{ab}^{ij}(p_{t-1}^i - b_{t-1}^i) \\ + \theta_{aa^f}^i(a_{t-1}^f - p_{t-1}^f) + \theta_{ab^f}^i(p_{t-1}^f - b_{t-1}^f) + \epsilon_{a^i,t} \end{aligned} \tag{3}$$

$$\begin{aligned} p_t^i - b_t^i = -\beta_i + \nu_{i,b}(p_{t-1}^i - m_{t-1}^i) + \sum_{i=1}^{n} \phi_{ba}^{i}(a_{t-1}^i - p_{t-1}^i) + \sum_{i=1}^{n} \phi_{bb}^{ij}(p_{t-1}^i - b_{t-1}^i) \\ + \theta_{ba^f}^{ij}(a_{t-1}^f - p_{t-1}^f) + \theta_{bb^f}^i(p_{t-1}^f - b_{t-1}^f) + \epsilon_{b^i,t} \end{aligned} \tag{4}$$

where β_i coefficient captures the average distance between the two quotes of each underlying asset i; it must always be positive. Equations (1), (2), (3), and (4) are the price dynamics of an underlying asset i.

3.1.2 Price Dynamics of an ETF

The unobserved long-run value (NAV) of an ETF is measured as the weighted sum of all the efficient values (m_t^i) of the underlying assets, i.e,

$$v_t = \sum_{i=1}^{n} w_i m_t^i \tag{5}$$

where w_i is the weight for underlying asset i. The ETF arbitrage trading mechanism is defined as the difference between the ETF price and its long-run value (NAV).

$$p_t^f - v_t = \sum_{i=1}^{n} \eta_i (p_{t-1}^i - m_{t-1}^i) + \nu_f (p_{t-1}^f - v_{t-1}) + \lambda_{f,a}(a_{t-1}^f - p_{t-1}^f) + \\ + \lambda_{f,b}(p_{t-1}^f - b_{t-1}^f) + \epsilon_{p^f,t} \tag{6}$$

where p_t^f is the mid-quote representing the price of an ETF at time t. The ETF arbitrage component is a function of the lagged underlying arbitrage components $(p_{t-1}^i - m_{t-1}^i)$, lagged ETF arbitrage component $(p_{t-1}^f - v_{t-1})$, and lagged transitory components of the ETF $(a_{t-1}^f - p_{t-1}^f, \ \ p_{t-1}^f - b_{t-1}^f)$.

The transitory components of an ETF are modeled as:

$$a_t^f - p_t^f = \beta_f + \sum_{i=1}^{n} \psi_{a^f a}^i (a_{t-1}^i - p_{t-1}^i) + \sum_{i=1}^{n} \psi_{a^f b}^i (p_{t-1}^i - b_{t-1}^i) + \nu_{f,a}(p_{t-1}^f - v_{t-1}) \\ + \psi_{a^f a^f}(a_{t-1}^f - p_{t-1}^f) + \psi_{a^f b^f}(p_{t-1}^f - b_{t-1}^f) + \epsilon_{a^f,t} \tag{7}$$

$$p_t^f - b_t^f = -\beta_f + \sum_{i=1}^{n} \psi_{b^f a}^i (a_{t-1}^i - p_{t-1}^i) + \sum_{i=1}^{n} \psi_{b^f b}^i (p_{t-1}^i - b_{t-1}^i) + \nu_{f,b}(p_{t-1}^f - v_{t-1}) \\ + \psi_{b^f a^f}(a_{t-1}^f - p_{t-1}^f) + \psi_{b^f b^f}(p_{t-1}^f - b_{t-1}^f) + \epsilon_{b^f,t} \tag{8}$$

where β_f is a positive coefficient that captures the average distance between the two quotes of the ETF and $(\epsilon_{a^f,t}, \epsilon_{b^f,t})$ are ETF supply and demand liquidity shocks, respectively. Therefore, equations (6), (7), and (8) determine the price dynamics of an ETF. The ETF connects with the underlying assets through the ETF arbitrage and transitory components.

3.1.3 Dynamic State-Space Model

Combining equations (1), (2), (3), and (4) for n underlying assets and equations (6), (7), and (8) for an ETF, I obtain the dynamic state-space model. I expect β_i's and β_f to be close to half of the average bid-ask spreads of each underlying asset i and the ETF, respectively. The parameter δ_i measures the effect of the ETF arbitrage mechanism on the price of an underlying asset i. Similarly, θ parameters measure the ETF transitory effect on the transitory components of asset i, ν_f captures the speed of the ETF arbitrage mechanism, and λ_fs capture the transitory effects on the ETF price. Finally, ϕ parameters capture the informational trading effects resulting from a fundamental or a random shock to the underlying asset markets[6]. I define ETF demand and supply ('liquidity') shocks as $\epsilon_{a^f,t}$ and $\epsilon_{b^f,t}$ [7].

[6] The parameter $\nu_{i,p}$ measures the speed of the underlying asset i's arbitrage component $(p_t^i - m_t^i)$ and λ_is measure the transitory effect on the price of the same underlying asset i. η_i measures the effect of underlying asset i arbitrage transactions on the price of an ETF, and ψs capture the transitory effect on the ETF transitory components

[7] Similarly, underlying asset i 'liquidity' shocks are $\epsilon_{a^i,t}$ and $\epsilon_{b^i,t}$, and fundamental shock to underlying asset i markets is denoted by $\epsilon_{m^i,t}$. Finally, $\epsilon_{p^i,t}$ and $\epsilon_{p^f,t}$ are

In matrix form, the dynamic state-space model is given as follows (see Appendix A1 for the full model):

$$\boldsymbol{X}_t = \boldsymbol{A} + \boldsymbol{B}\boldsymbol{X}_{t-1} + \bar{\epsilon}_t \tag{9}$$

where $\boldsymbol{X}_{t-1}$ is $(4n+3) \times \boldsymbol{T}$ vector of lagged variables of equations (1), (2), (3), and (4) for each underlying asset and equations (6), (7), and (8) for the ETF, $\boldsymbol{A}$ is $(4n+3) \times 1$ vector of constants, $\boldsymbol{B}$ is $(4n+3) \times (4n+3)$ matrix of coefficients, and $\bar{\epsilon}_t$ is $(4n+3) \times 1$ vector of noises with $\bar{\epsilon}_t \sim N(0,\ \Sigma_t)$. The number of observations for each lagged variable is T, i.e., $t = 1, 2, \ldots, T$ seconds. The vector of innovations of the dynamic state-space model (9) is assumed to be jointly normally distributed with $\bar{\epsilon}_t \sim N\left(0_{(4n+3)\times 1}, \Sigma_{t_{(4n+3)\times(4n+3)}}\right)$. In the framework of the dynamic state-space model, normality will allow one to estimate the unobserved factors $(m_t^i,\ v_t)$, for $i = 1, 2, \ldots, n$, using an augmented Kalman filter. The model parameters are then estimated using the error prediction decomposition of the augmented Kalman filter via maximum likelihood estimation.

3.2 Model Interpretation

3.2.1 ETF Arbitrage Trading Effects

The high-frequency ETF arbitrage component is defined as the difference between the ETF price and NAV of the ETF. This subsection explains how the ETF arbitrage transaction works in the dynamic state-space model. Suppose there is positive 'liquidity' shock to the ETF market, i.e., $\epsilon_{a^f,t}$, as a result of more investors demanding ETF shares. This leads to an increase in $a_t^f - p_t^f$ and then causes the ETF price p_t^f to deviate from its NAV v_t. An authorized participant (AP), through the creation/redemption process, trades between the ETF shares and the portfolio of the underlying assets to correct the pricing error. The AP buys the 'basket' of the underlying assets and then sells the ETF shares to return the ETF price in line with its NAV. In doing so, the underlying assets may be affected and inherit the demand shock from the ETF market differently. First, there may be a cross-market transient effect in which the increase in $a_t^f - p_t^f$ may affect $a_t^i - p_t^i$ and $p_t^i - b_t^i$ the next period if $\theta_{b^f a}^i$ and $\theta_{b^f b}^i$ are not zero. Secondly, there may be a direct ETF arbitrage effect on the price of an underlying asset i via the arbitrage trading strategy taken by the AP if δ_i is not zero in the next period. Hence, the ETF liquidity shock may propagate to the underlying asset markets via ETF arbitrage trading.

3.2.2 Informational Trading Effects

After receiving or observing the news of a fundamental or a random shock, traders of the underlying assets start to trade their assets in the direction of the shock, but not simultaneously. This trading strategy by the underlying asset traders may cause the underlying assets to affect each other. For instance, trading the underlying assets in the direction of a random shock from the ETF market means that more liquidity is injected into the underlying market, which may cause $a_t^i - p_t^i$ to increase for some assets, if not all. Then the increase in $a_t^i - p_t^i$ may affect $p_t^i - b_t^i$ the next period if ϕ_{ba}^i, ϕ_{ba}^{ij}, and ϕ_{ba}^{ji} are different from zero, for $i = 1, 2, \ldots, n$ and $j = 2, \ldots, n$. The above

---

some random shocks which are neither transitory nor fundamental shocks that may come to the underlying asset i and ETF markets, respectively.

process is the informational trading effect (second channel) on the underlying markets due to a random shock from the ETF. Conversely, suppose traders start trading one of the underlying assets frequently, say asset 1, then there is a positive liquidity shock to $a_t^1 - p_t^1$ that increases it. Traders of the other underlying assets $j = 2, 3, \ldots, n$ will start trading in the same direction as the initial shock to asset 1, leading to asset 1 affecting other underlying assets and other underlying assets affecting each other. Thus, a shock to $a_t^1 - p_t^1$ may affect $p_t^j - b_t^j$ the next period if ϕ_{ba}^1 and ϕ_{ba}^{j1}'s are different from zero. Similarly, other assets may affect each other, i.e., a shock to $a_t^j - p_t^j$ may affect $p_t^j - b_t^j$ the next period if ϕ_{ba}^{2j}, ϕ_{ba}^{3j}, ..., ϕ_{ba}^{nj} are not zero, for $j = 2, 3, \ldots, n$. This informational trading effect results from a shock to one underlying asset.

3.2.3 Propagation of ETF Liquidity Shocks to Underlying Assets

In this subsection, I show how ETF liquidity shocks propagate to the underlying assets of an ETF in the dynamic model. I do this through the premium/discount of the ETF and variance of the returns of the underlying assets. In the dynamic model, the definition of the ETF arbitrage transaction component coincides with the definition of ETF premium given by

$$\begin{aligned}
\pi_t &= p_t^f - v_t \\
\Delta\pi_t &= \sum_{i=1}^{n} \eta_i \Delta(p_{t-1}^i - m_{t-1}^i) + \nu_f \Delta(p_{t-1}^f - v_{t-1}) + \lambda_{f,a}\Delta(a_{t-1}^f - p_{t-1}^f) + \\
&\quad \lambda_{f,b}\Delta(p_{t-1}^f - b_{t-1}^f) + \Delta\epsilon_{p^f,t} \qquad (10)
\end{aligned}$$

Returns of an ETF NAV and its variance are as follows:

$$\begin{aligned}
v_t - v_{t-1} &= \sum_{i=1}^{n} w_i (m_t^i - m_{t-1}^i) \\
var(v_t - v_{t-1}) &= \sum_{i=1}^{n} w_i^2 \sigma_{m^i,t}^2 \qquad (11)
\end{aligned}$$

Equation (11) is obtained since $(v_t - v_{t-1})$ and $(m_t^i - m_{t-1}^i)$ are stationary processes. The variance of the ETF NAV returns is the weighted sum of the fundamental variances of the underlying assets. The ETF returns can be shown as follows (see Appendix B2 for the derivation):

$$\begin{aligned}
p_t^f - p_{t-1}^f &= \Delta v_t + \sum_{i=1}^{n} \eta_i \Delta(p_{t-1}^i - m_{t-1}^i) + \nu_f \Delta p_{t-1}^f + \nu_f \Delta v_{t-1} + \\
&\quad \lambda_{f,a}\Delta(a_{t-1}^f - p_{t-1}^f) + \lambda_{f,b}\Delta(p_{t-1}^f - b_{t-1}^f) + \epsilon_{p^f,t} \qquad (12)
\end{aligned}$$

Variance of the ETF returns is given by

$$var(\Delta p_t^f) = D_1 \left[\sum_{i=1}^{n} (1+\nu_f^2+\eta_i^2\delta_i^2) w_i^2 \sigma^2_{\Delta m^i,t} + \lambda^2_{f,a}\sigma^2_{\Delta a^f,t} + \lambda^2_{f,b}\sigma^2_{\Delta b^f,t} + \sigma^2_{\Delta p^f,t} \right.$$
$$\left. + \sum_{i=1}^{n} (1+\nu_{i,p}^2)\eta_i^2 \sigma^2_{\Delta p^i,t} + \lambda^2_{i,a} \sum_{i=1}^{n} \eta_i^2 \sigma^2_{\Delta a^i,t} + \lambda^2_{i,a} \sum_{i=1}^{n} \eta_i^2 \sigma^2_{\Delta b^i,t} \right] \tag{13}$$

where $D_1 = \frac{1}{(1-\nu_f^2-\delta_i^2\sum_{i=1}^n \eta_i^2)}$. Similarly, the variance of the returns for underlying asset i is given by

$$var(\Delta p_t^i) = D_2 \left[\left(1+\nu_{i,p}^2 + \frac{\delta_i^2}{1-\nu_f^2} \sum_{i=1}^{n} \eta_i^2 \right) \sigma^2_{\Delta m^i,t} + \lambda^2_{i,a}\sigma^2_{\Delta a^i,t} + \lambda^2_{i,b}\sigma^2_{\Delta b^i,t} + \right.$$
$$\left. \sigma^2_{\Delta p^i,t} + \frac{1+\nu_f^2}{1-\nu_f^2}\sigma^2_{\Delta p^f,t} + \frac{\lambda^2_{f,a}\delta_i^2}{1-\nu_f^2}\sigma^2_{\Delta a^f,t} + \frac{\lambda^2_{f,b}\delta_i^2}{1-\nu_f^2}\sigma^2_{\Delta b^f,t} \right]. \tag{14}$$

where $D_1 = \frac{1}{(1-\nu_{i,p}^2-\frac{\delta_i^2}{1-\nu_f^2}\sum_{i=1}^n \eta_i^2)}$.

In equation (10), the deviations between ETF price and its NAV are contemporaneously correlated with the ETF returns, and its NAV returns volatility. Similarly, the deviations between the ETF price and its NAV are contemporaneously correlated with the underlying assets returns and their respective NAV returns. This can be seen from equation (14) where a fraction of the volatility of ETF liquidity, i.e., $\sigma^2_{\Delta a^f,t}$ and $\sigma^2_{\Delta b^f,t}$ finds its way into the underlying assets returns. This means that from the dynamic state-space model, the ETF arbitrage mechanism is a channel through which liquidity shocks may propagate from the ETF market to the underlying assets.

[8]I also find that an ETF may exhibit staleness in its NAV. The first term of the equation (13) is the variance of the ETF total returns, and the second and third terms are the variances of the ETF liquidity shocks to the ETF price. In addition, the fourth term is the variance from the random shock to the ETF price, and the last three terms represent the variances of the underlying asset's liquidity shocks to the ETF price. Comparing equations (11) and (13), the variance of the ETF price will always be greater than the variance of the ETF NAV. This means there is staleness in the ETF NAV, and the higher the degree of staleness in the NAV, the larger the variance of liquidity shocks. This confirms the argument by Madhavan and Sobczyk (2016), Evans et al. (2018), and Israeli et al. (2017), that the larger the variance of demand or supply shock to a stock, the higher the degree of stalenes in the NAV of the stock.

3.3 Measures for Analyzing Results

3.3.1 Measure for Excess Volatility

To determine whether ETF demand and supply (liquidity) shocks affect the underlying asset prices, I measure the intensity of shocks through the variances of returns of the underlying assets. The last two terms in equation (14) are the excess volatility brought

[8] Variance of the ETF fundamental returns is the weighted-sum of all the variances in the fundamental returns of the underlying assets.

to the price of underlying asset i when there is a liquidity shock to an ETF market. This excess volatility is permanent and transitory (non-fundamental). Similarly, the last two terms in equation (13) are the permanent transitory volatility brought to the ETF price if there is a liquidity shock to underlying asset i.

3.3.2 Measure for Speed of ETF Arbitrage

The prolonged periods of ETF pricing error are measured by the parameter ν_f. The parameter ν_f measures the speed at which AP acts to correct the ETF pricing errors, i.e., the deviations between the ETF price and its NAV. Since the ETF arbitrage component is a stationary process, the expected h seconds ahead of the premium from period t is given by $E[\Delta\pi_{t+h}] = \nu_f^h \Delta\pi_t$. Assuming half-life for the pricing error, the time horizon h needed to halve the deviation between the ETF price and its NAV is obtained through $0.5 = \nu_f^h$. Therefore, the time horizon an AP acts to keep the ETF price in line with its NAV is given by

$$h = \frac{\log(0.5)}{\log(|\nu_f|)} \tag{15}$$

The intuition is that the closer the value of the arbitrage parameter ν_f to 1, the slower the correction of the ETF pricing error by an AP. In the case of $\nu_f = 0$, equation (15) implies automatic or immediate correction of the ETF pricing error.

