



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Comerton-Forde, C;Marta, T

Title:

ETF effects: The role of primary versus secondary market activities

Date:

2025

Citation:

Comerton-Forde, C. & Marta, T. (2025). ETF effects: The role of primary versus secondary market activities. *Journal of financial markets*, 75, <https://doi.org/10.1016/j.finmar.2025.100983>.

Persistent Link:

<https://hdl.handle.net/11343/356784>

License:

[CC BY-NC-ND](#)

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

## Journal of Financial Markets

journal homepage: [www.elsevier.com/locate/finmar](http://www.elsevier.com/locate/finmar)ETF effects: The role of primary versus secondary market activities<sup>☆</sup>Carole Comerton-Forde<sup>a</sup>, Thomas Marta<sup>b</sup> \*<sup>a</sup> University of Melbourne, Australia<sup>b</sup> Wilfrid Laurier University - Lazaridis School of Business and Economics, Canada

## ARTICLE INFO

## JEL classification:

G14

G15

G23

## Keywords:

Exchange traded funds

HFT

Volatility

Arbitrage

Liquidity

## ABSTRACT

High-frequency traders (HFTs) dominate secondary market trading in exchange-traded funds (ETFs) but do not engage in ETF arbitrage. By contrast, primary market arbitrageurs enforce the law of one price, but their activities are infrequent and limited by arbitrage costs. We find that primary market activity is associated with increased volatility and illiquidity in overweighted ETF constituent stocks, while HFT activity is linked to narrower bid-ask spreads. Using a quasi-natural experiment in Japan, we show that while ETF primary market activity can temporarily disrupt market quality, the liquidity benefits of secondary market trading ultimately outweigh these negative effects.

## 1. Introduction

Assets under management (AUM) in exchange-traded funds (ETFs) have expanded from \$1 trillion in 2009 to \$15 trillion in December 2024, surpassing hedge funds and private equity.<sup>1</sup> This rapid growth has been accompanied by ETFs becoming a larger fraction of overall trading activity, accounting for 30% of trading volume in the United States (U.S.) and substantial shares of activity in European and Asia-Pacific markets.<sup>2</sup> ETFs are widely regarded as attractive investment vehicles, offering inexpensive access to liquid and diversified portfolios. However, the growing prominence of ETFs has raised concerns among regulators and academics about their potential to introduce non-fundamental volatility, herding behavior, and market fragility through their arbitrage mechanism (Ben-David et al., 2018; Bhattacharya and O'Hara, 2018; Brogaard et al., 2024).

A defining feature of ETFs is their reliance on authorized participants (APs), specialized arbitrageurs who engage in the creation and redemption of ETF units to align ETF prices with their net asset value (NAV). These primary market activities enforce the law of one price by ensuring that ETFs trade in line with their underlying assets. However, this process also requires APs to manage substantial inventories in the constituent stocks, which can affect market quality through adverse selection, volatility transmission, and liquidity extraction. Separately, ETFs are heavily traded in secondary markets, where high-frequency traders (HFTs) play a dominant role. Unlike APs, HFTs operate solely in secondary markets and focus on short-term profit opportunities, often managing

<sup>☆</sup> We gratefully acknowledge helpful comments and suggestions from the editor (Paolo Pasquariello), two anonymous referees, Diego Amaya, James Brugler, Jérôme Dugast, Thierry Foucault, Arseny Gorbenko, Katya Malinova, Sophie Moinas, Fabricio Perez, Fabrice Riva, Andriy Shkilko, Zhuo Zhong, and seminar participants at the RBA-FIRN Central Bank Workshop 2022, FMA 2022, McMaster, WLU, and UNSW Ph.D. microstructure workshop. We thank Jiayuan (Shin) Chen for his research assistance. Comerton-Forde is a Research Fellow at the Centre for Economic Policy Research, an economic consultant for the Australian Securities and Investments Commission and an Academic Advisor to the Plato Partnership.

\* Corresponding author.

E-mail addresses: [carole.comerton-forde@unimelb.edu.au](mailto:carole.comerton-forde@unimelb.edu.au) (C. Comerton-Forde), [tmarta@wlu.ca](mailto:tmarta@wlu.ca) (T. Marta).

<sup>1</sup> Authors' calculations based on Global ETF/ETP Database provided by [ETFGL.com](https://www.etf.com)

<sup>2</sup> [Brogaard et al. \(2024\)](https://www.brogaard.com) report that ETFs' share of trading activity in U.S. markets rose from 3% in 2002 to 30% in 2020. iShares reports that for the first quarter of 2023, the ETFs' share of trading activity is 11% in Europe and 13% in Asia-Pacific, see [iSharesGlobalETFfacts, Apr13, 2023](https://www.ishares.com/GlobalETFfacts/Apr13,2023).

<https://doi.org/10.1016/j.finmar.2025.100983>

Received 23 February 2024; Received in revised form 7 April 2025; Accepted 7 April 2025

Available online 9 May 2025

1386-4181/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

inventory risks through low-latency trading strategies.<sup>3</sup> We argue that these distinct roles—APs in primary markets and HFTs in secondary markets—have fundamentally different implications for the market quality of ETF constituent stocks. While the individual impacts of APs and HFTs are well researched, their collective impacts remain unexplored.

Two competing mechanisms explain how ETFs and HFTs influence the market quality of constituent stocks: the arbitrage channel and the information linkage channel. The arbitrage channel suggests that ETF primary market activities by APs directly affect constituent stocks by enforcing the law of one price but may also increase volatility and reduce liquidity through adverse selection and demand shocks (Budish et al., 2015; Malamud, 2015; Foucault et al., 2017). By contrast, researchers examining the information linkage channel argue that ETF secondary market activities, particularly those involving HFTs, improve market quality by indirectly diffusing information via price and quote updates (Grossman, 1992; Madhavan, 1996). However, the benefits of this channel may vary, as Bhattacharya and O'Hara (2018) show potential adverse effects in the ETF context. These contrasting hypotheses underscore the need to disentangle the impacts of ETF primary and secondary markets on constituent stock quality.

To better understand the complex ETF market dynamics, we examine three central questions. First, what factors drive ETF primary and secondary market activities, and what is their relative importance? Second, how do these activities influence the market quality of the constituent stocks? Third, do ETF primary market activities have a causal impact on the market quality of these stocks?

To answer the first two questions, we use trade-level data from Euronext Paris, where we are able to identify the activities of HFTs at a trade level. We document substantial differences in the frequency and determinants of primary and secondary market activities. ETF primary market activity occurs infrequently on only 32% of stock-days and is equivalent to only 13% of the ETF secondary market trading volume. By contrast, HFTs dominate ETF secondary markets, contributing to 74% of ETF trading volume, compared to 50% in the underlying stocks. We find that primary market activity is closely linked to ETF mispricing, as APs create or redeem ETF units to exploit arbitrage opportunities and help to enforce the law of one price. By contrast, HFT activity is not associated with daily ETF mispricing and primary market activity. These findings demonstrate the fundamentally distinct motivations and strategies of APs and HFTs.

To answer the second question, we use trade-level data for the constituent stocks of the main French index, the CAC 40. We show that the ETF exhibits higher liquidity than the weighted average liquidity of the constituent stocks, in part due to HFT activity. We examine the association between primary and secondary market activities and key measures of market quality, such as volatility, quoted spreads, and price efficiency. Our results show that ETF primary market activities are associated with increased volatility, wider quoted spreads, and reduced price efficiency for constituent stocks. These findings are consistent with the arbitrage channel, where arbitrage activity extracts liquidity from constituent stocks. Conversely, we find that HFT activity in the ETF secondary market is negatively associated with quoted spreads, suggesting improved liquidity, and is not associated with increased volatility or reduced price efficiency. This is consistent with the information linkage channel. These results demonstrate that the ETF primary and secondary markets have distinct and often opposing effects on the market quality of the constituent stocks.

The French data do not allow us to say anything causal about these associations. Therefore, we take the unusual step of considering a different market setting in which we can implement an identification strategy. To establish causality, we leverage a quasi-natural experiment in the Japanese market, where the Bank of Japan's (BoJ) quantitative easing program involved large-scale purchases of equity index-backed ETFs. The experiment provides plausibly exogenous cross-sectional heterogeneity in the BoJ's demand for stocks due to the fact that one of the ETF products being purchased by the BoJ, replicates the Nikkei, a price-weighted index. Therefore, BoJ purchases imply the over-weighted activity of APs on some Nikkei stocks compared to their market capitalization.

Adopting the identification strategy of Barbon and Gianinazzi (2019) and Charoenwong et al. (2021), we extend the analysis by incorporating a counterfactual test. Consistent with the arbitrage channel, we show that on days with BoJ purchases, the implied arbitrage activity of the APs in the primary market increases in volatility and decreases in liquidity in overweighted constituent stocks. However, these effects are localized in time and concentrated in certain stocks. Over a longer horizon, the improved liquidity in ETF secondary markets, largely facilitated by HFTs, appears to mitigate the detrimental effects of primary market arbitrage.

Overall, our paper provides novel insights into the distinct roles of APs and HFTs in the ETF markets and we examine their respective impacts on market quality. By disentangling the effects of primary and secondary market activities, our work offers a more nuanced understanding of ETF market dynamics and can help to inform ongoing debate about their broader financial market implications.

Our findings contribute to several strands of the literature. First, we contribute to resolving the debate on the effects of ETFs on their constituent stocks. Previous empirical studies highlight both detrimental effects, such as increased volatility and reduced liquidity (Ben-David et al., 2018; Brogaard et al., 2024), and beneficial effects, such as improved liquidity and price efficiency and information diffusion (Saglam et al., 2019; Box et al., 2021). Our results reconcile these conflicting views by demonstrating that the ETF primary market can have detrimental effects under extreme conditions, while the secondary market, dominated by HFTs, provides significant liquidity benefits. We also show that the ETF primary market is linked to an index arbitrage strategy, as predicted in Kumar and Seppi (1994). As such, ETF arbitrage contributes to price efficiency by enforcing the law of one price (Pasquariello, 2017). However, we find that the benefits of arbitraging small ETF mispricing are insufficient to offset the costs of arbitrage, which connects our study to the literature on limits to arbitrage (Shleifer and Vishny, 1997; Gromb and Vayanos, 2010). These limits, in turn, reduce the frequency of ETF primary market activities, limiting their importance relative to the substantial trading activities of HFTs in the ETF secondary market.

<sup>3</sup> We acknowledge that there are firms, such as Citadel, that act as both APs and HFTs, but we focus on the distinct impact of these two different types of business that they operate.

Second, our paper advances the understanding of liquidity provision in ETFs. While prior research has focused on APs as the primary drivers of ETF liquidity (Malamud, 2015; Foucault et al., 2017), we show that HFTs dominate ETF secondary markets and play a critical role in reducing quoted spreads and enhancing liquidity. These findings extend the literature on liquidity provision by highlighting the unique contributions HFTs make to the ETF ecosystem.