3.3.3 Measure for Price Discovery

The relevance of price discovery is one of the critical aspects of trading ETFs. ETF investors are concerned about buying at significant premiums or selling at enormous discounts. I examine whether ETF arbitrage transactions contribute to price discovery in the ETF market using equation (6), i.e., the ETF premium at any time t. The price discovery component (PDC) of the premium is the portion of the total variance of the premium that is not attributable to transitory noise shocks (Madhavan and Sobczyk, 2016). From equation (6), the measure for price discovery for the ETF arbitrage transaction is given by;

$$PDC = 1 - \frac{\sum_{i=1}^{n} \eta_i^n \sigma_{p^i,t}^2}{(1-\nu_f^2)\sigma_{\pi,t}^2} \tag{16}$$

where σ_π is the standard deviation of the premium $\pi = p_t^f - v_t$, and
$0 \leq PCD \leq 1$. Intuitively, the higher the price discovery component (PDC), the more contribution to price discovery by the ETF arbitrage transaction. Thus, the closer the arbitrage parameter ν_f to 1, the more the ETF arbitrage transaction contributes to price formation in the ETF market.

4. Data

I use tick data provided by TAQ for two ETFs: SPDR Dow Jones Industrial Average ETF (DIA) and VanEck Vectors Semiconductor ETF (SMH) listed on the NYSE. DIA and SMH have 30 and 25 underlying assets, respectively. DIA is an exchange-traded fund incorporated in the U.S. that tracks the Dow Jones Industrial Average Index. The ETF holds the 30 large-cap U.S. stocks that represent the Index. This ETF is registered as a Unit Investment Trust and pays a monthly dividend representing dividends paid by the underlying stocks. The ETF's holdings are price-weighted[9]. Similarly, SMH is an exchange-traded fund incorporated in the U.S. that tracks the performance results of the MVIS US-listed Semi-conductor 25 Index. The fund invests in the largest and most liquid companies listed in the U.S. that are active in the semiconductor sector[10].

The datasets span from January to March 2018 and are recorded on second-by-second frequency. The data contain the National Best Bid and Offer (NBBO) from NSYE for an ETF and each of the underlying assets of the ETF. The NBBO measures the best prices prevailing across all markets to focus on market price discovery.

DIA and SMH ETFs were selected based on the asset under management (AUM), market capitalization, and data availability. I implement the model for an ETF with higher AUM and market cap and an ETF with less AUM and market cap. DIA and SMH have AUM of $252.87B and $0.907B, respectively, and the average market cap for DIA is $ 21.32B, and SMH is $1.07B during the first quarter of 2018. Tables (5) and (6) in the Appendix report the descriptive statistics for DIA and SMH ETFs and their underlying assets, respectively. Comparing both ETFs, mid-quotes prices are higher in DIA than in SMH. I also report market capitalization for each of the underlying assets for both ETFs. Typically, assets with large market capitalization have higher prices.

I provide statistics on time-weighted bid-ask spreads in dollars and as a percentage of the prevailing quoted mid-points using the TAQ NBBO data sampled at seconds frequency. Spreads increase in dollar units and percentage terms from large to small assets. Spreads are likely to play an essential role in demand or supply liquidity decisions. However, spreads calculated based on displayed liquidity may overestimate the effective spreads received due to non-displayed orders. Furthermore, I compute the average mid-points for both ETFs and their underlying assets on second-by-second frequency over the three months. I present the average daily trading volume over the three months of the datasets. The average per day trading volume is higher for DIA ($251.18 million) compared to the trading volume for SMH at $1.26 million over the three months considered.

Similarly, I report the average daily trading volumes over the three months for each underlying asset for both DIA and SMH ETFs. For the underlying assets of DIA, Microsoft Corporation (MSFT) had the highest average trading volume at $52.83 million, and Travelers Companies Inc (TRV) recorded the lowest trading volume at $1.51 million. On the other hand, for the underlying assets of SMH, Advanced Micro Devices (AMD) recorded the highest average trading volume at $64.48 million, and ASML Holding NV (ASML) had the lowest trading volume at $1.35 million.

5. Empirical Results

The main empirical results focus on the following variables of interest; ETF arbitrage trading effect, informational trading effect, speed of ETF arbitrage, and price discovery in the ETF market. Tables (7) and (11) in Appendix C2 report selected estimated parameters of interest from the dynamic state-space model for DIA and SMH ETFs,

[9] https://www.bloomberg.com/quote/DIA:US
[10] https://www.bloomberg.com/quote/SMH:US

respectively[11]. Thus, column 4 in both tables (7) and (11) present the estimated parameters that capture the effect of ETF arbitrage on the respective underlying assets[12]. Column 5 reports the estimated parameters that measure the speed of ETF arbitrage and the underlying arbitrage transactions. Finally, columns 10 to 13 in tables (7) and (11) report the estimated parameters for cross-market transitory effects, i.e., ETF transitory effect on the underlying asset. Table (8) estimates the informational trading effect of the underlying assets on DIA ETF. Likewise, tables (9) and (10) report the estimates of the informational trading effect of two randomly selected underlying assets of DIA ETF (BA and AAPL) on all the underlying assets. Due to space limitations, see the Online Appendix for the estimates of the informational trading effect for the rest of the underlying assets of DIA ETF, SMH ETF, and all the underlying assets of SMH ETF. The estimated parameters are statistically significant at the 5% level for all or almost all ETFs and their underlying assets.

5.1 ETF Arbitrage Trading Effects

I find that the estimates for $\boldsymbol{\delta}_i$ measuring the direct effect of ETF arbitrage trading on the price of underlying asset $\boldsymbol{i}$ for both ETFs were not zero in the next period. As shown in equation (14) from section 3, the dynamic model in this study demonstrates how the ETF arbitrage mechanism serves as a channel through which shocks from the ETF market propagate to the prices of the underlying assets and vice versa. Column 4 of tables (7) and (11) in Appendix C2 provide evidence that any 'liquidity' shock to DIA and SMH ETFs propagates to their respective underlying assets via the ETF arbitrage mechanism since the estimates of $\boldsymbol{\delta}$ for the underlying assets of both ETFs are not zero. Hence, there are ETF arbitrage trading effects that are introduced directly into the underlying asset markets via the arbitrage trading component[13]. This confirms the argument by Ben-David et al. (2018) that any liquidity shocks to an ETF propagate to the underlying assets through the ETF arbitrage mechanism, and these shocks affect the underlying assets in the short term. In addition, I find that both DIA and SMH ETFs have a transitory effect on their respective underlying assets, i.e., $\boldsymbol{\theta}$ parameters for the underlying assets of DIA and SMH were not zero in the next period. The estimates in columns 10 to 13 of tables (7) and (11) are significantly not zero. This implies cross-market transitory effects on the underlying assets due to ETF arbitrage trading. Thus, liquidity shock to an ETF market affects the transitory components of the underlying assets through the ETF arbitrage mechanism. Hence, this confirms that an ETF arbitrage transaction is one of the channels responsible for the propagation of shocks from the ETF markets to the underlying asset markets.

Finally, I find that liquidity shock from an ETF market brings extra transitory volatility to each of the underlying asset prices of both ETFs. These new layers of volatility stay permanently in the prices of the underlying assets. However, the underlying assets do not receive an equal layer of the additional layer of transitory volatility. Thus, assets with higher weights in the portfolio of the underlying assets get a more significant portion

11 For the full estimated parameters of the dynamic state-space model for both ETFs and their corresponding underlying assets, see the Online Appendix (https://github.com/BoaduSebbe/onlineAppendix.git).

12 Similarly, column 3 presents the estimated parameters that capture the effect of underlying asset arbitrage on the respective ETFs.

13 Similarly, since $\sum_{i=1}^{n} \boldsymbol{\eta}_i$ is not zero in the next period, there is an underlying arbitrage trading effect that is introduced in the respective ETF markets, see Online Appendix.

of the excess transitory volatility from the ETF liquidity shocks. The first columns of tables (2) and (13) provide the estimated excess transitory volatility brought to the underlying assets when there is a liquidity shock to DIA and SMH ETFs and vice versa. For underlying assets of DIA ETF, BA is highly affected with extra transitory volatility of 0.925, and VZ is less affected with extra transitory volatility of 0.101 when one unit of liquidity shock propagates to the underlying assets from DIA. Similarly, for the underlying assets of SMH ETF, TSM is highly affected with extra transitory volatility of 0.954, and NVDA is less affected with extra transitory volatility of 0.237 when one unit of SMH liquidity shock propagates to the underlying assets.

One can also observe the above finding in the estimated mean impulse responses for the prices of the underlying assets of DIA and SMH ETFs in tables (1) and (12). Columns 2 to 4 of tables (1) and (12) report the estimates of IRFs of the underlying asset prices from the dynamic state-space model. The estimates in tables (1) and (12) can be interpreted as the average price impact of liquidity and fundamental shocks emanating from the respective ETFs to their underlying assets. For the underlying assets of both ETFs, a unit of ETF liquidity shock adds permanent non-fundamental volatility to the prices of these assets. This also confirms the argument by Ben-David et al. (2018) that a shock that propagates from the ETF markets to the underlying asset markets adds extra transitory volatility to the underlying assets' prices. I further provide graphs of the impulse responses, i.e., figures (1) and (2), for four randomly selected [14] underlying assets of both ETFs to gain a better insight into the effect of ETF liquidity shocks. One can see that ETF liquidity shocks to the underlying asset prices do not die out entirely, resulting from the excess volatility brought to the underlying markets through the ETF arbitrage mechanism.

5.2 Informational Trading Effect

Through the activities of investors of the underlying assets, I find informational trading effects on the underlying assets resulting from random or fundamental shocks. Tables (8) to (10) in the Appendix and table (1) to (28) on the Online Appendix report estimated informational trading effect parameters for the underlying assets of DIA. Similarly, tables (29) to (54) in the Online Appendix provide estimates for the informational trading effect on the underlying assets of SMH ETF. For both ETFs, the estimates for the parameters ϕ^{ij}_{aa}, ϕ^{ij}_{ab}, ϕ^{ij}_{ba}, and ϕ^{ij}_{bb}, $i, j = 1, 2, \dots n$ are not zero in the next period. These non-zero estimates of the informational trading effect show that the investors of the underlying assets, upon receiving or observing the news of a fundamental or random shock, start to trade their assets in the same direction of the initial shock, but not simultaneously, thus causing the underlying assets to affect each other. In effect, this means that fundamental or random shocks (which may include demand and supply shocks to the ETF) propagate to the underlying assets through the adjustment in the ask and bid quotes of the underlying assets of the ETF. This finding partially confirms that of Box et al. (2019), but they conclude that any shock to an ETF market has no arbitrage pressures on the underlying asset; rather, the shock propagates through the adjustment in the ask and bid quotes of the assets.

In contrast to Box et al. (2019), I find that ETF liquidity shocks propagate to the underlying assets of an ETF via the ETF arbitrage mechanism and also through the

[14] For all graphs of impulse responses for prices of the rest of the underlying assets, see the Online Appendix.

adjustment in the ask and bid quotes of the assets (transitory components of the assets). This suggests that there are ETF arbitrage pressures on the underlying assets of both ETFs. Although the channel that propagates a more significant portion of shocks to the underlying assets is not the focus of this paper, I believe the ETF arbitrage mechanism propagates a larger portion of the shock to the underlying assets. This is because, on average, the estimated parameters for the ETF arbitrage trading effect (δs) are higher in magnitude than those of the informational trading effect (ϕs).

Table 1: **DIA's Underlying Asset Price Impulse Response Function**

Price of Stock	*Ask Quote of DIA*	*Bid Quote of DIA*	*Price of DIA*
AAPL	0.045^{**}	-0.229^{**}	-0.026^{**}
CVX	0.025^{*}	-0.310^{*}	0.044^{*}
AXP	0.066^{**}	-0.024^{**}	0.017^{**}
BA	0.052^{**}	-0.049^{**}	-0.029^{**}
CSCO	0.023^{**}	-0.105^{**}	-0.022^{**}
CAT	-0.027^{**}	0.024^{**}	0.021^{**}
DIS	0.017^{**}	-0.017^{**}	0.098^{**}
IBM	0.010^{*}	-0.019^{*}	0.039^{*}
DOW	-0.020^{**}	0.015^{**}	-0.014^{**}
GS	0.029^{**}	-0.028^{**}	0.085^{**}
KO	0.010^{**}	-0.026^{**}	-0.052^{**}
PFE	0.020^{**}	-0.016^{**}	0.039^{**}
MFST	0.019^{**}	-0.026^{**}	0.029^{**}
NKE	0.023^{**}	-0.022^{**}	-0.030^{**}
PG	0.039^{**}	-0.018^{**}	0.026^{**}
HD	0.024^{**}	-0.020^{**}	0.012^{**}
JPM	0.022^{**}	-0.034^{**}	0.033^{**}
INTC	0.011^{**}	-0.017^{**}	0.020^{**}
JNJ	0.049^{**}	0.012^{**}	0.010^{**}
MMM	0.014^{**}	-0.014^{**}	0.015^{**}
MCD	0.030^{**}	-0.014^{**}	0.011^{**}
MRK	0.023^{**}	-0.021^{**}	0.023^{**}
WBA	0.019^{**}	-0.025^{**}	0.039^{**}
UTX	0.075^{**}	-0.032^{**}	0.017^{**}
TRV	0.035^{**}	-0.022^{**}	0.026^{**}
XOM	0.015^{**}	-0.028^{**}	0.031^{**}
WMT	0.018^{**}	-0.026^{**}	0.036^{**}
UNH	0.023^{*}	-0.022^{*}	0.031^{*}
V	0.021^{**}	-0.025^{**}	-0.050^{**}
VZ	-0.007^{**}	0.010^{**}	0.017^{**}

The table reports underlying asset day-ahead average impulse response functions (IRFs) from the dynamic state-space model developed in section 3 for the underlying assets of DIA ETF. The sample period is from January 2 to March 30, 2018, and the sampling frequency is 1-second. Columns 2 to 4 represent the variables being shocked by a unit, and rows correspond to the variables (underlying assets) affected by the unit shock from columns 2 to 4. The underlying assets' prices at one-second frequency are the mid-points of the best NBBs and best NBOs for the respective underlying assets. Observations include all the best NBBs and NBOs between 9:30 and 16:30 EST. The IRFs are estimated for one trading day ahead (representing 9 hours or equivalently 32,400 seconds). For the estimates of the IRFs, a *(**) next to the estimates indicates that the estimated IRF differs from zero and is statistically significant at the 1% (5%) level using clustered standard errors by an asset and by day.

5.3 Speed of ETF Arbitrage

Using equation (**??**), the corresponding estimated time in seconds for an AP to correct parity between the ETF prices and their NAVs are 231 (approximately 4 minutes) for DIA and 577 (approximately 10 minutes) for SMH. These time estimates are obtained through the parameter that measures the speed of ETF arbitrage (ν_f). The intuition is that the closer the value of ν_f to 1, the slower it takes an AP to act to correct the pricing error between an ETF price and its NAV. In the case of ν_f, being zero implies an immediate correction of the pricing error. As reported in column 5 of tables (7) and (11), $\hat{\nu_f}$ is 0.99699 for DIA and 0.99883 for SMH. In general, APs take a longer time to correct parities between the two ETF prices and their respective NAVs since $\hat{\nu_f}$ for both ETFs are closer to 1. Nonetheless, APs are likely to take less time to correct pricing errors for DIA than SMH assuming half-life for the unobserved parity between the ETF prices and their NAVs.

5.4 Contribution to Price Discovery

Using equation (16), the estimated PDC for DIA and SMH ETFs are 0.737 and 0.669, respectively. Intuitively, these estimates make sense since approximately 74%, and 67% of the variations in the premiums of DIA and SMH are due to price discovery, respectively. This result suggests that ETF premiums or discounts reflect staleness in ETF NAVs and transitory liquidity shocks in both ETF markets. This means that the ETF arbitrage transaction speeds up the price discovery process. Since transitory liquidity shocks contribute to the price discovery processes of both DIA and SMH ETFs, it implies that the volatility of the prices of DIA and SMH vary with the state of their respective bid/ask quotes.

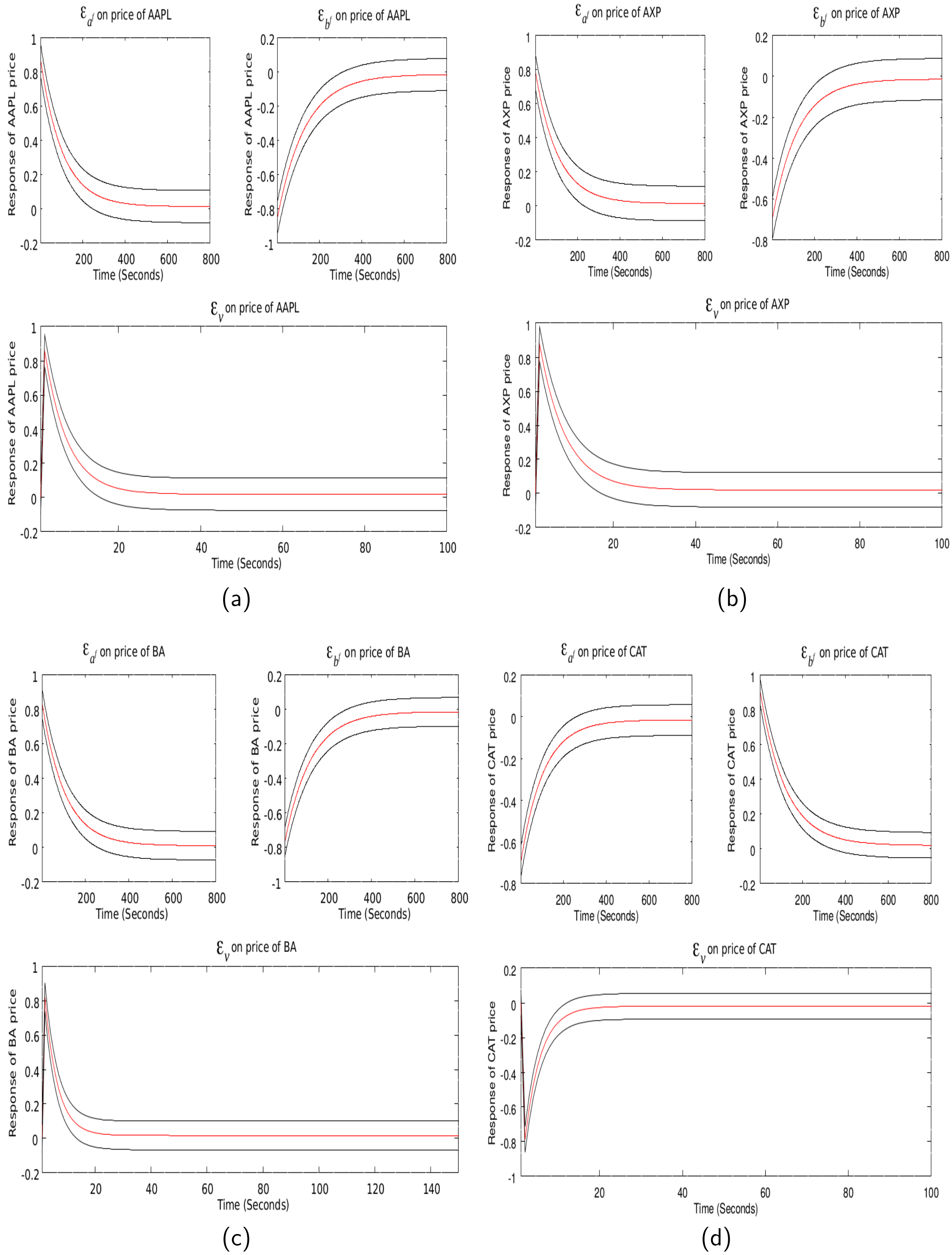
Finally, for the estimates of correlation coefficients in the covariance matrix Σ_t, see tables in the Online Appendix D. I find estimates that are consistent with the empirical findings of Pascual and Veredas (2010) and Hansen and Lunde (2006b). The negative correlations between ϵ_j and ϵ_k, $j \in \{m^i\}$, $k \in \{a^i, b^i, a^f, b^f\}$ imply that if there is an arrival of new information, then the best ask and bid quotes in both ETFs and their underlying assets adjust asymmetrically. The transitory component of the ask quotes increase (decrease), and the transitory component of the bid quotes decrease (increase) when there is a negative (positive) shock to the efficient prices of the ETFs and their underlying assets. This is the sign of adjustment of the best ask and bid quotes in the same direction as the efficient values. In addition, the arrival of a negative (positive) shock leading to the adjustment in the quoted prices is explained by traders that do not continuously monitor the market (Pascual and Veredas, 2009, 2010). Due to limited space, the estimated auxiliary parameters $\hat{a}_{j,k}$ and $\hat{b}_{j,k}$ of the GARCH (1,1) processes of the elements in the correlation matrix R_t are not reported, but are available upon request from the author.

Table 2: **Estimated Extra Volatility and Speed of Arbitrage: DIA**

	Extra Volatility	*Speed of Arbitrage (h)*
DIA		**230.7**
AAPL	**0.579** (0.199)	**58.81**
CVX	**0.542** (0.106)	**53.24**
AXP	**0.547** (0.067)	**45.68**
BA	**0.925** (0.047)	**21.67**
CSCO	**0.688** (0.036)	**34.03**
CAT	**0.542** (0.043)	**49.04**
DIS	**0.583** (0.055)	**25.95**
IBM	**0.781** (0.065)	**27.76**
DOW	**0.620** (0.041)	**25.46**
GS	**0.110** (0.034)	**51.85**
KO	**0.289** (0.013)	**52.20**
PFE	**0.155** (0.020)	**37.96**
MSFT	**0.386** (0.023)	**52.04**
NKE	**0.412** (0.132)	**52.31**
PG	**0.127** (0.094)	**51.08**
HD	**0.496** (0.135)	**50.47**
JPM	**0.267** (0.032)	**38.97**
INTC	**0.155** (0.017)	**41.08**
JHJ	**0.139** (0.012)	**31.99**
MMM	**0.456** (0.070)	**50.66**
MCD	**0.317** (0.015)	**52.32**
MRK	**0.249** (0.094)	**51.83**
WBA	**0.297** (0.117)	**51.81**
UTX	**0.305** (0.068)	**51.03**
TRV	**0.744** (0.007)	**52.04**
XOM	**0.611** (0.147)	**50.88**
WMT	**0.601** (0.013)	**50.85**
UNH	**0.108** 0.035	**52.32**
V	**0.419** (0.013)	**34.62**
VZ	**0.101** (0.042)	**51.84**

The table provides estimates for excess transitory volatility brought to the underlying assets of DIA and the speed of arbitrage trading. The first column presents the extra volatility that is brought to the underlying asset when one unit of demand and supply (liquidity) shocks the ETF. Parenthesized values in the first column are the respective additional transitory volatility brought to the price of the DIA ETF when there is a unit of liquidity shock to the respective underlying assets. On the other hand, column two shows the estimated time in seconds for the speed of arbitrage trading, assuming a half-life for the pricing error.

Figure 1: **Price Impulse Responses for Selected Underlying Assets of DIA**



6. Conclusion

In this study, I quantitatively examine the effects of ETF arbitrage trading on the underlying holdings of an ETF. In particular, the study focuses on answering the following questions: Do ETF liquidity demand shocks propagate to the underlying assets, and if so, through which channel(s)? If ETF liquidity shocks propagate to the underlying assets, do these shocks add any extra volatility to the prices of the underlying assets? At what speed can deviations between an ETF price and its NAV be corrected by an authorized participant (AP)? Finally, does the ETF arbitrage mechanism speed up or slow down price discovery in the ETF market? I develop a dynamic state-space model combining the price dynamics of an ETF, its underlying assets, and an ETF arbitrage mechanism to answer the empirical questions.

I apply high-frequency datasets spanning from January to March 2018 to estimate the model individually for Dow Jones Industrial Average ETF (DIA) investing in the 30 stocks of the DOW Jones index and VanEck Vectors Semiconductor ETF (SMH) with a portfolio of 25 stocks of MVIS semi-conductor index. The contribution of the study is to provide quantitative evidence to investigate the argument put forward by Ben-David et al. (2018) that liquidity demand shocks to an ETF propagate to the underlying assets through the ETF arbitrage mechanism. My results show that ETF liquidity (demand and supply) shocks propagate to the underlying assets via the ETF arbitrage transaction mechanism. First, I find that investors of the underlying assets, upon receiving news of a shock to the ETF or any of the underlying stocks, start to trade their assets in the same direction of the initial shock, but not simultaneously. This leads to the underlying assets affecting each other and indirectly inheriting the liquidity shocks from the ETF market or the shock from another underlying asset. Hence, some shocks, including ETF liquidity shocks, propagate to the underlying assets via the adjustment in the ask and bid quotes of the underlying assets. This is the informational trading effect. Secondly, I find that the ETF liquidity shocks that propagate to the underlying assets add permanent non-fundamental volatility to all the prices of the underlying assets and vice versa. However, the shocks do not equally affect the underlying assets of the ETFs. Assets with higher weights are highly affected as compared to less-weighted assets.

Furthermore, the estimates of the speed of ETF arbitrage transactions suggest that APs correct deviations between the DIA ETF and its NAV using approximately 4 minutes. Similarly, APs correct deviations between the SMH ETF and its NAV using 10 minutes. This means that, on average, APs act slower to correct any parity between the ETFs and their NAVs. However, APs act relatively slower in keeping the SMH price in line with its NAVs than correcting parity between the DIA price and its NAV. Finally, I show that observed premiums/discounts can be decomposed into transitory liquidity and price discovery components. The estimates of the price discovery components for the two ETFs suggest that premiums/discounts reflect staleness in the ETFs' NAVs against the transitory liquidity pressures in the ETF markets. These findings show that an ETF arbitrage transaction speeds up the price discovery process in the ETF market. Hence, the ETF arbitrage contributes price discovery in the ETF market.

More attention needs to be paid to how underlying asset prices are formed when new information arrives in the ETF markets. Thus, searching for factors that contribute to price discovery among the underlying assets of an ETF will be of great interest for future research. In most cases, researchers investigate price discovery between only two related asset markets or exchanges. It would be interesting to investigate price discovery among many assets simultaneously to see if other factors might contribute to asset price formation among these assets.

References

Agapova, A., Shelton, A., and Volkov, N. (2018). Bond etfs and price volatility of underlying securities.

Agarwal, V., Hanouna, P., Moussawi, R., and Stahel, C. W. (2018). Do etfs increase the commonality in liquidity of underlying stocks?

Bandi, F. M. and Russell, J. R. (2008). Microstructure noise, realized variance, and optimal sampling. *The Review of Economic Studies*, 75(2):339–369.

Ben-David, I., Franzoni, F., and Moussawi, R. (2017). Exchange traded funds. *Annual Review of Financial Economics*, (9):165–185.

Ben-David, I., Franzoni, F., and Moussawi, R. (2018). Do etfs increase volatility? *The Journal of Finance*, 73(6):2471–2535.

Box, T., Davis, R., Evans, R. B., and Lynch, A. A. (2019). Intraday arbitrage between etfs and their underlying portfolios. *Social Science Research Network SSRN*, (3322400).

Broman, M. S. and Shum, P. (2018). Relative liquidity, fund flows and short-term demand: Evidence from exchange-traded funds. *Financial Review*, 53(1):87–115.

Charteris, A., Chau, F., Gavriilidis, K., and Kallinterakis, V. (2014). Premiums, discounts and feedback trading: evidence from emerging markets' etfs. *International Review of Financial Analysis*, 35:80–89.

Da, Z. and Shive, S. (2013). When the bellwether dances to noise: Evidence from exchange-traded funds. Technical report, Working Paper, University of Notre Dame.

Dannhauser, C. D. (2017). The impact of innovation: Evidence from corporate bond exchange-traded funds (etfs). *Journal of Financial Economics*, 125(3):537–560.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350.

Engle, R. F. and Sarkar, D. (2006). Premiums-discounts and exchange traded funds. *The Journal of Derivatives*, 13(4):27–45.

Evans, R. B., Moussawi, R., Pagano, M. S., and Sedunov, J. (2018). Etf short interest and failures-to-deliver: Naked short-selling or operational shorting?

Hansen, P. R. and Lunde, A. (2006a). Consistent ranking of volatility models. *Journal of Econometrics*, 131(1-2):97–121.

Hansen, P. R. and Lunde, A. (2006b). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics*, 24(2):127–161.

Hasbrouck, J. (1996). 22 modeling market microstructure time series. *Handbook of statistics*, 14:647–692.

Hasbrouck, J. (2003). 22 modeling market microstructure time series. *Handbook of statistics*, 21.

Hilliard, J. (2014). Premiums and discounts in etfs: An analysis of the arbitrage mechanism in domestic and international funds. *Global Finance Journal*, 25(2):90–107.

Israeli, D., Lee, C. M., and Sridharan, S. A. (2017). Is there a dark side to exchange traded funds? an information perspective. *Review of Accounting Studies*, 22(3):1048–1083.

Lin, A. and Chou, A. (2006). The tracking error and premium/discount of taiwan's first exchange traded fund. *Web Journal of Chinese Management Review*, 9(3):1–21.

Lou, D. (2012). A flow-based explanation for return predictability. *The Review of Financial Studies*, 25(12):3457–3489.

Madhavan, A. and Sobczyk, A. (2016). Price dynamics and liquidity of exchange-traded funds. *Journal of Investment Management*, 14(2):1–17.

Menkveld, A. J., Koopman, S. J., and Lucas, A. (2007). Modeling around-the-clock price discovery for cross-listed stocks using state space methods. *Journal of Business & Economic Statistics*, 25(2):213–225.

Nam, J. (2017). Market accessibility, corporate bond etfs, and liquidity. *Kelley School of Business Research Paper*, (18-1).

Newey, K. and McFadden, D. (1994). Large sample estimation and hypothesis. *Handbook of Econometrics, IV, Edited by RF Engle and DL McFadden*, pages 2112–2245.

O'Hara, M. (2003). Presidential address: Liquidity and price discovery. *The Journal of Finance*, 58(4):1335–1354.

Pascual, R. and Veredas, D. (2009). What pieces of limit order book information matter in explaining order choice by patient and impatient traders? *Quantitative Finance*, 9(5):527–545.

Pascual, R. and Veredas, D. (2010). Does the open limit order book matter in explaining informational volatility? *Journal of Financial Econometrics*, 8(1):57–87.

Perlin, M., Dufour, A., and Brooks, C. (2014). The determinants of a cross market arbitrage opportunity: theory and evidence for the european bond market. *Annals of Finance*, 10(3):457–480.

Pontiff, J. (1996). Costly arbitrage: Evidence from closed-end funds. *The Quarterly Journal of Economics*, 111(4):1135–1151.

Vayanos, D. and Woolley, P. (2013). An institutional theory of momentum and reversal. *The Review of Financial Studies*, 26(5):1087–1145.

Appendix

Online Appendix: https://github.com/BoaduSebbe/onlineAppendix.git

Appendix A1: Dynamic State-Space Model-Full Version

The dynamic state-space model that combines the prices dynamics of an ETF, the price dynamics of the underlying holdings of the ETF, and the ETF arbitrage mechanism is defined in the full version as follows:

$$
\begin{bmatrix} m_t^1 \\ p_t^1 - m_t^1 \\ a_t^1 - p_t^1 \\ p_t^1 - b_t^1 \\ \vdots \\ m_t^n \\ p_t^n - m_t^n \\ a_t^n - p_t^n \\ p_t^n - b_t^n \\ p_t^f - v_t \\ a_t^f - p_t^f \\ p_t^f - b_t^f \end{bmatrix}
=
\begin{bmatrix} \mu_{m_1} \\ 0 \\ \beta_1 \\ -\beta_1 \\ \vdots \\ \mu_{m_n} \\ 0 \\ \beta_n \\ -\beta_n \\ 0 \\ \beta_f \\ -\beta_f \end{bmatrix}
+
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \nu_{1,p} & \lambda_{1,a} & \lambda_{1,b} & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \delta_1 & 0 & 0 \\
0 & \nu_{1,a} & \phi_{aa}^1 & \phi_{ab}^1 & 0 & 0 & \phi_{aa}^{12} & \phi_{ab}^{12} & \dots & 0 & 0 & \phi_{aa}^{1n} & \phi_{ab}^{1n} & 0 & \theta_{aa^f}^1 & \theta_{ab^f}^1 \\
0 & \nu_{1,b} & \phi_{ba}^1 & \phi_{bb}^1 & 0 & 0 & \phi_{ba}^{12} & \phi_{bb}^{12} & \dots & 0 & 0 & \phi_{ba}^{1n} & \phi_{bb}^{1n} & 0 & \theta_{ba^f}^1 & \theta_{bb^f}^1 \\
 & & & & & & & & \vdots & & & & & & & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & \nu_{n,p} & \lambda_{n,a} & \lambda_{n,b} & \delta_n & 0 & 0 \\
0 & 0 & \phi_{aa}^{n1} & \phi_{ab}^{n1} & 0 & 0 & \phi_{aa}^{n2} & \phi_{ab}^{n2} & \dots & 0 & \nu_{n,a} & \phi_{aa}^n & \phi_{ab}^n & 0 & \theta_{aa^f}^n & \theta_{ab^f}^n \\
0 & 0 & \phi_{ba}^{n1} & \phi_{bb}^{n1} & 0 & 0 & \phi_{ba}^{n2} & \phi_{bb}^{n2} & \dots & 0 & \nu_{n,b} & \phi_{ba}^n & \phi_{bb}^n & 0 & \theta_{ba^f}^n & \theta_{bb^f}^n \\
0 & \eta_1 & 0 & 0 & 0 & \eta_2 & 0 & 0 & \dots & 0 & \eta_n & 0 & 0 & \nu_f & \lambda_{f,a} & \lambda_{f,b} \\
0 & 0 & \psi_{a^f a}^1 & \psi_{a^f b}^1 & 0 & 0 & \psi_{a^f a}^2 & \psi_{a^f b}^2 & \dots & 0 & 0 & \psi_{a^f a}^n & \psi_{a^f b}^n & \nu_{f,a} & \psi_{a^f} & \psi_{a^f b^f} \\
0 & 0 & \psi_{b^f a}^1 & \psi_{b^f b}^1 & 0 & 0 & \psi_{b^f a}^2 & \psi_{b^f b}^2 & \dots & 0 & 0 & \psi_{b^f a}^n & \psi_{b^f b}^n & \nu_{f,b} & \psi_{b^f a^f} & \psi_{b^f}
\end{bmatrix}
\begin{bmatrix} m_{t-1}^1 \\ p_{t-1}^1 - m_{t-1}^1 \\ a_{t-1}^1 - p_{t-1}^1 \\ p_{t-1}^1 - b_{t-1}^1 \\ \vdots \\ m_{t-1}^n \\ p_t^n - m_t^n \\ a_{t-1}^n - p_{t-1}^n \\ p_{t-1}^n - b_{t-1}^n \\ p_{t-1}^f - v_{t-1} \\ a_{t-1}^f - p_{t-1}^f \\ p_{t-1}^f - b_{t-1}^f \end{bmatrix}
+
\begin{bmatrix} \epsilon_{m^1,t} \\ \epsilon_{p^1,t} \\ \epsilon_{a^1,t} \\ \epsilon_{b^1,t} \\ \vdots \\ \epsilon_{m^n,t} \\ \epsilon_{p^n,t} \\ \epsilon_{a^n,t} \\ \epsilon_{b^n,t} \\ \epsilon_{p^f,t} \\ \epsilon_{a^f,t} \\ \epsilon_{b^f,t} \end{bmatrix}
$$