Third, we contribute to the growing body of work on high-frequency trading. While HFTs have been shown to improve market quality in equities (Biais et al., 2015; Foucault et al., 2016), their role in ETFs has received limited attention. Our study documents that despite not engaging in ETF arbitrage, HFTs contribute to improved liquidity in ETF constituent stocks, supporting the predictions of the information linkage channel (Fremault, 1991; Madhavan, 1996). These findings highlight the broader market benefits of HFT-driven liquidity in ETFs.

The remainder of the paper is structured as follows. In Section 2, we present details of the ETF primary market and HFT secondary market activities and examine their determinants. In Section 3, we analyze how these activities affect the market quality of constituent stocks. Section 4 provides estimates of the causal impact of primary market activity. In Section 5, we investigate which effect dominates using a time trend analysis. Concluding remarks are in Section 6.

## 2. ETF primary market and HFT secondary market activities

Our first analysis focuses on the French market where we are able to directly identify both primary market activity in ETFs and the trading activity of HFTs in the ETF secondary market. We begin by documenting the level of primary market activity and HFT activity in the ETF secondary market. We then examine the determinants of these two types of activity.

We use the Autorité des Marchés Financiers (AMF) Euronext Paris data provided by Eurofidai-Bedofih.<sup>4</sup> The data provides microsecond time-stamped orders and trades for stocks trading on Euronext Paris. We obtain the database for the calendar year 2015 for all ETFs and stocks trading on Euronext. The database provides an HFT flag to identify when an HFT is on one or both sides of the trade. Further details about the data set and the flag are provided in Section A1 of the Internet Appendix.

### 2.1. Sample selection, variable definitions, and descriptive statistics

We identify all European ETFs trading on Euronext Paris in 2015. Many ETFs on Euronext have low AUM and seldom trade, so we use two filters to obtain a sample of large liquid ETFs. First, we require that the market capitalization at the end of 2015 be greater than €50 million euros. Second, we exclude ETFs with an average daily trading volume below €1 million. This provides a sample of 34 ETFs that represent 93% of the trading volume of European ETFs listed on Euronext Paris. The market value and volume traded for the ETFs in our sample are reported in Table IA1 of the Internet Appendix. These statistics show that the French ETF market is heavily concentrated, with the top 10 ETFs accounting for 81% of AUM and 44% of trading activity. Market value is only weakly correlated with trading activity.<sup>5</sup>

Using data from Bloomberg and Datastream, we proxy ETF primary market activities using the change in the number of ETF units outstanding each day expressed as a percentage. Indeed, creating (redeeming) ETF units requires the delivery (receipt) of constituent stocks and therefore cross-market arbitrage. For the ETF  $e$  and the day  $t$ , we compute ETF primary market activities as:

$$PMActivity_{e,t} = \frac{\#UnitsETF_{e,t} - \#UnitsETF_{e,t-1}}{\#UnitsETF_{e,t-1}}, \quad (1)$$

where  $\#UnitsETF$  denotes the number of ETF units outstanding at the end of the day. We use  $PMActivity$  to measure the signed magnitude of primary activity and its absolute value  $|PMActivity|$  for its overall magnitude.

To account for the fact that ETF primary market activities do not occur every day, we also compute ETF primary market activities using an ETF-day indicator variable. The indicator variable measures the frequency rather than the magnitude of the primary market. This indicator variable is set to one for the days when the ETF experiences a creation or redemption, and to zero otherwise. Formally, we have:

$$\mathbb{1}_{PMActivity_{e,t}} = \begin{cases} 1, & \text{if } |PMActivity_{e,t}| > 0 \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

Our main proxy for HFT activity, constructed using the Bedofih data, is based on the number of ETF units traded. For each ETF  $e$  and day  $t$ , we compute HFT activity as the total number of units traded for which at least one of the counterparties is flagged as an HFT. Unlike APs in the primary market, HFTs are active every day. Hence, instead of looking at the frequency of HFTs' trades, we focus on the intensity of their activities. Our main proxy of HFTs' activities is the logarithm of the absolute number of units traded by HFTs:

$$HFTActivity_{e,t} = \log(1 + \#UnitsTradedHFT_{e,t}), \quad (3)$$

where  $\#UnitsTradedHFT$  denotes the number of units traded by HFTs.

<sup>4</sup> These data have been used in a range of studies including (Colliard and Hoffmann, 2017; Anagnostidis et al., 2020; Donatoni et al., 2022; Métais and Roger, 2022).

<sup>5</sup> ETF concentration is not specific to France. As of May 2023, the ten most traded ETFs in the U.S. accounted for 57% of total trading activity (Authors' calculations using ETFdb.com data.)

**Table 1**

Variable definitions.

This table reports definitions for the variables used throughout the empirical analysis. Variables are constructed at the ETF-day or stock-day frequency, unless otherwise noted. Variables are winsorized at the 1% and 99% levels.

Variable	Definition
<i>PMA</i> Activity	Change in the number of ETF units, expressed as a percentage.
$\mathbb{1}_{PM}$ Activity	Indicator variable equal to one if an ETF experiences primary market activity on a given day, zero otherwise.
<i>ETFMispricing</i>	Difference between the ETF closing price and its NAV, expressed as a percentage of the NAV and standardized at the ETF level.
$\mathbb{1}_{HighMispricing,Median}$	Indicator variable equal to one if $ ETFMispricing $ is above the ETF median, zero otherwise.
$\mathbb{1}_{HighMispricing,TopTercile}$	Indicator variable equal to one if $ ETFMispricing $ is in the top tercile of the ETF distribution, zero otherwise.
<i>MarketValue</i>	Log of the market capitalization.
<i>HFT%</i> <sub>alldays</sub>	HFT trading volume as a percentage of ETF trading volume on all trading days.
<i>HFT%</i> <sub>PMdays</sub>	HFT trading volume as a percentage of ETF trading volume on days with primary market activity.
<i>HFT%</i> <sub>non-PMdays</sub>	HFT trading volume as a percentage of ETF trading volume on days without primary market activity.
<i>HFT</i> Activity	Logarithm of one plus the number of ETF units traded by HFTs.
<i>SignedHFT</i> Activity	Net trading activity of HFTs, signed according to net buying or selling.
<i>HFT</i> StockActivity	Logarithm of one plus the number of constituent shares traded by HFTs.
<i>HFT</i> LiquidityProvision	Log of one plus the number of ETF units traded by HFTs when acting as liquidity providers.
<i>HFT</i> LiquidityTaking	Log of one plus the number of ETF units traded by HFTs when acting as liquidity takers.
<i>ETF</i> Own	ETF ownership of a stock, expressed as a percentage of its market capitalization.
<i>ETF</i> Own $\times$ $ HFT$ Activity	Interaction of ETF ownership with absolute primary market activity.
<i>ETF</i> Own $\times$ <i>HFT</i> Activity	Interaction of ETF ownership with HFT trading activity.
<i>Volatility</i>	Annualized standard deviation of 5 min returns, in percentage.
<i>QuotedSpread</i>	Average quoted spread based on 5 min intervals, in bps.
<i>Depth</i>	Log of the average best-limit order book depth based on 5 min intervals.
<i>Resiliency</i>	Depth resiliency, based on Kempf et al.'s (2009) dynamic model.
<i>Inefficiency</i>	Absolute value of first-order 5 min returns autocorrelations.
<i>ETF</i> QuotedSpread	Average quoted spread of the ETF based on 5 min intervals, in bps.
<i>QuotedSpreadDiff</i>	Difference between the weighted average of the quoted spread of the constituent stocks (based on their index weights) and the ETF's quoted spread.
<i>PM</i> StockActivity	Value of daily stock purchases due to ETF primary market activity, divided by stock market capitalization, in percentage.
<i>LastPM</i> StockActivity	Most recent non-zero value of <i>PM</i> StockActivity on non-BoJ purchase days.
<i>Return</i>	Daily close-to-close return adjusted for dividends, in percentage.

**Table 2**

Descriptive statistics — Euronext Paris ETFs.

This table reports summary statistics for ETF-day level measures of HFT and primary market activity. Statistics are shown for all days, days with primary market activity, and days without. Variable definitions are provided in Table 1. The sample includes 34 large ETFs listed on Euronext Paris over the period January 2015 to December 2015.

Statistic	N	Mean	St. Dev.	Min	Pct(25)	Median	Pct(75)	Max
<i>PMA</i> Activity	6,645	0.06	2.23	-10.53	0.00	0.00	0.00	9.98
$\mathbb{1}_{PM}$ Activity	6,645	0.32	0.47	0.00	0.00	0.00	1.00	1.00
<i>ETFMispricing</i>	6,645	0.005	0.37	-1.55	-0.06	0.01	0.10	1.37
$ ETFMispricing $	6,645	0.21	0.31	0.00	0.03	0.09	0.22	1.55
<i>MarketValue</i>	6,645	812	1,450	24	142	260	918	8,034
<i>HFT%</i> <sub>alldays</sub>	6,645	73.81	30.16	0.00	55.75	87.80	97.82	100.00
<i>HFT%</i> <sub>PMdays</sub>	2,322	78.36	27.92	0.00	67.75	92.02	97.95	100.00
<i>HFT%</i> <sub>non-PMdays</sub>	4,323	71.37	31.02	0.00	50.58	84.38	97.71	100.00
<i>HFT</i> Activity	6,645	9.98	2.97	0.00	8.36	10.16	12.04	15.72

ETF mispricing, constructed using data from Bloomberg, is the daily ETF premium or discount to NAV. Following Dannhauser (2017) and Pan and Zeng (2019), we use the ETF closing price and its NAV. The daily ETF mispricing is defined as follows<sup>6</sup>:

$$ETFMispricing_{e,t} = \frac{Price_{e,t} - NAV_{e,t}}{NAV_{e,t}}. \quad (4)$$

In some specifications, we are interested in the magnitude of the price deviation irrespective of its direction (discount and premium). In these instances, we use the absolute value of *ETF Mispricing*. All variables are winsorized at the 1% and 99% levels to limit the impact of outliers.

In Table 1, we summarize the definitions of all variables used in the empirical analysis, and in Table 2, we present summary statistics for the Euronext Paris ETF sample. ETF primary market activities are infrequent, with a median of 0 for  $\mathbb{1}_{PM}Activity_{e,t}$

<sup>6</sup> To mitigate potential biases from different ETF NAV conventions, such as fixed-income ETFs using bid prices while most equity ETFs use mid prices, we also standardize mispricing at the ETF level by subtracting each ETF's mean mispricing.

confirming that for more than half of the days, there is no creation/redemption of ETF units. The average of 0.32 indicates that ETF creations/redemptions occur, on average, only on 32% of the days. To illustrate the low frequency of ETF primary market activities, in Fig. 1 Panel A, we plot the percentage of ETFs that experience primary market activities each day.