Appendix A2: The Covariance Matrix Σ_t

Based on the assumptions on the entries of the covariance matrix $\boldsymbol{\Sigma}_t$, the covariance matrix is given by

$$\boldsymbol{\Sigma}_t = \begin{bmatrix}
\boldsymbol{\sigma}^2_{m^1,t} & \boldsymbol{\sigma}_{m^1p^1,t} & \boldsymbol{\sigma}_{m^1a^1,t} & \boldsymbol{\sigma}_{m^1b^1,t} & \boldsymbol{\sigma}_{m^1m^2,t} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\sigma}_{m^1m^n,t} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\sigma}_{m^1p^f,t} & \boldsymbol{\sigma}_{m^1a^f,t} & \boldsymbol{\sigma}_{m^1b^f,t} \\
 & \boldsymbol{\sigma}^2_{p^1,t} & \boldsymbol{\sigma}_{p^1a^1,t} & \boldsymbol{\sigma}_{p^1b^1,t} & \boldsymbol{\sigma}_{p^1m^2,t} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\sigma}_{p^1m^n,t} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\sigma}_{p^1p^f,t} & \boldsymbol{\sigma}_{p^1a^f,t} & \boldsymbol{\sigma}_{p^1b^f,t} \\
 & & \boldsymbol{\sigma}^2_{a^1,t} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\sigma}_{a^1p^f,t} & \mathbf{0} & \mathbf{0} \\
 & & & \boldsymbol{\sigma}^2_{b^1,t} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\sigma}_{b^1p^f,t} & \mathbf{0} & \mathbf{0} \\
 & & & & \boldsymbol{\sigma}^2_{m^2,t} & \boldsymbol{\sigma}_{m^2p^2,t} & \boldsymbol{\sigma}_{m^2a^2,t} & \boldsymbol{\sigma}_{m^2b^2,t} & \cdots & \boldsymbol{\sigma}_{m^2m^n,t} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\sigma}_{m^2p^f,t} & \boldsymbol{\sigma}_{m^2a^f,t} & \boldsymbol{\sigma}_{m^2b^f,t} \\
 & & & & & \boldsymbol{\sigma}^2_{p^2,t} & \boldsymbol{\sigma}_{p^2a^2,t} & \boldsymbol{\sigma}_{p^2b^2,t} & \cdots & \boldsymbol{\sigma}_{p^2m^n,t} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\sigma}_{p^2p^f,t} & \boldsymbol{\sigma}_{p^2a^f,t} & \boldsymbol{\sigma}_{p^2b^f,t} \\
 & & & & & & \boldsymbol{\sigma}^2_{a^2,t} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\sigma}_{a^2p^f,t} & \mathbf{0} & \mathbf{0} \\
 & & & & & & & \boldsymbol{\sigma}^2_{b^2,t} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \sigma_{b^2p^f,t} & \mathbf{0} & \mathbf{0} \\
 & & & & & & & & \vdots & & \vdots & & & & \vdots & \\
 & & & & & & & & & \boldsymbol{\sigma}^2_{m^n,t} & \boldsymbol{\sigma}_{m^np^n,t} & \boldsymbol{\sigma}_{m^na^n,t} & \boldsymbol{\sigma}_{m^nb^n,t} & \boldsymbol{\sigma}_{m^np^f,t} & \boldsymbol{\sigma}_{m^na^f,t} & \boldsymbol{\sigma}_{m^nb^f,t} \\
 & & & & & & & & & & \boldsymbol{\sigma}^2_{p^n,t} & \boldsymbol{\sigma}_{p^na^n,t} & \boldsymbol{\sigma}_{p^nb^n,t} & \boldsymbol{\sigma}_{p^np^f,t} & \boldsymbol{\sigma}_{p^na^f,t} & \boldsymbol{\sigma}_{p^nb^f,t} \\
 & & & & & & & & & & & \boldsymbol{\sigma}^2_{a^n,t} & \mathbf{0} & \boldsymbol{\sigma}_{a^np^f,t} & \mathbf{0} & \mathbf{0} \\
 & & & & & & & & & & & & \boldsymbol{\sigma}^2_{b^n,t} & \boldsymbol{\sigma}_{b^np^f,t} & \mathbf{0} & \mathbf{0} \\
 & & & & & & & & & & & & & \boldsymbol{\sigma}^2_{p^f,t} & \boldsymbol{\sigma}_{p^fa^f,t} & \boldsymbol{\sigma}_{p^fb^f,t} \\
 & & & & & & & & & & & & & & \boldsymbol{\sigma}^2_{a^f,t} & \mathbf{0} \\
 & & & & & & & & & & & & & & & \boldsymbol{\sigma}^2_{b^f,t}
\end{bmatrix}$$

Microstructure models have shown that the risk of information asymmetries partially determines the size of the bid-ask spreads. Hence, the transitory components of the ask and bid quotes in the ETF and all the underlying assets markets may be correlated with the innovations to the long-run values (v_t, m_t^i) in the presence of adverse selection. I analyze the daily cross-correlations between SPDR Dow Jones Industrial Average ETF (DIA) and its 30 underlying assets and the daily cross-correlations among the 30 underlying assets of DIA. A similar daily cross-correlations analysis is performed on SMH ETF and its 25 underlying assets (see Online Appendix A1 and A2 for these cross-correlations). I find results consistent with the results documented in the realized volatility literature, i.e., the cross-correlations vary with time, hence, the reason for assuming time varying Σ_t, see e.g., Bandi and Russell (2008),Hansen and Lunde (2006a), and Hansen and Lunde (2006b).

I do not restrict the correlation between the innovations to the long-run efficient prices ($\epsilon_{m^i,t}$) and the innovations to the transitory components in both the ETF and the underlying assets markets ($\epsilon_{a^i,t}$, $\epsilon_{b^i,t}$) and ($\epsilon_{a^f,t}$, $\epsilon_{b^f,t}$). Nonetheless, I restrict the contemporaneous correlation between the innovations to the transitory components in the ETF and underlying asset markets. Assuming that there is demand or supply shock on one side of a market, then the shock is expected to cause a lagged rather than a contemporaneous effect on the other side of the same market as captured by the dynamic model. Next, I restrict the contemporaneous correlations between the innovations to the quoted prices of the underlying assets and the innovations to the prices of other underlying assets. Thus, there is no correlation between $\epsilon_{a^i,t}$, $\epsilon_{b^i,t}$ and $\epsilon_{p^j,t}$ for $j \neq i$. However, I allow for the correlation between $\epsilon_{a^i,t}$, $\epsilon_{b^i,t}$ and $\epsilon_{p^j,t}$ for $j = i$, and also allow for the correlation between $\epsilon_{a^f,t}$, $\epsilon_{b^f,t}$ and $\epsilon_{p^f,t}$. Furthermore, I restrict the contemporaneous cross-correlations between the innovations to the underlying asset i's quotes ($\epsilon_{a^i,t}$, $\epsilon_{b^i,t}$) and innovations to the ETF quotes ($\epsilon_{a^f,t}$, $\epsilon_{b^f,t}$). The covariance matrix is given by Σ_t.

With the covariance matrix Σ_t, the dynamic model specification becomes an n-variate volatility model. I follow Engle (2002) in modeling the covariance terms of Σ_t using Dynamic Conditional Correlation (DCC) processes. The DCC represents a good compromise in terms of the number of parameters, flexibility, and capturing the true features of the data. I only need to estimate a few extra parameters of the model conditional variances. The diagonal entries of Σ_t are made up of informational volatility ($\sigma^2_{m^i,t}$) and transitory volatility $\left(\sigma^2_{j,t}, \;\; j \in \{a^i, b^i, p^i, p^f, a^f, b^f\}\right)$.

Informational volatility $\sigma^2_{m^i,t}$ for an underlying asset i is allowed to have both deterministic and dynamic components, which are assumed to follow the EGARCH process below;

$$\sigma^2_{m^i,t} = \exp\left(\tau_0^i + \tau_1^i(\sigma^2_{m^i,t-1}) + \tau_2^i\left(\frac{\epsilon_{m^i,t-1}}{\sigma_{m^i,t-1}}\right) + \tau_3^i\left|\frac{\epsilon_{m^i,t-1}}{\sigma_{m^i,t-1}}\right|\right)$$

where $\frac{\epsilon_{m^i,t-1}}{\sigma_{m^i,t-1}}$ is a standardized shock to the underlying asset i. The processes for $\sigma^2_{m^i,t}$ include intercepts τ_0^i; persistence parameters τ_1^i; sign or leverage parameters τ_2^i capture the asymmetric effects of good and bad news of prior shocks in respective markets; and magnitude parameters τ_3^i measure the effects of the size of these same shocks in respective markets.

Furthermore, I assume that the conditional variances of the arbitrage and transitory components $\sigma^2_{j,t}$, $j \in \{a^i, b^i, p^i, p^f, a^f, b^f\}$ are functions of past shocks, leading to the following $EGARCH(0,1)$ model:

$$\sigma^2_{j,t} = \exp\left(\kappa^0_j + \kappa_j \left|\frac{\epsilon_{j,t-1}}{\sigma_{j,t-1}}\right|\right), \; j \in \{a^i, b^i, p^i, p^f, a^f, b^f;\; i = 1, 2, \ldots n\}$$

where κ_j measures the effect of prior shocks through the absolute value of the standardized shocks in respective markets, which are transitory in nature.

Appendix A3: Estimation Procedure

In this Appendix section, I describe the estimation procedure of the dynamic state-space model. Following Newey and McFadden (1994) and Engle (2002), the dynamic model parameters are estimated using a two-step approach using the Augmented Kalman filter technique. The dynamic state space model is represented by

$$X_t = A + BX_{t-1} + \epsilon_t, \qquad \bar{\epsilon}_t \sim N(0,\ \Sigma_t)$$

Since I assume that the variances of the vector of innovations, ϵ_t, follow dynamic conditional correlation (DCC), then

$$\Sigma_t = D_t R_t D_t^T z_t$$

where $D_t = diag\left\{\sqrt{\sigma^2_{j,t}}\right\}$ for $j \in \{m^i, p^i, a^i, b^i, p^f, a^f, b^f;\; i = 1, 2, ...n\}$, R_t is a conditional correlation matrix of the standard disturbances $\epsilon_{j,t}$, and z_t is a vector of *iid* errors such that $E[z_t] = 0$ and $E[z_t z_t^T] = I_{4n+3}$. The covariance matrix Σ_t should be symmetric and positive definite, and to ensure that both D_t and R_t must be symmetric and positive definite. By construction, D_t is symmetric and positive definite since D_t is a diagonal matrix and the diagonal entries are the square root of variances. R_t is structured to have its major diagonal entries 1 and off-diagonal elements as the respective correlation coefficients. The off-diagonal elements of R_t are given by

$$\rho_{j,k,t} = \frac{q_{j,k,t}}{\sqrt{q_{j,j,t}q_{k,k,t}}}, \qquad j,k \in \{m^i, p^i, a^i, b^i, p^f, a^f, b^f,\;\; i = 1, 2, \ldots n\}$$

where the auxiliary variables $q_{j,k,t}$ follow GARCH (1,1) process as follows:

$$q_{j,k,t} = \Gamma_{j,k}(1 - a_{j,k} - b_{j,k}) + a_{j,k}\left(\epsilon_{j,k,t}\epsilon^T_{j,k,t}\right) + b_{j,k}q_{j,k,t-1} \tag{17}$$

with $\Gamma_{j,k}$ being unconditional correlation between $\epsilon_{j,t}$ and $\epsilon_{k,t}$. To ensure that R_t is positive definite, parameters $a_{j,k}$ and $b_{j,k}$ must satisfy $a_{j,k},\ b_{j,k} > 0$ and $a_{j,k} + b_{j,k} \leq 1$. By construction, both D_t and R_t are symmetric and positive definite, hence Σ_t is symmetric and positive definite. Under the normality assumption, I estimate the model parameters and the correlation coefficients by augmented Kalman filter following a two-step estimation approach. I denote parameters in matrix D_t by Δ and additional parameters in the correlation matrix R_t by Q. The log-likelihood for the estimator is expressed as :

$$\begin{aligned}
\boldsymbol{y}_t|\boldsymbol{I}_{t-1} &\sim \boldsymbol{N}(\boldsymbol{0}_{4n+3}, \Sigma_t) \\
\boldsymbol{l} &= -\frac{1}{2}\sum_{t=1}^{T}\left(\boldsymbol{n}\log(2\boldsymbol{\pi}) + \log|\Sigma_t| + \boldsymbol{y}_t\Sigma_t^{-1}\boldsymbol{y}_t\right) \\
&= -\frac{1}{2}\sum_{t=1}^{T}\left(\boldsymbol{n}\log(2\boldsymbol{\pi}) + 2\log|\boldsymbol{D}_t| + \boldsymbol{y}_t^T\boldsymbol{D}_t^{-1}\boldsymbol{D}_t^{-1}\boldsymbol{y}_t - \boldsymbol{e}_t^T\boldsymbol{e}_t\right. \\
&\quad \left. + \log|\boldsymbol{R}_t| + \boldsymbol{e}_t^T\boldsymbol{R}_t\boldsymbol{e}_t\right)
\end{aligned}$$

Due to the large size of the covariance matrix Σ_t, I follow the two-step approach of Newey and McFadden (1994). I denote parameters in matrix $\boldsymbol{D}_t$ by $\boldsymbol{\Delta}$ and additional parameters in the correlation matrix $\boldsymbol{R}_t$ by $\boldsymbol{Q}$. The log-likelihood function is expressed as the sum of the volatility part and correlation part:

$$l(\boldsymbol{\Delta}, \boldsymbol{Q}) = l_v(\boldsymbol{\Delta}) + l_c(\boldsymbol{\Delta}, \boldsymbol{Q})$$

where $l_v(\boldsymbol{\Delta})$ is the log-likelihood for the volatility part and $l_c(\boldsymbol{\Delta}, \boldsymbol{Q})$ is the log-likelihood for the correlation part. The volatility part is given by

$$l_v(\boldsymbol{\Delta}) = -\frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{n}\log(2\boldsymbol{\pi}) + \log(|\boldsymbol{D}_t|^2) + \boldsymbol{y}_t^T\boldsymbol{D}_t^{-2}\boldsymbol{y}_t), \tag{18}$$

and correlation part expressed as

$$l_c(\boldsymbol{\Delta}, \boldsymbol{Q}) = -\frac{1}{2}\sum_{t=1}^{T}(\log(|\boldsymbol{R}_t|) + \boldsymbol{e}_t^T\boldsymbol{R}_t^{-1}\boldsymbol{e}_t - \boldsymbol{e}_t^T\boldsymbol{e}_t) \tag{19}$$

where $\boldsymbol{e}_t = \boldsymbol{D}_t^{-1}\boldsymbol{y}_t$, and $\boldsymbol{y}_t = [\boldsymbol{X}_t - (\boldsymbol{A} + \boldsymbol{B}\boldsymbol{X}_{t-1})]$.

The first stage in the two-step approach is to maximize the log-likelihood of the volatility term $\hat{\boldsymbol{\Delta}} = \boldsymbol{argmax}\,\{l_v(\boldsymbol{\Delta})\}$ and in the second stage the correlation parameters $\boldsymbol{Q}$ are estimated by taking $\hat{\boldsymbol{\Delta}}$ as given and then maximizing the log-likelihood of the correlation term:$\hat{\boldsymbol{Q}} = \boldsymbol{argmax}\left\{l_v(\hat{\boldsymbol{\Delta}}, \boldsymbol{Q})\right\}$. The consistency of the correlation parameters $\boldsymbol{Q}$ is ensured by the consistency of the volatility term parameters $\boldsymbol{\Delta}$. The maximum of the second stage is a function of the first stage parameter estimates. Hence, if the estimated model parameters are consistent, then the correlation parameters will be consistent as long as the log-likelihood of the correlation term is continuous in the neighborhood of the true parameters. The dynamic state-space model has $2(\boldsymbol{n}+1)(2\boldsymbol{n}+7)+3\boldsymbol{n}$ parameters excluding the auxiliary parameters of the GARCH (1,1) processes of the off-diagonal elements of $\boldsymbol{R}_t$ matrix and also exclude conditional correlation coefficients of $\boldsymbol{R}_t$.

I formulate the DCC process for the framework of this study as follows:

$$\begin{aligned}
&y_t|I_{t-1} \sim N(\mathbf{0}_n, D_t^T R_t D_t), \quad y_t = [X_t - (A + BX_{t-1})] \quad \textit{and} \\
&\qquad I_{t-1} \in \{X_{t-1}, D_{t-1}, R_{t-1}, y_{t-1}, ...\} \\
&D_t = diag\left\{\sqrt{\sigma^2_{j,t}}\right\}, j \in \{m^i, p^i, a^i, b^i, v, p^f, a^f, b^f; i = 1, 2, \ldots n\} \\
&e_t = D_t^{-1} y_t \\
&R_t = diag\{Q_t\}^{-1} Q_t diag\{Q_t\}^{-1}
\end{aligned}$$

where the entries of Q_t matrix are the auxiliary variables $q_{j,k,t}$ expressed in equation (17). Under the normality assumption of the first equation above, I obtain a log-likelihood function. The second matrix equation assumes that each asset follows a univariate EGARCH process.

Appendix B1: Derivation of the Variance of ETF NAV Returns (11)

From equation (5), the ETF NAV returns is given by

$$\begin{aligned}
v_t - v_{t-1} &= \sum_{i=1}^{n} w_i(m_t^i - m_{t-1}^i) \\
&= \sum_{i=1}^{n} w_i(\Delta m_{t-1}^i + \Delta\epsilon_{m^i,t})
\end{aligned}$$

The variance of ETF NAV returns is

$$var(v_t - v_{t-1}) = \sum_{i=1}^{n} w_i^2 \sigma^2_{\Delta m^i,t}$$

Appendix B2: Derivation of the Variance of ETF Returns (13)

Starting from equation (6), the price of an ETF is modeled as

$$\begin{aligned}
p_t^f &= v_t + \sum_{i=1}^{n} \eta_i(p_{t-1}^i - m_{t-1}^i) + \nu_f(p_{t-1}^f - v_{t-1}) + \lambda_{f,a}(a_{t-1}^f - p_{t-1}^f) + \\
&\lambda_{f,b}(p_{t-1}^f - b_{t-1}^f) + \epsilon_{p^f,t} \qquad (20)
\end{aligned}$$

Shifting forward by one period, I obtain the ETF return as

$$\begin{aligned}
\Delta p_t^f &= \Delta v_t + \sum_{i=1}^{n} \eta_i \Delta(p_{t-1}^i - m_{t-1}^i) + \nu_f \Delta p_{t-1}^f - \nu_f \Delta v_{t-1} + \\
&\lambda_{f,a}\Delta(a_{t-1}^f - p_{t-1}^f) + \Delta\lambda_{f,b}(p_{t-1}^f - b_{t-1}^f) + \Delta\epsilon_{p^f,t}
\end{aligned}$$

The variance of ETF returns is

$$var(\Delta p_t^f) = var(\Delta v_t) + \sum_{i=1}^{n} \eta_i var(\Delta(p_{t-1}^i - m_{t-1}^i)) + \nu_f var(\Delta(p_{t-1}^f - v_{t-1}))$$
$$+ \lambda_{f,a} var(\Delta(a_{t-1}^f - p_{t-1}^f)) + \Delta var(\lambda_{f,b}(p_{t-1}^f - b_{t-1}^f)) + var(\Delta\epsilon_{p^f,t})$$

$$var(\Delta p_t^f) = \sum_{i=1}^{n} w_i^2 \sigma_{\Delta m^i,t}^2 + \sum_{i=1}^{n} \eta_i^2 var(\Delta(p_{t-1}^i - m_{t-1}^i)) + \nu_f^2 var(\Delta p_{t-1}^f) +$$
$$\nu_f^2 + \sum_{i=1}^{n} w_i^2 \sigma_{\Delta m^i,t}^2 + \lambda_{f,a}^2 \sigma_{\Delta a^f,t}^2 + \lambda_{f,b}^2 \sigma_{\Delta b^f,t}^2 + \sigma_{\Delta p^f,t}^2 \qquad (21)$$

But;

$$\Delta(p_{t-1}^i - m_{t-1}^i) = \Delta(p_{t-2}^i - m_{t-2}^i) + \lambda_{i,a}(a_{t-2}^i - p_{t-2}^i) + \lambda_{i,b}(p_{t-2}^i - b_{t-2}^i) + \Delta\epsilon_{p^i,t}$$

Taking the variance of the last equation above and substituting it into equation (21), I obtain equation (13). A similar derivation can be done for the variance of the returns of underlying asset i using equation (14).

Appendix C1: Descriptive Statistics

Table 3: **Weights for the Underlying Assets of DIA**

	AAPL	CVX	AXP	BA	CSCO	CAT	DIS	IBM	DOW	GS
Weights	0.064	0.0297	0.0295	0.0896	0.011	0.035	0.0359	0.0327	0.0132	0.0535
	HD	KO	JPM	INTC	JNJ	MMM	MCD	NKE	MRK	MSFT
Weights	0.0577	0.0128	0.0314	0.0141	0.0319	0.0417	0.0473	0.0223	0.0206	0.0361
	WBA	UTX	PG	PFE	TRV	XOM	WMT	UNH	V	VZ
Weights	0.0151	0.0361	0.0294	0.0089	0.0327	0.0167	0.0294	0.0624	0.0438	0.0145

Source: etf.com
The table reports the average weights of the Underlying assets of DIA from January 2, 2018, to March 31, 2018.

Table 4: **Weights for the Underlying Assets of SMH**

	TSM	INTC	NVDA	AMD	ASML	TXN	QCOM	MU	AVGO
Weights	0.1268	0.1169	0.0624	0.0524	0.0483	0.048	0.048	0.0474	0.067
	NXPI	LRCX	AMAT	ADI	KLAC	XLNX	STM	MCHP	CDNS
Weights	0.0600	0.0477	0.0432	0.0417	0.0286	0.0274	0.0261	0.0232	0.0238
	SWKS	MXIM	TER	MRVL	QRVO	ON	OLED		
Weights	0.0200	0.0175	0.0125	0.0124	0.0112	0.0104	0.009		

Source: etf.com
The table reports the average weight of the Underlying assets of SMH from January 2, 2018, to March 31, 2018.

Table 5: **Descriptive Statistics for DIA and its Underlying Assets**

Symbol	Source	AUM (\$$B$)	$Market$ Cap (\$$B$)	$mid-quote$ (\$)	$Bid-Ask$ $spread$ (\$)	$Relative$ $Bid-Ask$ $spread$ (bps)	$Trading$ $Volume$ (\$$M$)	Observations
DIA	TAQ	252.87	21.32	258.71	1.300	0.005	251.18	1521000
AAPL	TAQ	367.8	860.03	174.45	0.289	0.0016	39.6	1521000
AXP	TAQ	179.96	70.63	99.33	0.447	0.0045	4.39	1521000
BA	TAQ	113.55	196.10	310.70	0.276	0.0014	5.00	1521000
CVX	TAQ	256.44	217.83	130.12	0.450	0.0035	6.50	1521000
CSCO	TAQ	131.51	208.38	41.47	0.733	0.0172	38.09	1521000
CAT	TAQ	78.99	88.12	72.18	0.904	0.0134	4.87	1521000
DIS	TAQ	97.94	155.70	111.37	1.03	0.0093	8.24	1521000
IBM	TAQ	125.29	140.84	167.60	0.084	0.0001	3.76	1521000
DOW	TAQ	130.01	31.21	27.25	0.454	0.0175	17.40	1521000
GS	TAQ	973.54	95.17	261.71	0.573	0.0022	3.12	1521000
HD	TAQ	44.53	206.53	196.84	0.489	0.0025	5.23	1521000
KO	TAQ	93.28	184.30	47.46	1.004	0.0213	12.40	1521000
JPM	TAQ	2609.79	374.42	112.95	0.943	0.0083	17.68	1521000
INTC	TAQ	128.6	210.99	46.99	1.065	0.0235	46.32	1521000
JNJ	TAQ	156.63	343.34	146.30	0.936	0.0063	7.23	1521000
MMM	TAQ	38.58	130.33	246.48	0.423	0.0017	2.85	1521000
MCD	TAQ	33.72	122.79	171.97	2.44	0.0141	4.04	1521000
NKE	TAQ	22.55	82.25	62.76	0.780	0.0127	9.55	1521000
MRK	TAQ	86.64	158.38	56.40	1.380	0.0244	12.14	1521000
MSFT	TAQ	245.5	241.68	85.97	0.830	0.0972	52.83	1521000
WBA	TAQ	70.82	68.32	73.19	2.810	0.0038	7.92	1521000
UTX	TAQ	55.43	100.66	126.36	2.840	0.0022	6.62	1521000
PG	TAQ	124.37	229.80	92.14	0.890	0.0097	11.77	1521000
PFE	TAQ	164.61	214.36	36.31	0.180	0.0049	24.88	1521000
TRV	TAQ	103.68	37.06	129.81	1.95	0.0142	1.51	1521000
XOM	TAQ	348.83	315.89	84.02	1.010	0.0125	15.97	1521000
WMT	TAQ	204.52	258.98	105.79	2.42	0.0229	9.14	1521000
UNH	TAQ	155.57	227.18	219.01	9.98	0.0456	3.99	1521000
V	TAQ	69.04	224.78	173.17	0.410	0.0237	10.02	1521000
VZ	TAQ	264.52	197.59	58.42	1.63	0.0271	16.25	1521000

Table (5) provides descriptive statistics for DIA ETF and its 30 underlying assets. I calculate the average of the following second-by-second statistics for the SMH ETF and each of its underlying assets. Bid–Ask spread is the average difference between the bid and ask quotes. The relative bid-ask spread is the average ratio of the quoted bid-ask spread divided by the quote midpoint, weighted by time; Volume is measured as the number of shares in dollars transacted. Market capitalization is measured at the end of March 2018.

Table 6: **Descriptive Statistics for SMH and its Underlying Assets**

Symbol	Source	AUM ($\$B$)	$Market$ Cap ($\$B$)	$mid{-}quote$ ($)	$Bid{-}Ask$ $spread$ ($)	$Relative$ $Bid{-}Ask$ $spread$ (bps)	$Trading$ $Volume$ ($\$M$)	Observations
SMH	TAQ	0.907	1.07	101.24	2.42	0.024	1.26	1521000
TSM	TAQ	70.08	707.75	43.30	3.413	0.0048	7.17	1521000
INTC	TAQ	128.6	210.99	46.99	1.065	0.0231	46.32	1521000
NVDA	TAQ	11.24	149.18	12.42	0.465	0.0371	24.26	1521000
AMD	TAQ	3.76	11.7	190.37	1.352	0.0071	64.48	1521000
ASML	TAQ	22.92	54.76	111.85	1.719	0.0155	1.35	1521000
TXN	TAQ	24.13	103.58	66.28	1.614	0.0247	5.43	1521000
QCOM	TAQ	64.13	74.92	17.79	2.360	0.1331	11.43	1521000
MU	TAQ	41.26	70.26	264.96	1.927	0.0734	65.62	1521000
AVGO	TAQ	54.54	110.06	118.98	1.923	0.0168	5.03	1521000
NXPI	TAQ	21.53	40.51	203.22	1.435	0.0071	5.73	1521000
LRCX	TAQ	10.77	31.23	62.123	1.309	0.0218	3.71	1521000
AMAT	TAQ	17.63	49.8	93.10	2.019	0.0229	15.32	1521000
ADI	TAQ	20.44	17.31	110.55	1.212	0.0114	3.60	1521000
KLAC	TAQ	5.61	16.77	70.97	1.541	0.0220	1.54	1521000
XLNX	TAQ	5.06	16.41	23.181	0.121	0.0052	2.63	1521000
STM	TAQ	10.07	11.15	22.86	1.765	0.0772	3.00	1521000
MCHP	TAQ	8.26	22.52	44.21	1.770	0.0401	3.11	1521000
CDNS	TAQ	2.49	10.37	96.86	2.078	0.0216	2.73	1521000
SWKS	TAQ	4.75	15.80	53.78	1.140	0.0217	1.95	1521000
MXIM	TAQ	3.68	17.40	45.87	1.503	0.0334	2.93	1521000
TER	TAQ	4.65	8.68	22.04	0.130	0.0059	1.90	1521000
MRVL	TAQ	9.98	10.77	67.33	0.119	0.0016	8.13	1521000
QRVO	TAQ	5.85	8.46	36.86	1.004	0.0272	1.73	1521000
ON	TAQ	1.01	9.59	21.65	1.182	0.0008	7.26	1521000
OLED	TAQ	1.06	7.58	175.74	1.087	0.0062	1.44	1521000

Table (6) provides descriptive statistics for SMH ETF and its 25 underlying assets. I calculate the average of the following second-by-second statistics for the SMH ETF and each of its underlying assets. Bid–Ask spread is the average difference between the bid and ask quotes. The relative bid-ask spread is the average ratio of the quoted bid-ask spread divided by the quote midpoint, weighted by time; Volume is measured as the number of shares in dollars transacted. Market capitalization is measured at the end of March 2018.

Appendix C2: Estimated Parameters for DIA ETF and its Underlying Assets.

Table 7: **Estimated dynamic state-space model for DIA and its 30 underlying assets**