The ETF primary market is small compared to the secondary market. Across our sample period, the total amount of ETF primary market activities represents only 13% of the ETF secondary market trading volume.<sup>7</sup> The low frequency of primary market activity is common across other markets but it is not well documented or considered in the papers examining the impact the primary market has on constituent stocks.<sup>8</sup>

HFTs by contrast account for a large fraction of total trading volume. On average, HFTs account for 73.8% of the daily trading volume for ETFs on Euronext. This is all the more striking given that the HFTs' flag excludes investment banks (flagged as MIX) from the HFT proxy despite similarities.<sup>9</sup> In Panel B of Fig. 1, we plot the HFTs' volume as a percentage of the total trading volume. To the best of our knowledge, we are the first to document the dominance of HFTs' trading in ETFs. HFTs' activities are much higher in ETFs than in the underlying CAC stocks where they account for 50% of the daily trading volume.<sup>10</sup>

Table 2 also reports that ETF mispricing is on average 0.005%, but highly variable. The standard deviation is 0.37% with a maximum discount of -1.55%. In Panels C and D of Fig. 1, we plot the market volatility and ETF mispricing, respectively. Visually, high market volatility in July and September seems related to high ETF mispricing (high discount or high premium).

## 2.2. Determinants of primary market activities

We next investigate the determinants of ETF primary market activities by estimating the following regression for each ETF  $e$  and day  $t$  in our sample of ETFs trading on Euronext:

$$\text{PrimaryActivity}_{e,t} = \alpha_e + \beta_1 \cdot \text{ETFMispricing}_{e,t} + \beta_2 \cdot \text{HFTActivity}_{e,t} + \phi' \cdot X_t + \eta_t + \varepsilon_{e,t}. \quad (5)$$

We use three alternative proxies for the *Primary Activity* described above:  $|\text{PMActivity}|$ ,  $\mathbb{1}_{\text{PM Activity}}$ , and  $\text{PMActivity}$ .

To account for time-invariant ETF heterogeneity, such as basket size, we incorporate ETF fixed effects, denoted as  $\alpha_e$ . To mitigate the effects of common shocks and trends, we introduce day fixed effects, represented by  $\eta_t$ . Additionally, to control for size and price, which the literature shows can impact market quality, we include the lagged ETF market capitalization and the inverse of the lagged ETF price as controls, denoted as  $X_t$ . Since day fixed effects can mask the results of interest, for some specifications, we replace the day fixed effects with month-year fixed effects. We also include the European equivalent of the macro controls used in Bessembinder et al. (2006): daily variation of market volatility (proxied by the European VIX: VSTOXXI), MSCI Europe equity index, European Ted Spread, and iBoxx corporate bond index.<sup>11</sup> Additionally, we include single lead and lag observations of the *Primary Activity* proxy to consider the possibility of potential front-running, anticipatory behaviors or delayed reactions.

The results in Table 3 confirm that there are more ETF primary market activities when an ETF is mispriced. Panel A reports the results using  $|\text{PMActivity}|$ . We find that ETF mispricing is positively related to the absolute value of daily ETF primary market activities. In the main specification presented in column (1), the coefficient 0.079 implies that a one percentage point increase in  $|\text{PMActivity}|$  is associated with an increase in absolute ETF mispricing of approximately 2.45 bps ( $0.079 \times 31$  bps, where 31 bps is the standard deviation of  $|\text{ETFMispricing}|$ ). Compared to the average  $|\text{ETFMispricing}|$  of 21 bps that is an 11.7% increase. In columns (1) to (3), the results remain qualitatively similar with and without the inclusion of time fixed effects, macro controls, and a single lead and lag of  $|\text{PMActivity}|$ , respectively.

To test for non-linearities, we introduce  $\mathbb{1}_{\text{HighMispricing}_{\text{Median}}}$  and  $\mathbb{1}_{\text{HighMispricing}_{\text{TopTercile}}}$ , which are indicator variables equal to one if the absolute value of the ETF-day mispricing is above the ETF median or top tercile, respectively, and zero otherwise. In columns (4) and (5) of Panel A in Table 3, the results show that ETF primary market arbitrage is larger with larger ETF mispricing. The coefficient of 0.143 indicates that ETFs with above-median mispricing tend to have, on average, 14.3 bps (0.143 percentage points) higher primary market activity. A larger mispricing (top 33%) is associated with a 12.1 basis point (0.121 percentage points) increase in primary market activity. These results relate to limits to arbitrage (Shleifer and Vishny, 1997; Gromb and Vayanos, 2010). Small mispricing does not provide sufficient benefits to arbitrageurs, given the trading costs implied by the primary market arbitrage.

Next, we examine the factors influencing the frequency of ETF primary market activities. Panel B of Table 3 shows that  $\mathbb{1}_{\text{PM Activity}}$  is also positively associated with absolute mispricing. The estimate of 0.025 in column (1) indicates that a one percentage point increase in  $|\text{ETFMispricing}|$  is associated with a 2.5 percentage point increase in the probability of primary market activity. This confirms that APs trade the ETF when its price is misaligned relative to the price of the constituent stocks. These results again confirm that there is more primary market activity when mispricing is more extreme.

<sup>7</sup> This is similar to the U.S. market, where for the first quarter of 2013, the gross primary market activity is only 11% of the secondary market activity of ETFs, see [iSharesGlobalETFfacts, Apr13, 2023](#).

<sup>8</sup> Fulkerson et al. (2022) show that the primary market is also infrequent in the case of U.S. ETFs, while Lettau and Madhavan (2018) report that the primary market is only a fraction of the secondary market for U.S. ETFs.

<sup>9</sup> Indeed, some desks of investment banks use their own capital to run high-frequency trading strategies and provide market access to high frequency trading firms.

<sup>10</sup> See Table IA2 of the Internet Appendix.

<sup>11</sup> Table IA3 of the Internet Appendix presents a correlation matrix of the macro controls, verifying that there is no multi-collinearity between the variables.

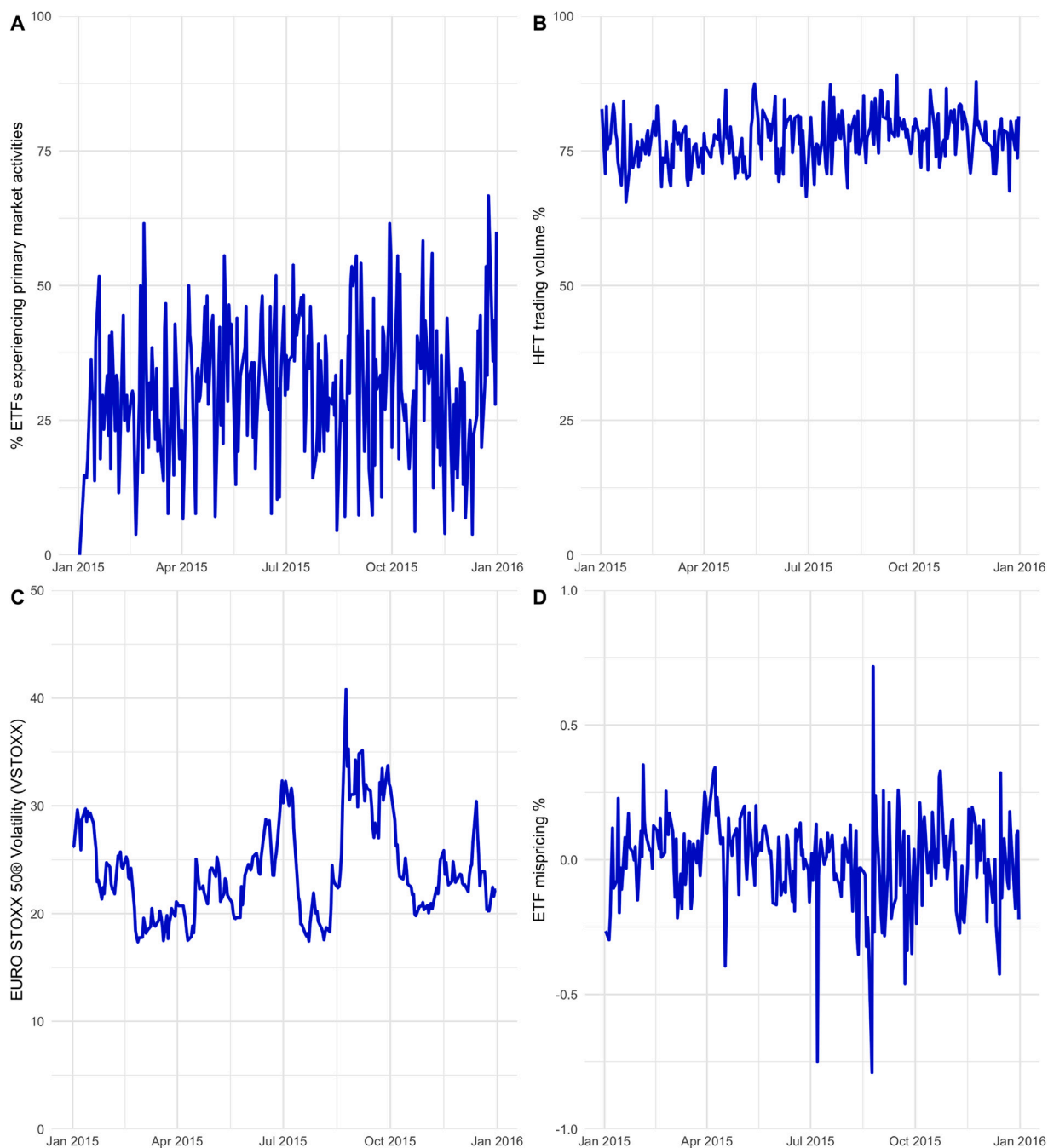


Fig. 1. Euronext Paris ETFs, volatility, and mispricing.

Panel A reports the percentage of ETFs experiencing primary market activities each day. Panel B reports HFTs' trading volume as a percentage of the daily number of ETF units traded where at least one of the counterparties is flagged as an HFT. Panel C reports the Eurostoxx 50 volatility index (VSTOXX). Panel D reports ETF mispricing as the difference between the ETF closing price and its NAV, expressed as a percentage of the NAV. The sample includes 34 large ETFs listed on Euronext Paris over the period from January 2015 to December 2015.

Panel C reports the results using the signed value of  $PMActivity$ .<sup>12</sup> This specification allows a distinction between ETF creations and ETF redemptions and therefore informs on the arbitrage direction. We again find that ETF mispricing is positively related to daily

<sup>12</sup> In this panel, the dependent variable increases with ETF unit creations and decreases with ETF unit redemptions. As a result, the indicator variables  $\mathbb{1}_{HighMispricing\_Median}$  and  $\mathbb{1}_{HighMispricing\_TopTercile}$  lose interpretability. Therefore, we do not include the alternative specification in columns (4) and (5).

**Table 3**

The determinants of primary market activity - Euronext Paris ETFs.