ETF		β_f			ν^f	ν_a^f	ν_b^f	λ_1^f	λ_2^f	ψ_{a^f}	$\psi_{a^f b^f}$	$\psi_{b^f a^f}$	ψ_{b^f}
DIA		0.066 (0.07)			0.997 (0.09)	0.922 (0.08)	0.998 (0.05)	0.186 (0.07)	−0.037 (0.09)	0.870 (0.27)	0.080 (0.07)	0.088 (0.09)	0.702 (0.24)
STOCK	μ_m	β	η	δ	ν_p	ν_a	ν_b	λ_a	λ_b	θ_{aa^f}	θ_{ab^f}	θ_{ba^f}	θ_{bb^f}
AAPL	2.587 (0.17)	0.023 (0.11)	−0.162 (0.10)	0.409 (0.04)	0.859 (0.19)	0.877 (0.03)	0.778 (0.01)	0.149 (0.51)	−0.171 (0.31)	0.157 (0.17)	0.147 (0.09)	0.127 (0.05)	0.172 (0.04)
CVX	4.695 (0.08)	0.136 (0.09)	0.407 (0.08)	0.579 (0.06)	0.930 (0.01)	0.958 (0.39)	0.889 (0.06)	0.189 (0.06)	−0.111 (0.06)	0.077 (0.09)	0.079 (0.05)	0.211 (0.03)	0.220 (0.10)
AXP	3.593 (0.12)	0.137 (0.08)	−0.072 (0.18)	0.368 (0.07)	0.929 (0.07)	0.977 (0.13)	0.878 (0.09)	0.139 (0.32)	−0.100 (0.05)	0.085 (0.05)	0.079 (0.09)	0.079 (0.20)	0.082 (0.08)
BA	2.489 (0.23)	0.254 (0.29)	0.417 (0.05)	0.409 (0.08)	0.979 (0.05)	0.980 (0.06)	0.909 (0.04)	0.149 (0.10)	−0.171 (0.17)	0.077 (0.03)	0.077 (0.06)	0.077 (0.05)	0.077 (0.31)
CSCO	−1.594 (0.19)	0.042 (0.08)	0.337 (0.15)	0.479 (0.06)	0.979 (0.07)	0.977 (0.08)	0.975 (0.02)	0.194 (0.06)	−0.171 (0.48)	0.077 (0.09)	0.077 (0.10)	0.077 (0.32)	0.077 (0.21)
CAT	5.999 (0.18)	0.174 (0.05)	0.257 (0.08)	0.629 (0.04)	0.979 (0.07)	0.978 (0.05)	0.978 (0.02)	0.194 (0.08)	−0.171 (0.07)	0.144 (0.21)	0.130 (0.23)	0.088 (0.22)	0.152 (0.10)
DIS	5.519 (0.08)	0.086 (0.10)	0.367 (0.05)	0.379 (0.03)	0.964 (0.35)	0.978 (0.05)	0.978 (0.07)	0.143 (0.06)	−0.126 (0.05)	0.078 (0.18)	0.551 (0.09)	0.079 (0.08)	0.077 (0.06)
IBM	5.697 (0.11)	0.090 (0.10)	0.377 (0.19)	0.479 (0.04)	0.938 (0.05)	0.979 (0.07)	0.977 (0.02)	0.134 (0.11)	−0.127 (0.32)	0.077 (0.10)	0.076 (0.15)	0.130 (0.11)	0.148 (0.10)
DOW	2.739 (0.09)	0.113 (0.34)	0.297 (0.06)	0.424 (0.35)	0.993 (0.22)	0.977 (0.08)	0.977 (0.01)	0.129 (0.30)	−0.033 (0.03)	0.077 (0.20)	0.077 (0.08)	0.077 (0.01)	0.077 (0.06)
GS	8.639 (0.08)	0.042 (0.04)	0.527 (0.39)	0.479 (0.04)	0.930 (0.04)	0.977 (0.32)	0.978 (0.01)	0.134 (0.25)	−0.087 (0.21)	0.077 (0.09)	0.077 (0.05)	0.077 (0.03)	0.077 (0.09)
HD	−10.579 (0.08)	0.143 (0.03)	0.527 (0.23)	0.439 (0.06)	0.949 (0.07)	0.957 (0.08)	0.979 (0.01)	−0.050 (0.04)	0.129 (0.05)	0.078 (0.08)	0.077 (0.03)	0.081 (0.09)	0.078 (0.05)
KO	2.660 (0.10)	0.028 (0.03)	−0.172 (0.06)	0.389 (0.03)	0.955 (0.01)	0.977 (0.09)	0.977 (0.03)	0.129 (0.11)	−0.092 (0.04)	0.081 (0.13)	0.078 (0.10)	0.078 (0.17)	0.077 (0.10)
JPM	5.639 (0.09)	0.052 (0.09)	0.417 (0.19)	0.599 (0.03)	0.965 (0.02)	0.977 (0.08)	0.977 (0.12)	0.129 (0.13)	−0.181 (0.05)	0.387 (0.06)	0.317 (0.12)	0.077 (0.07)	0.077 (0.27)
INTC	2.630 (0.12)	0.022 (0.50)	−0.182 (0.19)	0.549 (0.03)	0.925 (0.06)	0.977 (0.045)	0.979 (0.034)	0.129 (0.057)	−0.008 (0.076)	0.078 (0.07)	0.077 (0.21)	0.149 (0.03)	0.176 (0.03)
JNJ	5.979 (0.09)	0.100 (0.05)	0.929 (0.06)	0.499 (0.05)	0.989 (0.02)	0.977 (0.12)	0.977 (0.33)	0.129 (0.09)	−0.421 (0.09)	0.387 (0.09)	0.274 (0.06)	0.076 (0.03)	0.264 (0.04)
MMM	6.759 (0.34)	0.674 (0.04)	0.297 (0.07)	0.395 (0.05)	0.989 (0.05)	0.977 (0.08)	0.977 (0.04)	−0.091 (0.06)	0.129 (0.07)	0.208 (0.07)	0.204 (0.07)	0.201 (0.08)	0.193 (0.30)
MCD	7.829 (0.12)	0.151 (0.05)	0.407 (0.09)	0.359 (0.04)	0.997 (0.05)	0.978 (0.06)	0.986 (0.07)	−0.216 (0.04)	0.121 (0.02)	0.166 (0.05)	0.174 (0.31)	0.071 (0.07)	0.176 (0.10)
NKE	3.779 (0.09)	0.133 (0.05)	0.357 (0.04)	0.539 (0.32)	0.966 (0.12)	0.981 (0.23)	0.982 (0.03)	−0.191 (0.31)	0.141 (0.21)	0.178 (0.02)	0.177 (0.20)	0.167 (0.04)	0.247 (0.21)
MRK	3.819 (0.08)	0.062 (0.04)	0.177 (0.21)	0.484 (0.22)	0.995 (0.25)	0.977 (0.39)	0.977 (0.38)	−0.252 (0.04)	0.131 (0.03)	0.157 (0.22)	0.177 (0.10)	0.166 (0.26)	0.174 (0.10)
MSFT	5.829 (0.08)	0.039 (0.20)	0.307 (0.34)	0.429 (0.26)	0.983 (0.28)	0.981 (0.09)	0.978 (0.32)	−0.121 (0.04)	0.141 (0.03)	0.287 (0.16)	0.237 (0.05)	0.076 (0.23)	0.137 (0.09)
WBA	2.798 (0.12)	0.162 (0.06)	0.287 (0.02)	0.489 (0.12)	0.939 (0.23)	0.977 (0.03)	0.974 (0.23)	−0.081 (0.04)	0.133 (0.34)	0.074 (0.12)	0.077 (0.09)	0.076 (0.11)	0.137 (0.06)
UTX	5.919 (0.06)	0.360 (0.05)	0.407 (0.03)	0.689 (0.32)	0.998 (0.24)	0.979 (0.06)	0.981 (0.24)	−0.181 (0.05)	0.135 (0.06)	0.074 (0.10)	0.077 (0.08)	0.176 (0.09)	0.177 (0.08)
PG	4.789 (0.05)	0.052 (0.03)	0.187 (0.13)	0.458 (0.12)	0.986 (0.04)	0.984 (0.05)	0.979 (0.04)	0.191 (0.03)	−0.301 (0.09)	0.184 (0.05)	0.263 (0.05)	0.278 (0.04)	0.188 (0.10)
PFE	1.859 (0.08)	0.021 (0.08)	0.227 (0.18)	0.679 (0.16)	0.986 (0.10)	0.981 (0.16)	0.987 (0.37)	−0.104 (0.24)	0.141 (0.05)	0.276 (0.07)	0.177 (0.07)	0.367 (0.09)	0.312 (0.08)
TRV	5.684 (0.05)	0.989 (0.01)	0.427 (0.06)	0.563 (0.21)	0.956 (0.04)	0.983 (0.23)	0.978 (0.35)	−0.172 (0.03)	0.136 (0.07)	0.387 (0.08)	0.337 (0.21)	0.176 (0.07)	0.207 (0.30)
XOM	2.659 (0.11)	0.103 (0.05)	0.437 (0.03)	0.449 (0.03)	0.988 (0.23)	0.979 (0.33)	0.989 (0.32)	−0.131 (0.03)	0.144 (0.31)	0.247 (0.07)	0.230 (0.20)	0.186 (0.10)	0.247 (0.09)
WMT	4.685 (0.03)	0.089 (0.03)	0.337 (0.23)	0.729 (0.05)	0.985 (0.07)	0.979 (0.04)	0.984 (0.14)	−0.281 (0.23)	0.204 (0.12)	0.189 (0.05)	0.271 (0.06)	0.085 (0.07)	0.109 (0.06)
UNH	9.525 (0.04)	0.134 (0.09)	0.517 (0.04)	0.359 (0.22)	0.996 (0.11)	0.983 (0.03)	0.980 (0.12)	−0.121 (0.34)	0.141 (0.13)	0.085 (0.10)	0.122 (0.05)	0.187 (0.28)	0.273 (0.04)
V	7.642 (0.22)	0.109 (0.09)	0.297 (0.05)	0.479 (0.21)	0.998 (0.05)	0.982 (0.28)	0.984 (0.21)	0.194 (0.23)	−0.171 (0.05)	0.295 (0.20)	0.376 (0.09)	0.085 (0.08)	0.122 (0.04)
VZ	2.437 (0.09)	0.051 (0.04)	0.187 (0.23)	0.565 (0.04)	0.995 (0.04)	0.982 (0.14)	0.073 (0.06)	0.164 (0.04)	−0.171 (0.03)	0.099 (0.09)	0.080 (0.10)	0.187 (0.10)	0.273 (0.10)

Table 8: Estimated informational trading parameters for DIA ETF

var																														
	$\psi^{1}_{a^f a}$	$\psi^{1}_{a^f b}$	$\psi^{2}_{a^f a}$	$\psi^{2}_{a^f b}$	$\psi^{3}_{a^f a}$	$\psi^{3}_{a^f b}$	$\psi^{4}_{a^f a}$	$\psi^{4}_{a^f b}$	$\psi^{5}_{a^f a}$	$\psi^{5}_{a^f b}$	$\psi^{6}_{a^f a}$	$\psi^{6}_{a^f b}$	$\psi^{7}_{a^f a}$	$\psi^{7}_{a^f b}$	$\psi^{8}_{a^f a}$	$\psi^{8}_{a^f b}$	$\psi^{9}_{a^f a}$	$\psi^{9}_{a^f b}$	$\psi^{10}_{a^f a}$	$\psi^{10}_{a^f b}$	$\psi^{11}_{a^f a}$	$\psi^{11}_{a^f b}$	$\psi^{12}_{a^f a}$	$\psi^{12}_{a^f b}$	$\psi^{13}_{a^f a}$	$\psi^{13}_{a^f b}$	$\psi^{14}_{a^f a}$	$\psi^{14}_{a^f b}$	$\psi^{15}_{a^f a}$	$\psi^{15}_{a^f b}$
$a_t^f - p_t^f$	0.079 (0.09)	0.074 (0.05)	0.078 (0.10)	0.077 (0.23)	0.079 (0.08)	0.080 (0.09)	0.078 (0.17)	0.078 (0.06)	0.089 (0.20)	0.078 (0.05)	0.077 (0.10)	0.081 (0.09)	0.077 (0.0)	0.080 (0.20)	0.100 (0.06)	0.087 (0.09)	0.078 (0.05)	0.079 (0.10)	0.078 (0.05)	0.081 (0.17)	0.078 (0.09)	0.078 (0.25)	0.078 (0.06)	0.116 (0.20)	0.286 (0.05)	0.082 (0.10)	0.079 (0.09)	0.079 (0.0)	0.078 (0.20)	0.078 (0.03)
	$\psi^{16}_{a^f a}$	$\psi^{16}_{a^f b}$	$\psi^{17}_{a^f a}$	$\psi^{17}_{a^f b}$	$\psi^{18}_{a^f a}$	$\psi^{18}_{a^f b}$	$\psi^{19}_{a^f a}$	$\psi^{19}_{a^f b}$	$\psi^{20}_{a^f a}$	$\psi^{20}_{a^f b}$	$\psi^{21}_{a^f a}$	$\psi^{21}_{a^f b}$	$\psi^{22}_{a^f a}$	$\psi^{22}_{a^f b}$	$\psi^{23}_{a^f a}$	$\psi^{23}_{a^f b}$	$\psi^{24}_{a^f a}$	$\psi^{24}_{a^f b}$	$\psi^{25}_{a^f a}$	$\psi^{25}_{a^f b}$	$\psi^{26}_{a^f a}$	$\psi^{26}_{a^f b}$	$\psi^{27}_{a^f a}$	$\psi^{27}_{a^f b}$	$\psi^{28}_{a^f a}$	$\psi^{28}_{a^f b}$	$\psi^{29}_{a^f a}$	$\psi^{29}_{a^f b}$	$\psi^{30}_{a^f a}$	$\psi^{30}_{a^f b}$
	0.089 (0.09)	0.091 (0.05)	0.079 (0.10)	0.079 (0.08)	0.088 (0.18)	0.080 (0.09)	0.080 (0.19)	0.079 (0.06)	0.084 (0.20)	0.077 (0.05)	0.106 (0.10)	0.077 (0.09)	0.077 (0.0)	0.087 (0.20)	0.081 (0.04)	0.082 (0.09)	0.116 (0.05)	0.078 (0.10)	0.077 (0.08)	0.155 (0.29)	0.077 (0.09)	0.084 (0.11)	0.078 (0.06)	0.077 (0.20)	0.077 (0.05)	0.078 (0.10)	0.078 (0.09)	0.077 (0.0)	0.077 (0.20)	0.077 (0.07)
	$\psi^{1}_{b^f a}$	$\psi^{1}_{b^f b}$	$\psi^{2}_{b^f a}$	$\psi^{2}_{b^f b}$	$\psi^{3}_{b^f a}$	$\psi^{3}_{b^f b}$	$\psi^{4}_{b^f a}$	$\psi^{4}_{b^f b}$	$\psi^{5}_{b^f a}$	$\psi^{5}_{b^f b}$	$\psi^{6}_{b^f a}$	$\psi^{6}_{b^f b}$	$\psi^{7}_{b^f a}$	$\psi^{7}_{b^f b}$	$\psi^{8}_{b^f a}$	$\psi^{8}_{b^f b}$	$\psi^{9}_{b^f a}$	$\psi^{9}_{b^f b}$	$\psi^{10}_{b^f a}$	$\psi^{10}_{b^f b}$	$\psi^{11}_{b^f a}$	$\psi^{11}_{b^f b}$	$\psi^{12}_{b^f a}$	$\psi^{12}_{b^f b}$	$\psi^{13}_{b^f a}$	$\psi^{13}_{b^f b}$	$\psi^{14}_{b^f a}$	$\psi^{14}_{b^f b}$	$\psi^{15}_{b^f a}$	$\psi^{15}_{b^f b}$
$p_t^f - b_t^f$	0.080 (0.06)	0.078 (0.11)	0.078 (0.09)	0.076 (0.20)	0.081 (0.05)	0.078 (0.20)	0.078 (0.10)	0.078 (0.09)	0.078 (0.07)	0.078 (0.06)	0.079 (0.20)	0.079 (0.08)	0.078 (0.10)	0.078 (0.05)	0.079 (0.09)	0.083 (0.06)	0.078 (0.17)	0.078 (0.09)	0.078 (0.20)	0.083 (0.05)	0.078 (0.10)	0.082 (0.10)	0.077 (0.07)	0.078 (0.03)	0.078 (0.06)	0.089 (0.20)	0.078 (0.05)	0.079 (0.10)	0.078 (0.05)	0.097 (0.09)
	$\psi^{16}_{b^f a}$	$\psi^{16}_{b^f b}$	$\psi^{17}_{b^f a}$	$\psi^{17}_{b^f b}$	$\psi^{18}_{b^f a}$	$\psi^{18}_{b^f b}$	$\psi^{19}_{b^f a}$	$\psi^{19}_{b^f b}$	$\psi^{20}_{b^f a}$	$\psi^{20}_{b^f b}$	$\psi^{21}_{b^f a}$	$\psi^{21}_{b^f b}$	$\psi^{22}_{b^f a}$	$\psi^{22}_{b^f b}$	$\psi^{23}_{b^f a}$	$\psi^{23}_{b^f b}$	$\psi^{24}_{b^f a}$	$\psi^{24}_{b^f b}$	$\psi^{25}_{b^f a}$	$\psi^{25}_{b^f b}$	$\psi^{26}_{b^f a}$	$\psi^{26}_{b^f b}$	$\psi^{27}_{b^f a}$	$\psi^{27}_{b^f b}$	$\psi^{28}_{b^f a}$	$\psi^{28}_{b^f b}$	$\psi^{29}_{b^f a}$	$\psi^{29}_{b^f b}$	$\psi^{30}_{b^f a}$	$\psi^{30}_{b^f b}$
	0.102 (0.06)	0.089 (0.22)	0.080 (0.09)	0.167 (0.20)	0.078 (0.05)	0.077 (0.30)	0.079 (0.10)	0.082 (0.10)	0.095 (0.07)	0.079 (0.06)	0.084 (0.20)	0.077 (0.06)	0.081 (0.10)	0.138 (0.05)	0.080 (0.09)	0.077 (0.06)	0.096 (0.24)	0.169 (0.09)	0.079 (0.20)	0.095 (0.05)	0.077 (0.08)	0.086 (0.10)	0.077 (0.08)	0.078 (0.09)	0.077 (0.06)	0.078 (0.20)	0.077 (0.03)	0.077 (0.10)	0.077 (0.0)5	0.077 (0.09)