This table reports ETF-day level regressions of primary market activity on ETF mispricing and controls. Panel A reports results for  $|PMActivity|$ , the absolute change in ETF units (in %). Panel B reports results for the indicator  $\mathbb{1}_{PM\ Activity}$ , and Panel C reports results for  $PMActivity$ , the signed percentage change in ETF units. Macro controls include the Eurostoxx 50 volatility index, MSCI Europe equity index, European Ted Spread, and iBoxx corporate bond index. ETF controls include market value, inverse lagged price, return, log dollar volume, and bid-ask spread. Lead and lag refer to inclusion of the lead and lagged dependent variable. All specifications include ETF fixed effects; other controls vary by column. Variable definitions are provided in Table 1. Double-clustered standard errors are in parentheses. \*\*\*, \*\*, and \* indicate significance at the 1%, 5%, and 10% levels, respectively. The sample includes 34 large ETFs listed on Euronext Paris from January to December 2015.

Panel A: Magnitude					
	$ PMActivity $				
$ ETFMispricing $	0.079** (0.036)	0.101*** (0.034)	0.078** (0.035)		
$\mathbb{1}_{HighMispricing\_Median}$				0.143*** (0.035)	
$\mathbb{1}_{HighMispricing\_TopTercile}$					0.121*** (0.037)
$HFTActivity$	0.003 (0.009)	0.005 (0.006)	0.005 (0.008)	0.004 (0.009)	0.003 (0.009)
ETF fixed effects	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	No	Yes	Yes	Yes
Macro controls	NA	Yes	NA	NA	NA
Lead and lag	No	No	Yes	No	No
Observations	6,645	6,645	6,613	6,645	6,645
R <sup>2</sup>	0.158	0.108	0.159	0.158	0.158
Panel B: Frequency					
	$\mathbb{1}_{PM\ Activity}$				
$ ETFMispricing $	0.025** (0.010)	0.030*** (0.010)	0.024** (0.010)		
$\mathbb{1}_{HighMispricing\_Median}$				0.049*** (0.011)	
$\mathbb{1}_{HighMispricing\_TopTercile}$					0.041** (0.016)
$HFTActivity$	0.003 (0.003)	0.004 (0.003)	0.002 (0.002)	0.004 (0.003)	0.004 (0.003)
ETF fixed effects	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	No	Yes	No	No
Macro controls	NA	Yes	NA	NA	NA
Lead and lag	No	No	Yes	No	No
Observations	6,645	6,645	6,613	6,645	6,645
R <sup>2</sup>	0.210	0.137	0.217	0.137	0.137
Panel C: Signed					
	$PMActivity$				
$ETFMispricing$	0.245** (0.091)	0.255*** (0.087)	0.212*** (0.077)	0.260*** (0.090)	0.281*** (0.088)
$SignedHFTActivity$	-0.094 (0.071)	-0.087 (0.071)	-0.096 (0.070)	-0.091 (0.069)	-0.089 (0.069)
ETF fixed effects	Yes	Yes	Yes	No	No
Time fixed effects	Yes	No	Yes	Yes	No
Macro controls	NA	Yes	NA	No	No
Lead and lag	No	No	Yes	No	No
Observations	6,645	6,645	6,613	6,645	6,645
R <sup>2</sup>	0.073	0.017	0.097	0.069	0.011

ETF primary market activities. This suggests that APs create ETF units when the ETF trades at a premium (positive mispricing). By supplying ETF units in response to high demand — reflected in higher prices and greater mispricing — they help correct mispricing. Similarly, the findings suggest that APs redeem ETF units when the ETF trades at a discount, effectively reducing supply when demand is low. This shows the role of APs in correcting ETF mispricing rather than exacerbating it. Therefore, APs enforce the law of one price. A one percentage point increase in the ETF premium is associated with a 21 to 28 bps increase in ETF primary market

**Table 4**

Activity of high-frequency traders and ETF mispricing - Euronext Paris ETFs.

This table reports ETF-day OLS regressions of HFT activity on ETF mispricing, primary market activity, and controls. Panel A reports results for  $|PM\ Activity|$ , while Panel B reports results for  $\mathbb{1}_{PM\ Activity}$ . All specifications include ETF fixed effects. Macro controls comprise the Eurostoxx 50 volatility index, MSCI Europe equity index, European Ted Spread, and iBoxx corporate bond index. ETF-level controls include market value, inverse lagged price, return, log dollar volume, and bid-ask spread. Lead and lag refer to inclusion of the lead and lagged dependent variable. Variable definitions are provided in Table 1. Double-clustered robust standard errors are presented in parentheses. \*\*\*, \*\*, and \* indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample includes 34 large ETFs listed on Euronext Paris over the period January 2015 to December 2015.

Panel A: Magnitude					
	<i>HFTActivity</i>				
$ ETFMispricing $	0.167 (0.184)	0.193 (0.194)	0.138 (0.177)		
$\mathbb{1}_{HighMispricing,Median}$				0.085 (0.062)	
$\mathbb{1}_{HighMispricing,TopTercile}$					0.101 (0.071)
$ PM\ Activity $	0.003 (0.008)	0.004 (0.006)	0.001 (0.008)	0.004 (0.006)	0.004 (0.006)
ETF fixed effects	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	No	Yes	Yes	Yes
Macro controls	NA	Yes	NA	NA	NA
Lead and lag	No	No	Yes	No	No
Observations	6,645	6,645	6,645	6,645	6,645
R <sup>2</sup>	0.828	0.820	0.831	0.820	0.820
Panel B: Frequency					
	<i>HFTActivity</i>				
$ ETFMispricing $	0.166 (0.183)	0.190 (0.193)	0.136 (0.177)		
$\mathbb{1}_{HighMispricing,Median}$				0.092 (0.065)	
$\mathbb{1}_{HighMispricing,TopTercile}$					0.093 (0.068)
$\mathbb{1}_{PM\ Activity}$	0.026 (0.026)	0.038 (0.028)	0.016 (0.025)	0.023 (0.026)	0.024 (0.026)
ETF fixed effects	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	No	Yes	Yes	Yes
Macro controls	NA	Yes	NA	NA	NA
Lead and lag	No	No	Yes	No	No
Observations	6,645	6,645	6,645	6,645	6,645
R <sup>2</sup>	0.828	0.820	0.831	0.828	0.828

activities depending on the specification. A back-of-the-envelope calculation reveals the economic magnitude of the effect of ETF mispricing on arbitrage. For an ETF with AUM of €3.8 billion, such as the Lyxor CAC 40 (the principal ETF tracking the French large-cap index at the time of the study), the estimates imply that a 1 percentage point increase in the ETF premium is associated with the creation of €8.06 to €10.68 million worth of ETF units depending on the specification. Since the primary activity is linked to the order flow of the arbitrageurs, in Panel C, we compare it to the signed number of units traded by HFTs. Again, we find that HFT activity is unrelated to the primary market activity.

Across the three panels, the inclusion of ETF and time fixed effects, macro controls, and one lead and lagged observation of the dependent variables do not qualitatively change the results in any of the specifications.

Taken together, our results support the view that ETF mispricing is associated with ETF primary market activities and therefore to arbitrage with the constituent securities. This confirms Pan and Zeng (2019) findings on U.S. corporate bond ETFs for the European ETF market including equity, bond, commodities, leverage, and inverse products.

We undertake a number of robustness tests related to the measurement of HFT activity, the inclusion of additional lagged observations of the primary activity and mispricing, and positive/negative mispricing. We also consider alternative logit specifications for the frequency-based measure of primary market activity. We also consider a sample of only equity ETFs.

The results in Table 3 consistently show that neither the unsigned/signed magnitude nor the frequency of ETF primary market transactions are statistically associated with HFT activity. We test the robustness of these results to alternative proxies for HFT activity: (i) the number of ETF units traded by HFTs expressed as a percentage of the number of units traded, (ii) the log of the number of units traded by HFTs excluding observations with zero units traded by HFTs, and (iii) the log of the number of ETF units traded by HFTs where HFTs acted as liquidity providers. Section A2 of the Internet Appendix provides an explanation for

these proxies and their definitions. Table IA4 of the Internet Appendix shows that all of these proxies provide results that are qualitatively consistent with those of our main specification in Table 3.

The addition of five lags of primary activity and mispricing provides similar results (see Table IA5 of the Internet Appendix). These results indicate that the French market differs from the U.S. market. Evans et al. (2022) document a five-day delay between ETF mispricing and ETF primary market activity (indicating that in the U.S., APs may exploit the option to fail to deliver).

Consistent with limits to arbitrage, Table IA6 in the Internet Appendix shows that below-average mispricing (negative mispricing) significantly reduces the likelihood of primary market activity compared to above-average mispricing (positive mispricing). In this specification, we include an indicator variable equal to one when daily ETF mispricing is below its time-series average for that ETF.<sup>13</sup> The significant negative estimate suggests that arbitrage is less likely when ETFs trade at a discount, likely due to higher transaction costs and market frictions during sell-offs (Shleifer and Vishny, 1997). However, in Panel A, where we examine the absolute size of primary market activity, the coefficient for the negative mispricing indicator variable is not significant. This mixed evidence suggests that while negative mispricing discourages arbitrage activity, it does not directly affect the volume of primary market activity when arbitrage occurs.

Given that ETF primary market activities occur on only 32% of days, the data contain a substantial concentration of zero observations. To account for this, we re-estimate the specification using logit and conditional logit and present the results in Table IA7 of the Internet Appendix. These tests confirm our earlier OLS results. A one percentage point increase in  $|ETF\text{Mispricing}|$  increases the odds of ETF primary market activities by between 10% and 19% depending on the specification. This corresponds to an increase in  $\mathbb{1}_{PM\text{ Activity}}$  of 2.2 to 3.9 percentage points, compared to the 2.5 percentage points increase estimated using OLS.

Finally, we also reproduce the analysis using only equity ETFs and present the results in Table IA8 of the Internet Appendix. The results remain qualitatively unchanged, confirming that our main results are not driven by fixed income ETFs (Pan and Zeng, 2019).

### 2.3. Determinants of HFT activities

Having established that mispricing is associated with primary market activity, we next seek to understand if HFTs exhibit similar or different trading patterns to that of APs. We run the following panel regression:

$$HFT\text{activity}_{e,t} = \alpha_e + \beta_1 \cdot \text{PrimaryActivity}_{e,t} + \beta_2 \cdot |ETF\text{Mispricing}_{e,t}| + \varphi' \cdot X_t + \eta_t + \varepsilon_{e,t}. \quad (6)$$

The main variables are as defined for Eq. (5). We consider the same fixed effects, control variables, and a lead/lag of the dependent variable in alternative specifications.

The results in Table 4 confirm that the activity of HFTs is (i) not related to ETF primary market activities and (ii) not related to the price divergence of the ETF with regard to the constituent stocks. The coefficient in the first row shows no statistically significant association between the activity of HFTs and ETF mispricing (in absolute terms). In addition, the statistically insignificant estimate in the second row of Table 4 further confirms that the HFT activity is not linked to the ETF primary market. If HFTs are participating in the primary market, the coefficient should be negative and significant. The results are confirmed for both the magnitude (Panel A) and frequency (Panel B) of primary market activity. This confirms that HFTs are not APs or that their activity as APs are not substantial. HFTs behave as if they are not repricing the ETFs, nor participating in basis arbitrage at least at the daily frequency.

Our analysis in Table 4 is at a daily rather than intraday frequency, because ETF primary market activities can only be observed on a daily basis. HFTs are likely to react to short-term information that may not be visible at this frequency (Budish et al., 2015). However, despite the possibility that we may not be capturing all of the determinants of HFTs' activities due to the frequency of our data, our findings have important implications. Specifically, our results suggest that the HFTs' activities that dominate trading on the secondary market for ETFs are not significantly related to the primary market for ETFs. Therefore, if high frequency trading firms can participate in the ETF primary market, our findings suggest that this is not the predominant strategy used by high frequency trading firms. Similarly, since HFTs' activities are not related to the daily ETF mispricing it does not appear that index arbitrage is a primary focus of HFTs.

The role and impact of HFTs vary depending on whether they provide or take liquidity from the limit order book (Brogaard et al., 2014, 2019). Therefore, we also consider whether the activities of liquidity providers have a differential impact on ETF mispricing than liquidity takers. However, we find that there is no difference (see Table IA9 of the Internet Appendix). This finding confirms that HFTs are neither significantly active in the ETF primary market nor significantly contribute to the correction of daily ETF mispricing.

Overall, these results support the notion that the ETF primary market and the secondary market are distinct. APs arbitrage the ETF and the constituent stocks for the primary market, while HFTs' activities appear unrelated to the primary market and to the ETF mispricing. The ETF primary market helps align prices between the ETF and its basket of constituents with the law of one price, but this alignment occurs only when the mispricing is sufficiently large to cover transaction costs and other arbitrage expenses.

<sup>13</sup> We use each ETF's own mispricing average to account for the NAV convention, as noted in Footnote 6.

### 3. The influence of ETF activities on the market quality of constituent stocks

In this section, we continue to focus on the French market to study the association of ETF primary market activities and HFTs' ETF activity on the market quality of the constituent stocks using panel regressions.

We select the Lyxor CAC ETF and its constituent stocks, as it is the largest ETF on Euronext that holds French stocks. Moreover, the Lyxor CAC ETF is a physical ETF that reproduces its index (the CAC 40 Index) without sampling.<sup>14</sup> For each stock included in the ETF, we study the association between both the primary market activity and HFTs secondary market activity and the market quality of the Lyxor CAC ETF.

To create the market quality variables, we follow Colliard and Hoffmann (2017) and use five-minute intervals. We obtain the data from LSEG DataScope. We compute returns using the last transaction price of the interval. Daily volatility is defined as the standard deviation of the five-minute returns. We also compute a measure of price efficiency as the absolute value of first-order return autocorrelations. To measure liquidity, we compute the relative quoted spread, the depth, and the depth resiliency based on Kempf et al. (2009) dynamic model.

We continue to measure primary market activity and HFTs' ETF activity using  $|PMActivity|$  and  $HFTActivity$ .<sup>15</sup> To obtain variables at the constituent stock level  $i$ , we interact both variables with the stock's CAC ownership (i.e., its weight in the ETF).

$$ETFown_{i,t} \times |PMActivity_t| \quad (7)$$

$$ETFown_{i,t} \times HFTActivity_t, \quad (8)$$

where the ETF ownership per stock  $i$  is defined as:

$$ETFown_{i,t} = \frac{ETFholdings_{CAC,i,t}}{MarketValue_{i,t}}, \quad (9)$$

with ETF holdings computed as the product of the stock's weight in the Lyxor CAC ETF and the Lyxor CAC ETF market capitalization. By linking ETF primary market activities with the proportional weight of the stock in the ETF, the variable  $ETFown_{i,t} \times |PMActivity_t|$  measures the effect of ETF primary market activities on constituent stocks. Similarly,  $HFTActivity$  interacted with the stock's CAC ownership captures the effect of HFTs trading the ETF on constituent stocks. We investigate the association of primary and secondary market activity with the market quality of the constituent securities.

For each stock included in the Lyxor CAC ETF, we run the following OLS regressions:

$$Y_{i,t} = \alpha_i + \beta_1.(ETFown_{i,t} \times |PMActivity_t|) + \beta_2.(ETFown_{i,t} \times HFTActivity_t) + \eta_t + \varphi' . X_{i,t} + \varepsilon_{i,t}, \quad (10)$$

where  $Y_{i,t}$  denotes our market quality measures: the volatility, the liquidity, and the price efficiency. To account for time-invariant stock heterogeneity, the panel regression includes stock fixed effects,  $\alpha_i$ . The covariates,  $X_{i,t}$ , control for time-varying stock characteristics. We include the terms of the interaction separately:  $ETFown$ ,  $|PMActivity|$ , and  $HFTActivity$ . To disentangle the impact of HFTs directly trading the stocks from the impact of HFTs trading the ETF, the controls also include the logarithm of the number of shares traded by HFTs in the constituent stock ( $HFTStockActivity$ ). Consistent with the previous specifications, the covariates also include the lagged log market capitalization and the lagged price inverse. To control for common shocks and trends, we include day fixed effects. Table IA2 provides the descriptive statistics for the stocks included in the Lyxor CAC ETF.

Table 5 presents the results. In column (1), we examine the volatility of the stocks. Consistent with the arbitrage channel hypothesis, there is a significant positive association between ETF primary market activities and the volatility of the stocks. This indicates that when APs buy (sell) the stocks to create (redeem) the ETF, the constituent stocks tend to be more volatile. Conversely, the coefficient is negative and insignificant for the effect of HFTs' ETF activity on the volatility of the constituent stocks. Hence, the interaction effect of a stock being held by an ETF with high levels of HFT activity is not associated with stock volatility. The two types of ETF activity appear to affect the volatility of the constituent stocks differently.

Column (2) of Table 5 presents the results for the quoted spread of the stocks. The ETF primary market activities have a positive and significant effect on the quoted spread of the constituent stocks, implying a detrimental effect of the primary market on the liquidity of the stocks. The coefficient for *Depth* is not statistically different from zero, but the significant negative coefficient on the depth resiliency also indicates that ETF primary market activities are associated with reduced liquidity.

In stark contrast with the results for ETF primary market activities, there is a negative relation between HFT activity and the quoted spread for the constituent stocks. The relation is significant at the 1% level. These findings are consistent with the information leakage channel. HFTs appear to impact indirectly the constituent stocks, without directly trading them. The results show that HFTs' ETF activity is associated with an improvement in the liquidity of the constituent stocks. This finding supports Fremault's (1991) theoretical predictions that basket securities can make the market more complete by facilitating risk reallocation.

Finally, column (5) of Table 5 reports the results of a specification with the absolute return autocorrelation as the dependent variable. Consistent with the arbitrage channel hypothesis, we find that the ETF primary market activities are linked to an increase in

<sup>14</sup> Since the ETF replicates its index via physical replication without sampling, it follows that all the constituent stocks have to be traded in a creation/redemption. This framework favors the identification of the effect of arbitrage on constituent stocks.

<sup>15</sup> In Table IA10 of the Internet Appendix, we show that the results are corroborated when using an indicator variable  $\mathbb{1}_{PM\ Activity}$  that equals to one when the Lyxor CAC ETF undergoes creation or redemption amounting to over 3 million euros on a given day.

**Table 5**

PM and HFT ETF activities and the market quality of CAC stocks.

This table reports stock-day OLS regressions of market quality measures on ETF primary market activity and HFTs' ETF trading activity. Stock-level controls include lagged market value, inverse lagged price, and HFT trading activity in constituent shares. Variable definitions are provided in Table 1. Stock and day fixed effects are included. Double-clustered robust standard errors are presented in parentheses. \*\*\*, \*\*, and \* indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample includes the 38 stocks constantly composing the CAC Index over the period January 2015 and December 2015.

	<i>Volatility</i>	<i>QuotedSpread</i>	<i>Depth</i>	<i>Resiliency</i>	<i>Inefficiency</i>
<i>ETFown</i> × <i> PMAActivity </i>	0.029** (0.013)	0.004** (0.002)	0.017 (0.014)	-0.098** (0.043)	0.038** (0.018)
<i>ETFown</i> × <i>HFTActivity</i>	0.013 (0.010)	-0.006*** (0.001)	0.004 (0.008)	-0.068 (0.057)	-0.027* (0.014)
<i>ETFown</i>	-0.045 (0.136)	0.089*** (0.018)	-0.271** (0.114)	0.584 (0.840)	0.356* (0.192)
<i>HFTStockActivity</i>	0.052*** (0.004)	-0.007*** (0.0003)	-0.005 (0.003)	0.019*** (0.007)	-0.020*** (0.003)
Stock fixed effects	Yes	Yes	Yes	Yes	Yes
Day fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	9,455	9,455	9,455	9,455	9,455
R <sup>2</sup>	0.689	0.821	0.741	0.083	0.125

the absolute intraday return autocorrelation of the constituent stocks. Stocks held by an ETF being arbitrated are therefore negatively associated with price efficiency. By contrast, HFTs' ETF activity is positively related to the price efficiency of the constituent stocks at the 10% significance.

We again test whether the results differ when HFTs are providing versus taking liquidity. Table IA11 shows that the findings are qualitatively similar for both types of HFT activity.

Taken together, our results show that the ETF arbitrage for the primary market is associated with an increase in the constituent stocks' volatility, quoted spread, and autocorrelation, while HFTs' ETF activity is associated with a reduction in their quoted spread and a weakly significant improvement in price efficiency. The results are consistent with the hypothesis that the ETF primary market and the HFTs' activities on the ETF have distinct impacts on constituent stocks. Given that primary market activities occur infrequently and with minimal scale, it is likely that the benefits of ETF secondary market quality outweigh the negative impact of the primary market. However, due to concerns about simultaneity—such as the simultaneous occurrence of primary market arbitrage and mispricing shown in Table 3, and HFTs' activities and market volatility shown in Table 4—and other potential confounding factors, we cannot directly test for a net effect.

### 3.1. ETF-index liquidity differential

To further investigate how ETFs enhance the liquidity of their constituent stocks, we examine the liquidity differential between the CAC Index stocks and the Lyxor CAC ETF. Liquidity is proxied by the bid–ask spread. Spreads are computed daily for the CAC Index based on the constituent stocks and their respective weights. This is compared to the bid–ask spread for the Lyxor CAC ETF. The differential is calculated as the liquidity of the index minus that of the ETF. The summary statistics in Panel A of Table 6 show that the ETF is significantly more liquid than its index constituents. Specifically, the quoted spread differential of 2 bps indicates that the ETF consistently exhibits narrower bid–ask spreads, thereby enhancing its liquidity. The economic significance of this spread differential is notable as it exceeds the average bid–ask spread of the ETF itself, which is less than 2 bps.