Table 9: Estimated informational trading parameters for AAPL

var																														
	ϕ^{i}_{aa}	ϕ^{i}_{ab}	ϕ^{i2}_{aa}	ϕ^{i2}_{ab}	ϕ^{i3}_{aa}	ϕ^{i3}_{ab}	ϕ^{i4}_{aa}	ϕ^{i4}_{ab}	ϕ^{i5}_{aa}	ϕ^{i5}_{ab}	ϕ^{i6}_{aa}	ϕ^{i6}_{ab}	ϕ^{i7}_{aa}	ϕ^{i7}_{ab}	ϕ^{i8}_{aa}	ϕ^{i8}_{ab}	ϕ^{i9}_{aa}	ϕ^{i9}_{ab}	ϕ^{i10}_{aa}	ϕ^{i10}_{ab}	ϕ^{i11}_{aa}	ϕ^{i11}_{ab}	ϕ^{i12}_{aa}	ϕ^{i12}_{ab}	ϕ^{i13}_{aa}	ϕ^{i13}_{ab}	ϕ^{i14}_{aa}	ϕ^{i14}_{ab}	ϕ^{i15}_{aa}	ϕ^{i15}_{ab}
$a_t - p_t$	0.541 (0.09)	0.089 (0.15)	0.087 (0.08)	0.116 (0.07)	0.096 (0.07)	0.117 (0.09)	0.092 (0.10)	0.097 (0.09)	0.087 (0.05)	0.084 (0.06)	0.092 (0.04)	0.080 (0.07)	0.078 (0.06)	0.107 (0.09)	0.138 (0.10)	0.117 (0.09)	0.107 (0.08)	0.167 (0.07)	0.132 (0.06)	0.120 (0.09)	0.127 (0.11)	0.099 (0.04)	0.086 (0.07)	0.167 (0.12)	0.109 (0.09)	0.092 (0.07)	0.157 (0.06)	0.106 (0.04)	0.106 (0.08)	0.090 (0.08)
	ϕ^{i16}_{aa}	ϕ^{i16}_{ab}	ϕ^{i17}_{aa}	ϕ^{i17}_{ab}	ϕ^{i18}_{aa}	ϕ^{i18}_{ab}	ϕ^{i19}_{aa}	ϕ^{i19}_{ab}	ϕ^{i20}_{aa}	ϕ^{i20}_{ab}	ϕ^{i21}_{aa}	ϕ^{i21}_{ab}	ϕ^{i22}_{aa}	ϕ^{i22}_{ab}	ϕ^{i23}_{aa}	ϕ^{i23}_{ab}	ϕ^{i24}_{aa}	ϕ^{i24}_{ab}	ϕ^{i25}_{aa}	ϕ^{i25}_{ab}	ϕ^{i26}_{aa}	ϕ^{i26}_{ab}	ϕ^{i27}_{aa}	ϕ^{i27}_{ab}	ϕ^{i28}_{aa}	ϕ^{i28}_{ab}	ϕ^{i29}_{aa}	ϕ^{i29}_{ab}	ϕ^{i30}_{aa}	ϕ^{i30}_{ab}
	0.089 (0.07)	0.078 (0.085)	0.079 (0.07)	0.089 (0.08)	0.087 (0.05)	0.096 (0.04)	0.112 (0.067)	0.097 (0.06)	0.084 (0.05)	0.080 (0.08)	0.080 (0.04)	0.079 (0.089)	0.078 (0.07)	0.108 (0.09)	0.081 (0.06)	0.082 (0.07)	0.081 (0.057)	0.079 (0.06)	0.081 (0.07)	0.088 (0.06)	0.080 (0.06)	0.084 (0.054)	0.095 (0.07)	0.087 (0.08)	0.080 (0.08)	0.107 (0.10)	0.082 (0.079)	0.092 (0.07)	0.100 (0.09)	0.116 (0.10)
	ϕ^{i}_{ba}	ϕ^{i}_{bb}	ϕ^{i2}_{ba}	ϕ^{i2}_{bb}	ϕ^{i3}_{ba}	ϕ^{i3}_{bb}	ϕ^{i4}_{ba}	ϕ^{i4}_{bb}	ϕ^{i5}_{ba}	ϕ^{i5}_{bb}	ϕ^{i6}_{ba}	ϕ^{i6}_{bb}	ϕ^{i7}_{ba}	ϕ^{i7}_{bb}	ϕ^{i8}_{ba}	ϕ^{i8}_{bb}	ϕ^{i9}_{ba}	ϕ^{i9}_{bb}	ϕ^{i10}_{ba}	ϕ^{i10}_{bb}	ϕ^{i11}_{ba}	ϕ^{i11}_{bb}	ϕ^{i12}_{ba}	ϕ^{i12}_{bb}	ϕ^{i13}_{ba}	ϕ^{i13}_{bb}	ϕ^{i14}_{ba}	ϕ^{i14}_{bb}	ϕ^{i15}_{ba}	ϕ^{i15}_{bb}
$p_t - b_t$	0.057 (0.04)	0.447 (0.08)	0.127 (0.10)	0.092 (0.07)	0.086 (0.05)	0.100 (0.06)	0.026 (0.07)	0.098 (0.09)	0.108 (0.06)	0.097 (0.07)	0.111 (0.03)	0.079 (0.09)	0.085 (0.05)	0.127 (0.11)	0.095 (0.07)	0.139 (0.09)	0.167 (0.04)	0.107 (0.04)	0.081 (0.08)	0.083 (0.04)	0.107 (0.06)	0.102 (0.09)	0.096 (0.08)	0.147 (0.20)	0.149 (0.07)	0.122 (0.08)	0.087 (0.11)	0.069 (0.03)	0.086 (0.13)	0.101 (0.06)
	ϕ^{i16}_{ba}	ϕ^{i16}_{bb}	ϕ^{i17}_{ba}	ϕ^{i17}_{bb}	ϕ^{i18}_{ba}	ϕ^{i18}_{bb}	ϕ^{i19}_{ba}	ϕ^{i19}_{bb}	ϕ^{i20}_{ba}	ϕ^{i20}_{bb}	ϕ^{i21}_{ba}	ϕ^{i21}_{bb}	ϕ^{i22}_{ba}	ϕ^{i22}_{bb}	ϕ^{i23}_{ba}	ϕ^{i23}_{bb}	ϕ^{i24}_{ba}	ϕ^{i24}_{bb}	ϕ^{i25}_{ba}	ϕ^{i25}_{bb}	ϕ^{i26}_{ba}	ϕ^{i26}_{bb}	ϕ^{i27}_{ba}	ϕ^{i27}_{bb}	ϕ^{i28}_{ba}	ϕ^{i28}_{bb}	ϕ^{i29}_{ba}	ϕ^{i29}_{bb}	ϕ^{i30}_{ba}	ϕ^{i30}_{bb}
	0.082 (0.06)	0.098 (0.05)	0.080 (0.08)	0.096 (0.05)	0.077 (0.03)	0.088 (0.03)	0.085 (0.11)	0.082 (0.03)	0.079 (0.06)	0.093 (0.03)	0.086 (0.08)	0.080 (0.21)	0.081 (0.02)	0.080 (0.07)	0.078 (0.04)	0.102 (0.07)	0.090 (0.08)	0.079 (0.04)	0.082 (0.04)	0.077 (0.03)	0.083 (0.03)	0.086 (0.06)	0.080 (0.07)	0.117 (0.09)	0.083 (0.05)	0.087 (0.09)	0.078 (0.07)	0.082 (0.10)	0.084 (0.04)	0.099 (0.07)

For $\phi^{ij}_{..}$, $i = CVX$ and $j = CVX,\ AXP,\ BA, CSCO,\ CAT,\ DIS,\ IBM,\ DOW,\ GS,\ HD,\ KO,\ JPM,\ INTC,\ JNJ,\ MMM,\ MCD,\ NKE,\ MRK,\ MSFT,\ WBA,\ UTX,\ PG,\ PFE,\ TRV,\ XOM,\ WMT,\ UNH,\ V, VZ$

Table 10: Estimated informational trading parameters for BA

var																														
	ϕ^{i2}_{aa}	ϕ^{i2}_{ab}	ϕ^{i3}_{aa}	ϕ^{i3}_{ab}	ϕ^{i4}_{aa}	ϕ^{i4}_{ab}	ϕ^{i}_{aa}	ϕ^{i}_{ab}	ϕ^{i5}_{aa}	ϕ^{i5}_{ab}	ϕ^{i6}_{aa}	ϕ^{i6}_{ab}	ϕ^{i7}_{aa}	ϕ^{i7}_{ab}	ϕ^{i8}_{aa}	ϕ^{i8}_{ab}	ϕ^{i9}_{aa}	ϕ^{i9}_{ab}	ϕ^{i10}_{aa}	ϕ^{i10}_{ab}	ϕ^{i11}_{aa}	ϕ^{i11}_{ab}	ϕ^{i12}_{aa}	ϕ^{i12}_{ab}	ϕ^{i13}_{aa}	ϕ^{i13}_{ab}	ϕ^{i14}_{aa}	ϕ^{i14}_{ab}	ϕ^{i15}_{aa}	ϕ^{i15}_{ab}
$a_t - p_t$	0.076 (0.15)	0.082 (0.15)	0.078 (0.04)	0.077 (0.03)	0.081 (0.15)	0.079 (0.06)	0.563 (0.32)	0.118 (0.03)	0.080 (0.07)	0.084 (0.04)	0.062 (0.05)	0.082 (0.04)	0.083 (0.11)	0.077 (0.05)	0.103 (0.03)	0.077 (0.05)	0.078 (0.06)	−0.012 (0.09)	0.022 (0.04)	0.078 (0.04)	0.081 (0.08)	0.078 (0.25)	0.086 (0.04)	0.086 (0.04)	0.081 (0.02)	0.078 (0.07)	0.085 (0.07)	0.080 (0.25)	0.080 (0.04)	0.078 (0.06)
	ϕ^{i16}_{aa}	ϕ^{i16}_{ab}	ϕ^{i17}_{aa}	ϕ^{i17}_{ab}	ϕ^{i18}_{aa}	ϕ^{i18}_{ab}	ϕ^{i19}_{aa}	ϕ^{i19}_{ab}	ϕ^{i20}_{aa}	ϕ^{i20}_{ab}	ϕ^{i21}_{aa}	ϕ^{i21}_{ab}	ϕ^{i22}_{aa}	ϕ^{i22}_{ab}	ϕ^{i23}_{aa}	ϕ^{i23}_{ab}	ϕ^{i24}_{aa}	ϕ^{i24}_{ab}	ϕ^{i25}_{aa}	ϕ^{i25}_{ab}	ϕ^{i26}_{aa}	ϕ^{i26}_{ab}	ϕ^{i27}_{aa}	ϕ^{i27}_{ab}	ϕ^{i28}_{aa}	ϕ^{i28}_{ab}	ϕ^{i29}_{aa}	ϕ^{i29}_{ab}	ϕ^{i30}_{aa}	ϕ^{i30}_{ab}
	0.080 (0.06)	0.080 (0.07)	0.081 (0.03)	0.108 (0.09)	0.078 (0.05)	0.075 (0.04)	0.088 (0.15)	0.097 (0.08)	0.084 (0.08)	0.080 (0.06)	0.080 (0.04)	0.086 (0.06)	0.082 (0.04)	0.080 (0.07)	0.073 (0.07)	0.080 (0.08)	0.086 (0.09)	0.084 (0.06)	0.073 (0.05)	0.079 (0.07)	0.086 (0.03)	0.079 (0.05)	0.095 (0.06)	0.076 (0.03)	0.078 (0.04)	0.082 (0.07)	0.084 (0.03)	−0.004 (0.08)	0.080 (0.09)	0.136 (0.08)
	ϕ^{i2}_{ba}	ϕ^{i2}_{bb}	ϕ^{i3}_{ba}	ϕ^{i3}_{bb}	ϕ^{i4}_{ba}	ϕ^{i4}_{bb}	ϕ^{i}_{ba}	ϕ^{i}_{bb}	ϕ^{i5}_{ba}	ϕ^{i5}_{bb}	ϕ^{i6}_{ba}	ϕ^{i6}_{bb}	ϕ^{i7}_{ba}	ϕ^{i7}_{bb}	ϕ^{i8}_{ba}	ϕ^{i8}_{bb}	ϕ^{i9}_{ba}	ϕ^{i9}_{bb}	ϕ^{i10}_{ba}	ϕ^{i10}_{bb}	ϕ^{i11}_{ba}	ϕ^{i11}_{bb}	ϕ^{i12}_{ba}	ϕ^{i12}_{bb}	ϕ^{i13}_{ba}	ϕ^{i13}_{bb}	ϕ^{i14}_{ba}	ϕ^{i14}_{bb}	ϕ^{i15}_{ba}	ϕ^{i15}_{bb}
$p_t - b_t$	0.077 (0.08)	0.068 (0.03)	0.083 (0.02)	0.077 (0.05)	0.082 (0.08)	0.068 (0.03)	0.100 (0.21)	0.653 (0.32)	0.108 (0.09)	0.083 (0.05)	0.110 (0.02)	0.075 (0.04)	0.085 (0.09)	0.079 (0.04)	0.082 (0.09)	0.139 (0.26)	0.167 (0.11)	0.079 (0.05)	0.080 (0.02)	0.083 (0.06)	0.078 (0.01)	0.102 (0.11)	0.058 (0.05)	0.083 (0.06)	0.078 (0.04)	0.078 (0.04)	0.077 (0.06)	0.069 (0.08)	0.086 (0.05)	0.083 (0.08)
	ϕ^{i16}_{ba}	ϕ^{i16}_{bb}	ϕ^{i17}_{ba}	ϕ^{i17}_{bb}	ϕ^{i18}_{ba}	ϕ^{i18}_{bb}	ϕ^{i19}_{ba}	ϕ^{i19}_{bb}	ϕ^{i20}_{ba}	ϕ^{i20}_{bb}	ϕ^{i21}_{ba}	ϕ^{i21}_{bb}	ϕ^{i22}_{ba}	ϕ^{i22}_{bb}	ϕ^{i23}_{ba}	ϕ^{i23}_{bb}	ϕ^{i24}_{ba}	ϕ^{i24}_{bb}	ϕ^{i25}_{ba}	ϕ^{i25}_{bb}	ϕ^{i26}_{ba}	ϕ^{i26}_{bb}	ϕ^{i27}_{ba}	ϕ^{i27}_{bb}	ϕ^{i28}_{ba}	ϕ^{i28}_{bb}	ϕ^{i29}_{ba}	ϕ^{i29}_{bb}	ϕ^{i30}_{ba}	ϕ^{i30}_{bb}
	0.082 (0.08)	0.086 (0.05)	0.080 (0.07)	0.058 (0.04)	0.078 (0.05)	0.078 (0.05)	0.085 (0.06)	0.079 (0.08)	0.079 (0.07)	0.061 (0.04)	0.086 (0.06)	0.080 (0.07)	0.081 (0.07)	0.080 (0.07)	0.078 (0.08)	0.109 (0.25)	0.138 (0.21)	0.075 (0.05)	0.072 (0.03)	0.077 (0.04)	0.071 (0.03)	0.086 (0.04)	0.083 (0.04)	0.083 (0.07)	0.080 (0.07)	0.095 (0.03)	0.073 (0.05)	0.082 (0.05)	0.147 (0.07)	0.084 (0.04)

For $\phi^{ij}_{..}$, $i = BA$ and $j = AAPL,\ CVX,\ AXP,\ CSCO,\ CAT,\ DIS,\ IBM,\ DOW,\ GS,\ HD,\ KO,\ JPM,\ INTC,\ JNJ,\ MMM,\ MCD,\ NKE,\ MRK,\ MSFT,\ WBA,\ UTX,\ PG,\ PFE,\ TRV,\ XOM,\ WMT,\ UNH,\ V, VZ$

Appendix C3: Estimated Parameters for SMH and its Underlying Assets.

Table 11: **Estimated dynamic state-space model for SMH and its 25 underlying assets**