To investigate the determinants of the ETF's superior liquidity, we regress the quoted spread differential on ETF HFT activity and primary market activity. The results in Panel B of Table 6 show that HFT activity is positively and significantly associated with the ETF's superior liquidity. In alternative specifications (columns (3) to (6)), we find that this positive association holds regardless of whether HFTs provide or take liquidity, suggesting that HFT activity, beyond liquidity provision alone, contributes to the ETF's superior liquidity. By contrast, neither the frequency nor the magnitude of primary market activity is significantly related to the liquidity differential between the ETF and its index. These findings support the view that HFTs enhance ETF liquidity relative to its index, while APs do not play a significant role in this context.

Next, we turn to the liquidity of the constituent securities. In Panel C, we show that ETF liquidity, measured by *ETFQuotedSpread*, is positively and significantly related to the liquidity of its constituent stocks, measured by *QuotedSpread*. In Panel D, we analyze the total effect of the interaction term, *HFTActivity* × *ETFQuotedSpread*, on the quoted spread of the constituent stocks at various levels of HFT activity. This analysis is based on an alternative specification of the regression in Panel C, which includes the interaction term. We do not report the individual regression coefficients, as the standalone coefficient of *HFTActivity* is not meaningful given that *HFTActivity* is never zero. The results indicate that increased *HFTActivity* increase the association between the liquidity of the ETF and the constituent securities. While we cannot claim causality, the results are consistent with the view that ETF liquidity, driven by increased HFT activity, “spills over” to the constituent securities via HFT activity in the ETF.

Combined, the results indicate that HFTs dominate trading in the secondary ETF market (Table 2) and contribute to making the Lyxor CAC ETF more liquid than its constituent stocks (Table 6). Additionally, HFT activity in the ETF is associated with improved liquidity in the constituent stocks.

**Table 6**

ETF and constituent stock liquidity.

This table provides the results of an examination of liquidity differences between the CAC Index constituents and the Lyxor CAC ETF. The weighted average intraday quoted spread of the constituent stocks, *QuotedSpread*, is computed daily based on individual components and their index weights. *QuotedSpreadDiff* represents the average intraday quoted spread differential between *QuotedSpread* and the ETF's *ETFQuotedSpread*, expressed in bps. Panel A reports summary statistics for the liquidity differential. Panel B presents regressions of *QuotedSpreadDiff* on ETF HFT and primary market activity proxies. Panel C reports regressions of *QuotedSpread* on *ETFQuotedSpread* and controls. Panel D presents the total effect of the interaction  $HFTActivity \times ETFQuotedSpread$  at various levels of HFT activity. Variable definitions are provided in Table 1. The sample covers the period from January 2015 to December 2015.

Panel A: Summary statistics						
Statistic	N	Mean	St. Dev.	Pctl(25)	Median	Pctl(75)
<i>QuotedSpread</i>	255	4.74	0.52	4.35	4.65	5.04
<i>ETFQuotedSpread</i>	255	2.58	0.44	2.26	2.48	2.80
<i>QuotedSpreadDiff</i>	255	2.17	0.46	1.88	2.12	2.44

Panel B: HFT ETF activity and liquidity differential						
	<i>QuotedSpreadDiff</i>					
<i>HFTActivity</i>	0.306*** (0.086)	0.331*** (0.086)				
<i>HFTActivityLiquidityProvision</i>			0.242*** (0.073)	0.259*** (0.073)		
<i>HFTActivityLiquidityTaking</i>					0.321*** (0.075)	0.331*** (0.075)
<i>PMActivity</i>	-0.061 (0.070)		-0.061 (0.069)		-0.054 (0.064)	
$\frac{1}{PM} Activity$		-0.285* (0.152)		-0.264* (0.152)		-0.246 (0.151)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Macro controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	255	255	255	255	255	255
R <sup>2</sup>	0.211	0.222	0.203	0.212	0.233	0.240

Panel C: Constituent stock liquidity and ETF liquidity						
	<i>QuotedSpread</i>					
<i>ETFQuotedSpread</i>	0.511*** (0.075)	0.448*** (0.067)	0.508*** (0.066)	0.408*** (0.069)	0.725*** (0.091)	0.723*** (0.098)
<i>lag1_QuotedSpread</i>		0.403*** (0.024)	0.389*** (0.023)	0.427*** (0.023)		
<i>lead1_QuotedSpread</i>		0.333*** (0.023)	0.372*** (0.021)	0.344*** (0.023)		
<i>lag1_ETFQuotedSpread</i>		-0.081* (0.045)	-0.098** (0.042)	-0.132** (0.054)		
<i>HFTStockActivity</i>	-0.004*** (0.001)	-0.003*** (0.001)	-0.002*** (0.001)	-0.003*** (0.001)	-0.003***	
Stock fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Month fixed effects	Yes	Yes	Yes	No	No	No
Macro controls	Yes	Yes	No	Yes	No	No
Observations	9,455	9,380	9,380	9,380	9,455	9,493
R <sup>2</sup>	0.780	0.880	0.873	0.876	0.727	0.670

Panel D: Total effect of $HFTActivity \times ETFQuotedSpread$ on <i>QuotedSpread</i>						
Statistic	Min	1st Qu.	Median	Mean	3rd Qu.	Max
<i>HFTActivity</i>	9.813	11.626	12.173	12.158	12.683	14.632
Total Effect	0.149***	0.375***	0.444***	0.442***	0.508***	0.751***

**Table 7**

Descriptive statistics - Japanese stocks.

This table reports summary statistics for market quality and ETF primary market activity computed at the stock-day level. The sample is composed of 2,173 stocks composing the TOPIX and Nikkei between September 2010 to September 2016. Panel A includes only days with BoJ ETF purchases, while Panel B includes only days without purchases. Variable definitions are provided in Table 1.

Panel A: Days with BoJ ETF purchases								
Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
<i>PMStockActivity</i>	652,794	0.33	0.43	0.005	0.14	0.22	0.33	4.09
<i>Return</i>	652,794	-1.01	2.02	-5.89	-2.10	-0.87	0.04	6.78
<i>Volatility</i>	652,794	40.59	26.29	10.54	24.30	33.15	47.57	181.36
<i>QuotedSpread</i>	652,746	0.20	0.21	0.02	0.08	0.13	0.24	1.42
<i>Resiliency</i>	652,514	0.90	0.18	0.22	0.82	0.93	1.01	1.42
<i>Depth</i>	652,794	5.29	6.44	0.21	1.25	3.14	6.68	38.38
<i>Inefficiency</i>	652,471	0.20	0.15	0.002	0.08	0.17	0.29	0.68

Panel B: Days without BoJ ETF purchases								
Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
<i>LastPMStockActivity</i>	1,857,262	0.33	0.44	0.002	0.14	0.24	0.35	4.09
<i>Return</i>	1,876,040	0.42	1.92	-5.89	-0.64	0.27	1.35	6.78
<i>Volatility</i>	1,876,061	39.71	26.26	10.54	23.16	32.11	47.10	181.36
<i>QuotedSpread</i>	1,876,056	0.20	0.21	0.02	0.07	0.13	0.25	1.42
<i>Resiliency</i>	1,874,770	0.89	0.19	0.22	0.80	0.92	1.00	1.42
<i>Depth</i>	1,876,061	5.10	6.43	0.21	1.18	2.91	6.29	38.38
<i>Inefficiency</i>	1,874,622	0.21	0.15	0.002	0.09	0.18	0.31	0.68

#### 4. The causal impact of primary market activities on constituent stocks

All of our analysis to this point has documented associations between the primary market, HFTs' activities, and market quality. However, we ideally want to be able to demonstrate a causal relation between these activities. A quasi-natural experiment in the Japanese market provides us with an opportunity to examine the causal link between primary market activities and market quality. Unfortunately, Japanese data do not identify HFTs' activities, so we cannot address causation in the secondary market in Japan.

##### 4.1. Japanese setting, sample, and empirical approach

The quasi-natural experiment we utilize comes from the Bank of Japan's (BoJ's) purchases of ETFs for their quantitative easing program. The BoJ began purchasing large amounts of ETFs in 2010. On days when it intervenes, on average, the BoJ buys ¥70 billion per day of ETFs, compared to an average daily trading volume of these ETFs of around ¥25 billion. Given the extraordinary amount of BoJ purchases, ETF liquidity providers necessarily create ETF units by buying the constituent stocks. Therefore, we assume that every BoJ purchase leads to the creation of ETF units on the primary market.<sup>16</sup>

The BoJ mainly buys two types of ETFs: (i) ETFs tracking the TOPIX (a market cap-weighted index) and (ii) ETFs tracking the Nikkei (a price-weighted index). BoJ purchases of Nikkei stocks imply over-weighted activity in some Nikkei stocks relative to their market capitalization, providing plausibly exogenous variation in the BoJ's demand for stocks, which provides a unique opportunity to measure the effect of the ETF primary market activities on constituent stocks. In September 2016, the BoJ acknowledged that the purchases of Nikkei ETFs were generating abnormal activity in some stocks, and in response decreased the proportion of Nikkei ETFs in favor of TOPIX ETFs.

Since our identification strategy relies on the large purchases of Nikkei ETFs by the BoJ, our data set captures the period when the BoJ quantitative program involves purchases of both TOPIX and Nikkei ETFs in equal proportions, namely from September 2010 to September 2016. To obtain the dates of the BoJ ETF purchases and their amounts, we use the BoJ website. We retrieve the TOPIX and Nikkei Index composition and weights through time via LSEG. We use LSEG DataScope to obtain the five-minute interval trading and quotation data for the constituent stocks. To avoid rebalancing noise around index exclusion and inclusion, we only keep stocks when they are present in at least one of the two indexes. This dataset includes 2,173 stocks.

Consistent with Charoenwong et al. (2021) (CMW, 2021), we compute the daily primary market (PM) purchases per constituent stock  $i$ , implied by the BoJ ETF purchases per day  $t$ , as:

$$PMStockActivity_{i,t} = \frac{\sum_{ETF} BoJ\ Purchases_{ETF,t} \times StockWeightInETF_{ETF,t,i}}{MarketValue_{i,t-22}} \quad (11)$$

<sup>16</sup> Two pieces of evidence support this assumption. First, Figure IA1 of the Internet Appendix shows that the BoJ held 73% of the nation's ETFs as of March 2019. It follows that these ETF units had to be created. Second, the low tracking error of the ETFs also indicates that the purchased physical ETFs are composed of their constituent stocks. For instance, in May 2023, the largest Japanese ETF, the Nomura Topix ticker 1306, had a tracking difference of only 0.14% with its benchmark over 12 months.

Our variable of interest, *PMStockActivity*, quantifies daily ETF primary market purchases of the constituent stock necessary for APs to create ETF units for BoJ purchases. Importantly, for our identification strategy, we compute *PMStockActivity* for each stock relative to its market capitalization.

To create the market quality variables, we reproduce the method used for French stocks with LSEG DataScope. In line with the French data, we winsorize the variables at the 1% and 99% levels. We present the summary statistics in Table 7. Panel A (Panel B) reports the results for days with (without) BoJ purchases. The average (median) value of the daily stock purchases as a percentage of the market capitalization is 0.33% (0.22%). The average close-to-close return is  $-1.01\%$ , the average quoted spread is 0.2%, and the depth is ¥5.29 million. The statistics are similar across days with and without BoJ purchases except for close-to-close returns, which are negative on BoJ purchase days.

Our identification strategy relies on the disproportionate purchases of Nikkei stocks for the ETF primary market compared to their market capitalization. Since the Nikkei is a price-weighted index, Nikkei stocks with high prices relative to their market capitalization have high *PMStockActivity*. The identifying assumption is that the market quality of these high-price Nikkei stocks mainly differs from the rest of the stocks due to the higher magnitude of *PMStockActivity*. We also provide supporting counterfactual analysis by investigating the difference with days without BoJ purchases. Our empirical approach, therefore, exploits the heterogeneity of *PMStockActivity* in the cross-section of stocks.

The amounts of BoJ purchases per month are announced in advance; however, there is high variability in the choice of days the BoJ makes the ETF purchases. During our sample period, we find that during some weeks, the BoJ did not intervene at all, while at the maximum, the BoJ intervened 4 days per week. The BoJ tends to purchase ETFs when the market is declining (see CMW, 2021), but since the market return is arguably a martingale, the BoJ purchases are plausibly exogenous. To avoid the bias created by the BoJ purchasing ETFs when the market return is negative, we follow CMW, 2021 and run our regression only for the days  $d$  when the BoJ purchases ETFs.<sup>17</sup> As a result of this filter, the dataset is composed of 349 days from the 1,481 days of the original sample period.

Using this sample, we test whether the market quality of the constituent stocks is affected by the *PMStockActivity*. Specifically, for each stock  $i$  included in the TOPIX and Nikkei indexes, and day  $t$ , we run the following regression:

$$Y_{i,t} = \alpha_i + \lambda_t + \beta_1 .PMStockActivity_{i,t} + \beta' .X_{i,t} + \varepsilon_{i,t}, \quad (12)$$

where  $Y_{i,t}$  is one of our measures of market quality: return, volatility, or liquidity. To account for time-invariant stock heterogeneity and common shocks and trends, panel regression (12) includes stock,  $\alpha_i$ , and day,  $\lambda_t$ , fixed effects. The covariates,  $X_{i,t}$ , control for time-varying stock characteristics.<sup>18</sup>

#### 4.2. Japanese results and counterfactual test

We present the estimates of Eq. (12) in Table 8. In Panel A, we investigate the impact of primary market ETF creations on the market quality of constituent stocks. Following the BoJ ETF purchases, APs trade the constituent stocks, in particular, the Nikkei stocks with a high price relative to their market capitalization. In column (1), we show that when APs buy the constituent stocks to create ETF units (in response to the ETF purchases of the BoJ), the primary market increases the returns of the constituent stocks. This result confirms CMW, 2021 and Barbon and Gianinazzi (2019) findings that the BoJ intervention increases the prices of the stocks. In addition, this result also confirms Brown et al. (2021) findings that ETF creations increase the returns of the constituents via arbitrage.

Next, we turn to the effect of ETF arbitrage on the volatility and liquidity of the constituent stocks. The estimates are tabulated in columns (2) to (5). Consistent with the predictions of Malamud (2015), the ETF primary market activities have a statistically significant positive effect on the volatility and the illiquidity of the constituent stocks. When APs purchase for the ETF primary market a large amount of stock relative to its market capitalization, the stock becomes more volatile, and its illiquidity increases. The results are consistent across all our measures of liquidity. This novel result establishes that the ETF primary market is causing an increase in volatility and the illiquidity of the constituent stocks. Therefore, ETFs can have a detrimental impact on their constituent stocks when there is a high volume of primary market arbitrage activity.

The economic magnitude of the result is large. When APs create new ETF units for the BoJ, they trade on average 0.33% of the market capitalization of the constituent stocks. Back-of-the-envelope calculations indicate that an average stock purchase by APs increases its average return by 8%, its volatility by 2%, and its quoted spread by 2%. Additionally, the depth decreases by ¥0.31 million, a 6% reduction from the mean, and the depth resiliency is reduced by 0.66%. Pricing inefficiency declines by 0.05% but is significant only at the 10% level. The results are consistent with the arbitrage channel hypothesis and support the view that ETFs can increase the volatility and illiquidity of their constituent stocks through arbitrage activities.

*Counterfactual test:* To examine our identifying assumption—that the market quality of Nikkei stocks with high prices relative to their market capitalization does not differ significantly from that of other stocks, for reasons other than the APs stock purchases for

<sup>17</sup> On the full sample, our variable *PMStockActivity* would effectively include an indicator variable equal to one on the days of the BoJ purchases. This binary variable would capture the decision of the BoJ to proceed with the purchases of ETFs on these days. We further explore this issue by analyzing the days without BoJ purchases in a counterfactual test, with results presented later in Panel C of Table 8.

<sup>18</sup> Harada and Okimoto (2021) find that the BoJ tends to make more purchase in the afternoon, therefore we also test that our results are robust when using the afternoon trading session only. These results, reported in Table IA12 of the Internet Appendix, are qualitatively similar, except that pricing inefficiency becomes statistically significant only at the 10% level.

**Table 8**

ETF primary market effects - Japanese quasi-natural experiment.

The table reports estimates from OLS regressions of stock market quality proxies on Authorized Participants intervention per stock. Variable definitions are provided in Table 1. Panel A includes only days with BoJ ETF purchases; Panel B includes only days without purchases. Panel C reports Z-tests comparing the coefficients in Panels A and B. Stock and day fixed effects are included. Stock and day fixed effects are included. Double-clustered robust standard errors are presented in parentheses. \*\*\*, \*\*, and \* indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample is composed of 2,173 stocks composing the TOPIX and Nikkei between September 2010 to September 2016.

Panel A: Days with BoJ ETF purchases						
	<i>Return</i>	<i>Volatility</i>	<i>QuotedSpread</i>	<i>Depth</i>	<i>Resiliency</i>	<i>Inefficiency</i>
<i>PMStockActivity</i>	0.239*** (0.054)	1.901*** (0.505)	0.011** (0.005)	-0.943*** (0.214)	-0.018*** (0.002)	-0.003* (0.002)
Stock fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Day fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	652,794	652,794	652,746	652,794	652,514	652,471
R <sup>2</sup>	0.356	0.572	0.530	0.664	0.110	0.161

Panel B: Days without BoJ ETF purchases						
	<i>Return</i>	<i>Volatility</i>	<i>QuotedSpread</i>	<i>Depth</i>	<i>Resiliency</i>	<i>Inefficiency</i>
<i>LastPMStockActivity</i>	0.109*** (0.023)	0.903*** (0.298)	-0.010*** (0.002)	-0.177 (0.113)	-0.009*** (0.001)	-0.006*** (0.001)
Stock fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Day fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,857,262	1,857,262	1,857,257	1,857,262	1,855,993	1,855,851
R <sup>2</sup>	0.247	0.536	0.539	0.643	0.112	0.157

Panel C: Difference between days with and without BoJ purchases						
	<i>Return</i>	<i>Volatility</i>	<i>QuotedSpread</i>	<i>Depth</i>	<i>Resiliency</i>	<i>Inefficiency</i>
Z-test	2.21***	1.70**	3.97***	-3.16***	-3.53***	1.38*
p-value	0.01	0.03	0.00	0.00	0.00	0.08

the ETF primary market — we rerun our experiment with two important modifications. First, we focus exclusively on days without BoJ ETF purchases. Second, we substitute *PMStockActivity* with a placebo, *LastPMStockActivity*, defined as the last non-zero value of *PMStockActivity* on days of BoJ purchases. Given the potential lags in *PMStockActivity*, as highlighted by Evans et al. (2022), the estimates for *LastPMActivity* presented in Panel B are not the primary interest. The purchases of constituent stocks by APs may span multiple days, making *LastPMActivity* an imperfect placebo. Nonetheless, the comparative magnitude of estimates for days with and without BoJ purchases provides valuable insights. Therefore, we compare the estimates for days with BoJ purchases (Panel A) against those for placebo days (Panel B) using Z-tests, as follows:

$$Z\text{-test} = \frac{\hat{\beta}_{1,\text{Panel A}} - \hat{\beta}_{1,\text{Panel B}}}{\sqrt{\sigma^2(\hat{\beta}_{1,\text{Panel A}}) + \sigma^2(\hat{\beta}_{1,\text{Panel B}})}}. \quad (13)$$

The estimates of the Z-tests presented in Panel C of Table 8 confirm that our findings are linked to the BoJ purchases (and to the APs that have to create the ETF units by trading the constituent stocks). On the days when the BoJ purchases ETFs, the return, volatility, and illiquidity of the constituent stocks are significantly larger for the Nikkei stocks with a high price relative to their market capitalization. The comparison of the  $\hat{\beta}_1$  coefficients between the two panels allows for a more conservative estimation of the economic magnitude of the effects of ETF arbitrage. These estimates, which can be viewed as a lower bound, show that the magnitude of the changes is reduced, but remains large and statistically significant. The average stock purchase by APs due to the BoJ ETF purchases increases its return by 4%, its volatility by 1%, its quoted spread by 3%, and its price inefficiency by 1%. The depth decreases by ¥0.26 million, a 5% reduction from the mean, and depth resiliency is reduced by 0.33%. These results support our identification assumption that the Nikkei stocks with a high price relative to their market value differ from the other stocks due to ETF primary market activities. It is, therefore, APs' purchases that are responsible for the effects we document.

*A Bank of Japan effect or a primary market effect?* The objective of the asset purchase programs of the BoJ is to support the economy against deflation.<sup>19</sup> It is important to observe that the BoJ only buys ETFs and not constituent stocks. We, therefore,

<sup>19</sup> See the speech for the Bank of International Settlements (BIS) given by Mr. Masaaki Shirakawa, Governor of the Bank of Japan, "The Bank of Japan's efforts toward overcoming deflation," February 17, 2012: "The purpose of this operation is to encourage a decline in longer-term market interest rates and a reduction in various risk premiums so that financial conditions surrounding the ultimate borrowers of funds, such as firms and households, will become more accommodative."