ETF		β_f			ν^f	ν_a^f	ν_b^f	λ_1^f	λ_2^f	ψ_{a^f}	$\psi_{a^f b^f}$	$\psi_{b^f a^f}$	ψ_{b^f}
SMH		1.044 (0.03)			0.998 (0.04)	0.994 (0.06)	0.999 (0.05)	0.179 (0.02)	−0.161 (0.04)	0.883 (0.21)	0.090 (0.16)	0.094 (0.11)	0.790 (0.26)
STOCK	μ_m	β	η	δ	ν_p	ν_a	ν_b	λ_a	λ_b	θ_{aa^f}	θ_{ab^f}	θ_{ba^f}	θ_{bb^f}
TSM	−5.679 (0.09)	0.413 (0.08)	0.619 (0.02)	0.529 (0.12)	0.987 (0.05)	0.983 (0.23)	0.989 (0.32)	0.130 (0.04)	−0.211 (0.29)	0.080 (0.24)	0.084 (0.02)	0.083 (0.20)	0.086 (0.03)
INTC	3.716 (0.02)	0.033 (0.05)	0.677 (0.21)	0.474 (0.07)	0.989 (0.06)	0.958 (0.07)	0.989 (0.07)	0.217 (0.29)	−0.183 (0.01)	0.091 (0.04)	0.085 (0.04)	0.285 (0.06)	0.288 (0.06)
NVDA	2.801 (0.01)	1.883 (0.09)	0.907 (0.23)	0.474 (0.03)	0.996 (0.16)	0.984 (0.06)	0.981 (0.19)	0.131 (0.05)	−0.513 (0.32)	0.297 (0.07)	0.296 (0.02)	0.196 (0.09)	0.396 (0.03)
AMD	7.773 (0.12)	0.465 (0.22)	0.759 (0.28)	0.409 (0.08)	0.992 (0.18)	0.988 (0.08)	0.996 (0.05)	0.134 (0.09)	−0.511 (0.09)	0.280 (0.09)	0.176 (0.06)	0.221 (0.05)	0.273 (0.03)
ASML	3.917 (0.08)	0.770 (0.05)	0.561 (0.01)	0.429 (0.03)	0.997 (0.23)	0.987 (0.04)	0.984 (0.05)	−0.241 (0.05)	0.194 (0.05)	0.216 (0.07)	0.240 (0.04)	0.209 (0.03)	0.232 (0.06)
TXN	−3.859 (0.03)	1.719 (0.27)	0.629 (0.06)	0.429 (0.19)	0.996 (0.05)	0.982 (0.22)	0.983 (0.05)	−0.211 (0.04)	0.194 (0.29)	0.110 (0.05)	0.196 (0.08)	0.202 (0.08)	0.104 (0.07)
QCOM	9.649 (0.20)	1.614 (0.01)	0.753 (0.04)	0.509 (0.09)	0.998 (0.07)	0.978 (0.03)	0.998 (0.11)	0.164 (0.03)	−0.218 (0.09)	0.194 (0.03)	0.104 (0.05)	0.181 (0.09)	0.102 (0.07)
MU	2.833 (0.04)	2.360 (0.02)	0.743 (0.11)	0.594 (0.05)	0.991 (0.03)	0.992 (0.03)	0.984 (0.01)	0.134 (0.12)	−0.084 (0.29)	0.394 (0.05)	0.104 (0.04)	0.281 (0.08)	0.102 (0.04)
AVGO	2.349 (0.08)	1.928 (0.33)	0.801 (0.11)	0.314 (0.01)	0.998 (0.22)	0.928 (0.17)	0.971 (0.09)	0.132 (0.20)	−0.425 (0.17)	0.345 (0.06)	0.291 (0.05)	0.273 (0.08)	0.199 (0.03)
NXPI	9.621 (0.22)	1.924 (0.29)	0.868 (0.05)	0.609 (0.03)	0.997 (0.21)	0.977 (0.01)	0.978 (0.04)	0.138 (0.19)	−0.411 (0.09)	0.101 (0.06)	0.860 (0.08)	0.114 (0.10)	0.123 (0.04)
LRCX	−4.124 (0.27)	0.307 (0.05)	0.640 (0.08)	0.579 (0.03)	0.995 (0.30)	0.984 (0.10)	0.982 (0.28)	0.137 (0.05)	−0.292 (0.15)	0.112 (0.21)	0.105 (0.04)	0.095 (0.28)	0.296 (0.05)
AMAT	5.462 (0.11)	0.853 (0.30)	0.923 (0.20)	0.614 (0.29)	0.996 (0.09)	0.984 (0.22)	0.975 (0.23)	0.130 (0.21)	−0.4133 (0.23)	0.105 (0.08)	0.465 (0.08)	0.394 (0.09)	0.104 (0.06)
ADI	5.854 (0.21)	0.853 (0.04)	0.816 (0.05)	0.552 (0.03)	0.989 (0.23)	0.984 (0.11)	0.984 (0.05)	0.164 (0.13)	−0.607 (0.16)	0.181 (0.02)	0.102 (0.08)	0.172 (0.09)	0.104 (0.09)
KLAC	1.844 (0.04)	0.812 (0.04)	0.824 (0.05)	0.379 (0.05)	0.996 (0.21)	0.983 (0.25)	0.984 (0.06)	0.156 (0.09)	−0.213 (0.03)	0.296 (0.16)	0.296 (0.09)	0.294 (0.18)	0.296 (0.23)
XLNX	6.579 (0.07)	0.741 (0.29)	0.549 (0.02)	0.509 (0.09)	0.992 (0.04)	0.984 (0.17)	0.983 (0.11)	−0.906 (0.04)	0.164 (0.30)	0.102 (0.19)	0.104 (0.20)	0.156 (0.26)	0.112 (0.16)
STM	1.774 (0.04)	0.850 (0.14)	0.623 (0.03)	0.494 (0.23)	0.990 (0.05)	0.989 (0.05)	0.984 (0.23)	−0.767 (0.09)	0.611 (0.12)	0.197 (0.10)	0.195 (0.22)	0.197 (0.09)	0.345 (0.23)
MCHP	4.451 (0.01)	0.765 (0.14)	0.558 (0.20)	0.628 (0.34)	0.994 (0.30)	0.985 (0.04)	0.984 (0.03)	0.131 (0.21)	−0.105 (0.12)	0.197 (0.08)	0.395 (0.10)	0.197 (0.08)	0.245 (0.20)
CDNS	8.018 (0.21)	0.770 (0.04)	0.308 (0.08)	0.574 (0.05)	0.993 (0.03)	0.984 (0.08)	0.985 (0.05)	0.132 (0.21)	−0.092 (0.35)	0.117 (0.10)	0.197 (0.32)	0.171 (0.06)	0.185 (0.09)
SWKS	5.640 (0.04)	0.578 (0.06)	0.506 (0.22)	0.618 (0.11)	0.995 (0.24)	0.987 (0.09)	0.988 (0.04)	0.132 (0.33)	−0.225 (0.09)	0.296 (0.10)	0.197 (0.10)	0.197 (0.06)	0.495 (0.07)
MXIM	5.691 (0.02)	0.420 (0.05)	0.313 (0.06)	0.361 (0.05)	0.997 (0.09)	0.988 (0.09)	0.984 (0.29)	−0.108 (0.19)	0.146 (0.07)	0.166 (0.07)	0.101 (0.05)	0.380 (0.09)	0.310 (0.07)
TER	3.911 (0.07)	0.503 (0.04)	0.150 (0.04)	0.388 (0.32)	0.992 (0.23)	0.984 (0.05)	0.995 (0.04)	0.164 (0.03)	−0.376 (0.05)	0.294 (0.03)	0.104 (0.08)	0.102 (0.10)	0.104 (0.06)
MRVL	5.710 (0.03)	0.630 (0.07)	0.921 (0.02)	0.597 (0.10)	0.999 (0.03)	0.983 (0.06)	0.984 (0.05)	0.153 (0.04)	−0.018 (0.09)	0.074 (0.11)	0.194 (0.09)	0.193 (0.28)	0.102 (0.09)
QRVO	7.515 (0.06)	0.782 (0.16)	0.567 (0.06)	0.689 (0.06)	0.992 (0.05)	0.989 (0.09)	0.995 (0.08)	0.432 (0.13)	−0.425 (0.29)	0.193 (0.05)	0.102 (0.24)	0.294 (0.07)	0.294 (0.08)
ON	6.939 (0.05)	0.837 (0.07)	0.643 (0.03)	0.484 (0.04)	0.990 (0.03)	0.984 (0.20)	0.874 (0.02)	0.181 (0.08)	0.128 (0.12)	0.396 (0.08)	0.397 (0.106)	0.297 (0.03)	0.195 (0.05)
OLED	3.528 (0.09)	2.010 (0.05)	0.781 (0.03)	0.434 (0.04)	0.993 (0.09)	0.984 (0.29)	0.991 (0.03)	0.131 (0.04)	−0.092 (0.22)	0.149 (0.08)	0.167 (0.07)	0.106 (0.02)	0.161 (0.03)

Tables (7) and (11) reports selected estimated parameters for both DIA and SMH ETFs. Due to space limitations, not all the estimated parameters are reported here. The rest are reported in Appendix C2. β_f corresponds to half bid-ask spreads for respective ETFs, and βs correspond to half bid-ask spreads for respective underlying assets of the ETFs.

Table 12: **SMH's Underlying Asset Price Impulse Response Function**

Price of Stock	*Ask Quote of SMH*	*Bid Quote of SMH*	*Price of SMH*
TSM	0.041^{**}	-0.098^{**}	0.029^{**}
INTC	0.052^{**}	-0.101^{**}	-0.042^{**}
NVDA	0.026^{**}	-0.099^{**}	0.043^{**}
AMD	0.032^{**}	-0.100^{**}	0.049^{**}
ASML	-0.056^{**}	0.102^{**}	-0.052^{**}
TXN	0.063^{*}	-0.088^{*}	0.029^{*}
QCOM	0.039^{**}	-0.069^{**}	0.031^{**}
MU	0.072^{**}	-0.069^{**}	-0.051^{**}
AVGO	0.036^{**}	-0.073^{**}	0.041^{**}
NXPI	0.037^{**}	-0.040^{**}	0.039^{**}
LRCX	0.045^{**}	-0.071^{**}	0.021^{**}
AMAT	-0.052^{**}	0.059^{**}	0.029^{**}
ADI	0.063^{*}	-0.098^{*}	0.033^{*}
KLAC	-0.039^{**}	0.088^{**}	-0.045^{**}
XLNX	0.032^{**}	-0.090^{**}	0.034^{**}
STM	0.025^{**}	-0.078^{**}	0.032^{**}
MCHP	0.042^{**}	-0.035^{**}	0.025^{**}
CDNS	-0.032^{**}	0.038^{**}	-0.035^{**}
SWKS	0.052^{**}	-0.037^{**}	-0.045^{**}
MXIM	0.047^{*}	-0.073^{*}	0.042^{*}
TER	-0.039^{**}	0.051^{**}	-0.029^{**}
MRVL	0.033^{**}	-0.066^{**}	0.036^{**}
QRVO	0.071^{**}	-0.053^{**}	0.029^{**}
ON	-0.056^{*}	0.053^{*}	0.042^{*}
OLED	0.034^{**}	-0.072^{**}	0.051^{**}

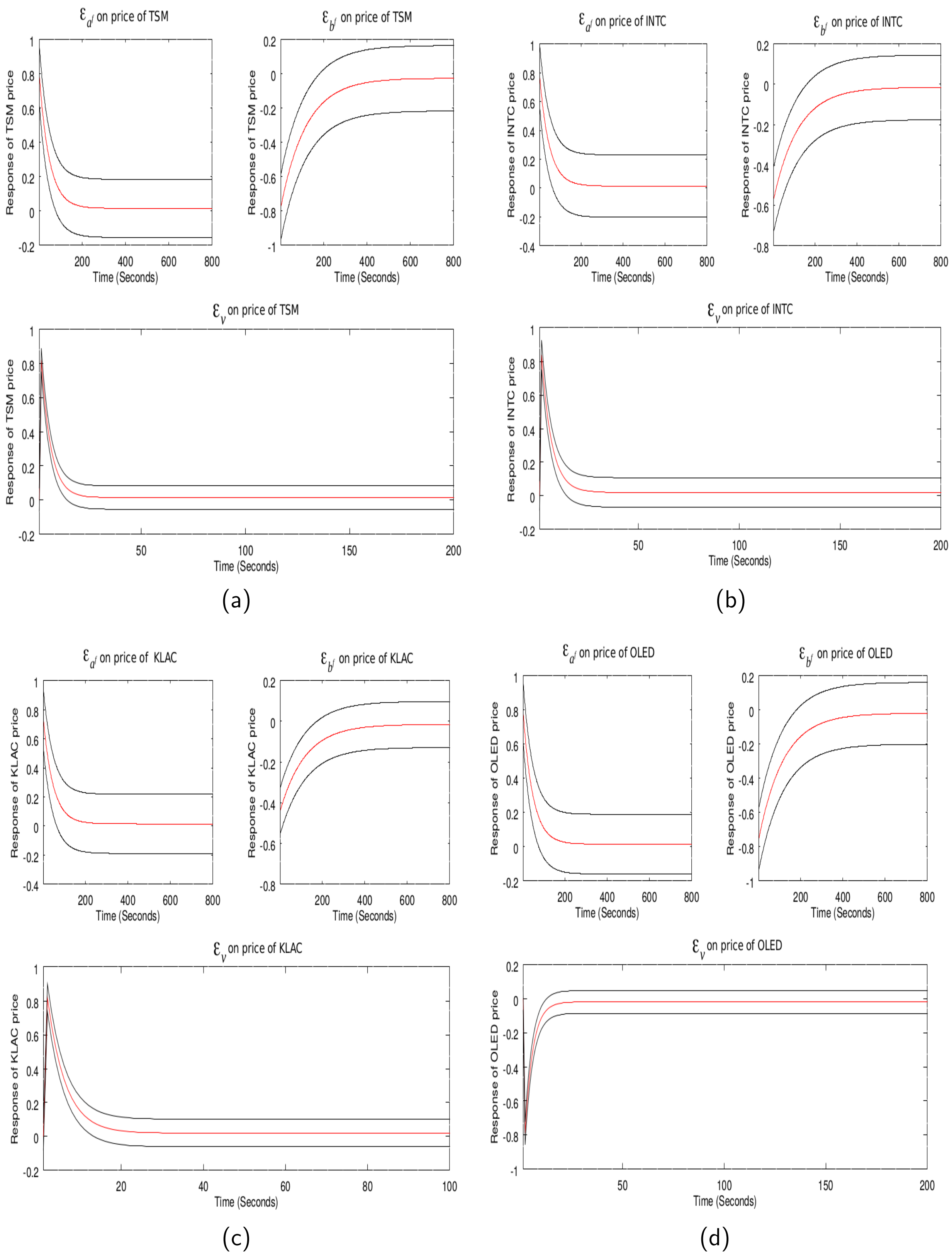
The table reports the underlying asset day-ahead average impulse response functions (IRFs) from the dynamic econometric state-space model developed in section 3 for the underlying assets of SMH ETF. The sample period is from January 2 to March 30, 2018, and the sampling frequency is 1-second. Columns 2 to 4 represent the variables being shocked by a unit, and rows correspond to the variables (underlying assets) affected by the unit shock from columns 2 to 4. The underlying asset prices at a one-second frequency are the mid-points of the underlying asset's best NBBs and best NBOs. Observations include all the best NBBs and NBOs between 9:30 and 16:30 EST. The IRFs are estimated for one trading day ahead (representing 9 hours or equivalently 32,400 seconds). For the estimates of the IRFs, a *(**) next to the estimates indicates that the estimated IRF differs from zero and is statistically significant at the 1% (5%) level using clustered standard errors by an asset and by day.

Table 13: **Estimated Extra Volatility and Speed of Arbitrage: SMH**

	Extra Volatility	*Speed of Arbitrage (h)*
SMH		**577.3**
TSM	**0.945** (0.046)	**40,81**
INTC	**0.798** (0.031)	**34.39**
NVDA	**0.432** (0.039)	**48.98**
AMD	**0.552** (0.076)	**25.19**
ASML	**0.472** (0.073)	**39.07**
TXN	**0.325** (0.040)	**33.43**
QCOM	**0.463** (0.064)	**30.29**
MU	**0.584** (0.025)	**40.01**
AVGO	**0.689** (0.057)	**29.30**
NXPI	**0.453** (0.030)	**35.42**
LRCX	**0.543** (0.058)	**34.32**
AMAT	**0.654** (0.068)	**39.58**
ADI	**0.543** (0.013)	**48.01**
KLAC	**0.338** (0.049)	**32.04**
XLNX	**0.751** (0.029)	**42.32**
STM	**0.630** (0.042)	**49.32**
MCHP	**0.535** (0.048)	**18.32**
CDNS	**0.399** (0.071)	**43.62**
SWKS	**0.469** (0.066)	**38.34**
MXIM	**0.685** (0.050)	**34.87**
TER	**0.453** (0.058)	**41.14**
MRVL	**0.409** (0.053)	**33.63**
QRVO	**0.556** (0.038)	**35.92**
ON	**0.563** (0.072)	**25.98**
OLED	**0.328** (0.040)	**21.09**

The table provides estimated excess transitory volatility brought to the 25 underlying assets of SMH and the speed of arbitrage trading. The first column presents the extra volatilities brought to the underlying asset when there is a liquidity shock to the SMH ETF. Parenthesized values in the first column are the respective additional transitory volatility brought to the price of SMH ETF when there is a liquidity shock to the respective underlying assets. On the other hand, column two shows the estimated time in seconds for the speed of arbitrage trading, assuming a half-life for the pricing error.

Figure 2: **Price Impulse Responses for Selected Underlying Assets of SMH**



Abstrakt

Důležitým aspektem obchodování s burzovně obchodovanými fondy (ETF z anglického Exchange-traded funds) je arbitrážní obchodní strategie prováděná autorizovanými účastníky trhu s cílem udržet ceny ETF v souladu s čistou hodnotou podkladových aktiv (NAV z anglického net asset value). ETF arbitráž je strategie, která využívá odchýlení ceny ETF od ceny podkladových aktiv. V této studii zkoumám vliv ETF arbitráže na podkladová aktiva ETF. Vytvářím dynamický state-space model, který sdruženě odhaduje dynamiku ceny ETF a jeho podkladových aktiv explicitním zahrnutím ETF arbitráže. Model odhaduji zvlášť pro Dow Jones Industrial Average ETF (DIA) a VanEck Vectors Semiconductor ETF (SMH). Empirické výsledky ukazují, že šoky ETF likvidity se šíří mezi podkladovými aktivy skrze mechanismus ETF arbitráže. Tyto šoky likvidity přidávají trvalou tranzitorní volatilitu k cenám podkladových aktiv. Zjišťuji, že jednotkový šok likvidity DIA vede k dodatečné volatilitě 0.1% až 0.93% cen podkladových aktiv DIA. Podobně jednotkový šok SMH přidá 0.33% až 0.95% dodatečné volatility k cenám podkladových aktiv SMH. Dále ukazuji, že korekce odchylky ceny ETF a NAV autorizovanými účastníky není okamžitá. V případě DIA trvá autorizovaným účastníkům přibližně 4 minuty provést ETF arbitráž, pro SMH je to přibližně 10 minut. Výsledky také naznačují, že ETF arbitráž urychluje proces objevování ceny na ETF trzích. Přibližně 74 % variace DIA a 67 % variace SMH je dána procesem objevování ceny.

Klíčová slova: burzovně obchodované fondy, ETF, podkladová aktiva, mechanismus ETF arbitráže, šoky likvidity, čistá cena aktiv, NAV, proces objevování ceny

Working Paper Series
ISSN 2788-0443

Individual researchers, as well as the on-line version of the CERGE-EI Working Papers (including their dissemination) were supported from institutional support RVO 67985998 from Economics Institute of the CAS, v. v. i.

Specific research support and/or other grants the researchers/publications benefited from are acknowledged at the beginning of the Paper.

Published by
Charles University, Center for Economic Research and Graduate Education (CERGE)
and
Economics Institute of the CAS, v. v. i. (EI)
CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.
Phone: + 420 224 005 153
Email: office@cerge-ei.cz
Web: https://www.cerge-ei.cz/

Editor: Byeongju Jeong

The paper is available online at https://www.cerge-ei.cz/working-papers/.

ISBN 978-80-7343-545-5 (Univerzita Karlova, Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-654-3 (Národohospodářský ústav AV ČR, v. v. i.)