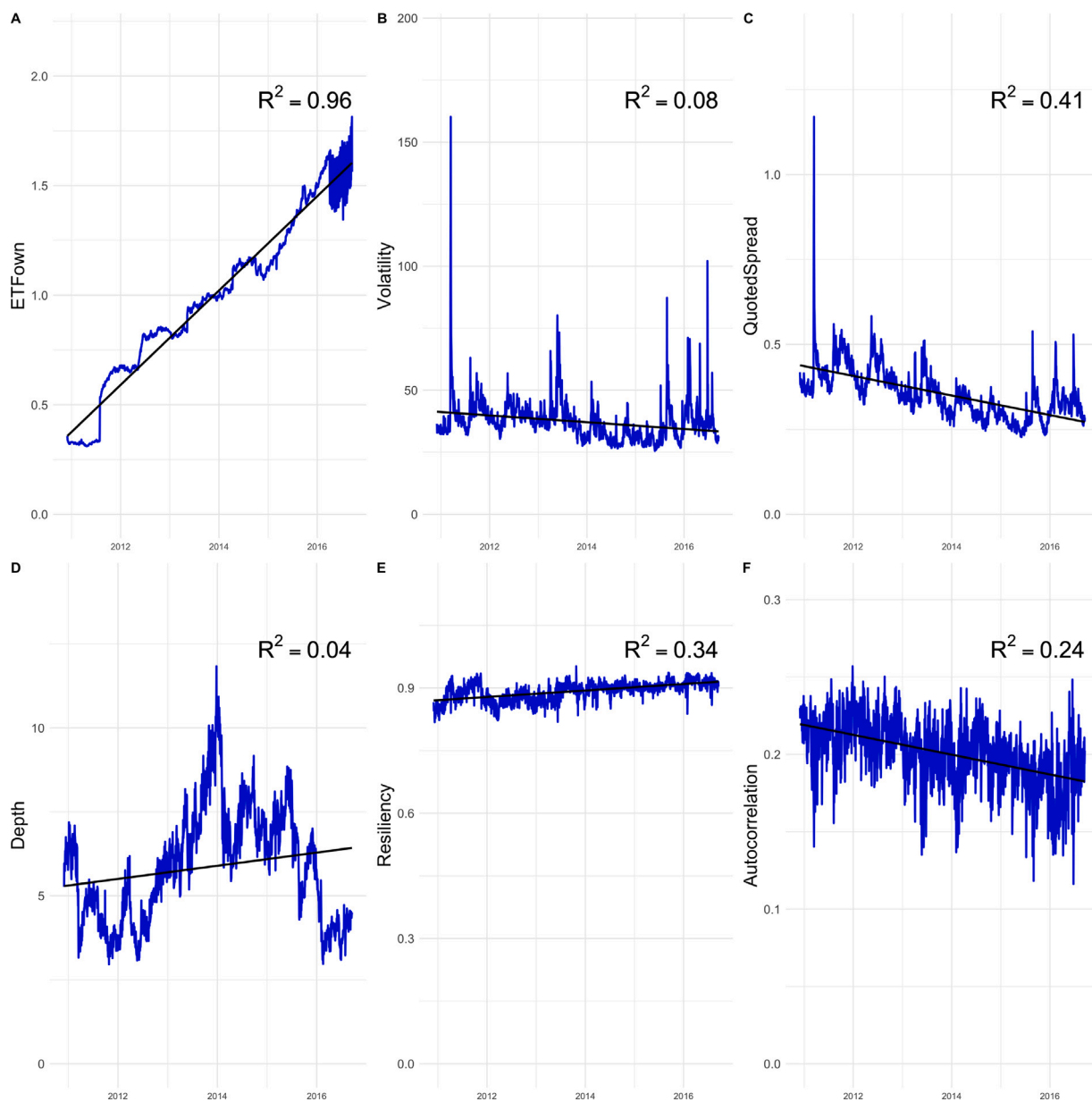


Fig. 2. Evolution of market quality and ETF ownership in Japanese stocks.

These charts report daily ETF ownership and five market quality for Japanese stocks over time. Panel A reports *ETFown*, Panel B *Volatility*, Panel C *QuotedSpread*, Panel D *Depth*, Panel E *Resiliency*, and Panel F *Autocorrelation*. Market quality variables are calculated using a market cap-weighted average. Variable definitions are provided in Table 1. The sample comprises 2,173 stocks composing the TOPIX and Nikkei between September 2010 and September 2016.

measure the effect of the ETF arbitrage channel for ETF creations. Irrespective of the objective of the BoJ, we show that when APs buy constituent stocks, their prices, volatility, and illiquidity all increase. A caveat of this experiment using the BoJ's purchases is that we document only the effect of the arbitrage for ETF creations. We cannot claim that ETF redemptions have the same effects. However, given that the effect is detrimental in the case of purchases, and that financial data tend to be skewed and leptokurtic, it is likely that the effect will be similar (if not stronger) for large sales.

## 5. What effect dominates?

Our results to this point provide conflicting evidence on the impact of ETFs on market quality in their constituent stocks. The obvious next question is which effect dominates. We are unable to offer a causal answer to this question, but we provide suggestive evidence by studying variation in Japanese stocks' market quality during the 2010–2016 period, when they experience a positive

**Table 9**

Evolution of market quality and ETF ownership in Japanese stocks.

This table presents OLS regression estimates of the evolution of market quality and ETF ownership over time. Daily market cap-weighted averages are computed for each market quality measure. All variables are regressed on annualized time to facilitate interpretation. Variable definitions are provided in Table 1. Statistical significance at the 1%, 5%, and 10% levels is indicated by \*\*\*, \*\*, and \*, respectively. The sample includes 2,173 stocks from the TOPIX and Nikkei indices over the period from September 2010 to September 2016.

Variable	Slope	Robust_SE	t-stat	p-value	Significance	R <sup>2</sup>
<i>ETFown</i>	0.19	0.02	11.5	0	***	0.96
<i>Volatility</i>	-1.48	0.5	-2.97	0	***	0.08
<i>QuotedSpread</i>	-0.03	0	-7.13	0	***	0.41
<i>Depth</i>	0.18	0.18	0.98	0.33		0.04
<i>Resiliency</i>	0.01	0	7.31	0	***	0.34
<i>Inefficiency</i>	-0.01	0	-7.81	0	***	0.24

shock to their ETF ownership. First, we compute a daily market-cap-weighted average for all market quality variables and ETF ownership. We then regress this average market quality on time to assess trends in market quality and ETF ownership for Japanese stocks over the sample period. Time is annualized for easier interpretation.

The estimates in Table 9 and plots in Fig. 2 reveal a notable increase in ETF ownership over the 2010 to 2016 period, accompanied by no deterioration in market quality. In fact, despite the 1.14 percentage point annual rise in ETF ownership, the market quality measures improve: volatility decreases by 1.2 percentage points per year, the quoted bid-ask spreads narrow by 3 bps annually, and depth resiliency and autocorrelation also reflect enhanced conditions.

These results show that despite a substantial positive ETF ownership shock, market quality does not deteriorate. While these improvements could also be influenced by other factors, they support the view that, given the low magnitude and frequency of the ETF primary market, the detrimental effects of arbitrage are unlikely to dominate overall. Despite some short-term and local deterioration in market quality due to ETF primary market arbitrage, overall the evidence aligns with the conclusion that ETFs do not degrade the market quality of their constituent stocks.

## 6. Conclusion

ETF primary market arbitrage activities, which have been the focus of the ETF literature, are limited. Consistent with evidence from the U.S. market, in our 2015 sample of French ETFs, creations/redemptions occur on fewer than one in three days and are equivalent to only 13% of secondary market trading volume. Most trading activity in the ETF secondary market is driven by HFTs, whose activity is mostly unrelated to the primary market.

We show that ETF primary and secondary market activities have different impacts on the constituent stocks of ETFs. Consistent with the arbitrage channel, extreme primary market activity, such as during BoJ ETF purchases, increases the volatility and reduces liquidity in overweighted constituent stocks. This aligns with predictions that ETF arbitrage can transmit illiquidity (Cespa and Foucault, 2014) and noise (Malamud, 2015; Ben-David et al., 2018) to the constituent stocks. By contrast, HFTs' secondary market activity is associated with narrower quoted spreads, suggesting an information linkage channel, where their role in enhancing ETF liquidity benefits the liquidity of constituent stocks.

These findings have implications for discussions on the effects of ETFs on their constituent stocks. Extreme primary market activity can reduce liquidity and amplify volatility in constituent stocks, while HFTs' secondary market activity is linked to enhanced liquidity. Given the infrequent nature and small scale of primary market activities in normal market conditions, the beneficial effect of secondary market activity likely dominates the negative primary market impact. We recommend that future research and regulators take into account the possible heterogeneous impact of the primary and secondary markets on market quality.

## CRedit authorship contribution statement

**Carole Comerton-Forde:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Thomas Marta:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization.

## Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.finmar.2025.100983>.

## Data availability

The authors do not have permission to share data.

## References

- Anagnostidis, P., Fontaine, P., Varsakelis, C., 2020. Are high-frequency traders informed? *Econ. Model.* 93, 365–383.
- Barbon, A., Gianinazzi, V., 2019. Quantitative easing and equity prices: evidence from the ETF program of the Bank of Japan. *Rev. Asset Pricing Stud.* 9 (2), 210–255.
- Ben-David, I., Franzoni, F., Moussawi, R., 2018. Do ETFs increase volatility? *J. Financ.* 73, 2471–2535.
- Bessembinder, H., Maxwell, W., Venkataraman, K., 2006. Market transparency, liquidity externalities, and institutional trading costs in corporate bonds. *J. Financ. Econ.* 82 (2), 251–288.
- Bhattacharya, A., O'Hara, M., 2018. Can ETFs increase market fragility? Effect of information linkages in ETF markets. Working paper, Baruch College (CUNY) and Cornell University.
- Biais, B., Foucault, T., Moinas, S., 2015. Equilibrium fast trading. *J. Financ. Econ.* 116 (2), 292–313.
- Box, T., Davis, R., Evans, R., Lynch, A., 2021. Intraday arbitrage between ETFs and their underlying portfolios. *J. Financ. Econ.* 141 (3), 1078–1095.
- Brogaard, J., Heath, D., Huang, D., 2024. ETF sampling and index arbitrage. p. 47, Unpublished working paper, University of Utah and Northeastern University.
- Brogaard, J., Hendershott, T., Riordan, R., 2014. High-frequency trading and price discovery. *Rev. Financ. Stud.* 27 (8), 2267–2306.
- Brogaard, J., Hendershott, T., Riordan, R., 2019. Price discovery without trading: evidence from limit orders. *J. Financ.* 74 (4), 1621–1658.
- Brown, D.C., Davies, S.W., Ringgenberg, M., 2021. ETF arbitrage, non-fundamental demand, and return predictability. *Rev. Financ.* 25 (4), 937–972.
- Budish, E., Cramton, P., Shim, J., 2015. The high-frequency trading arms race: Frequent batch auctions as a market design response. *Q. J. Econ.* 130 (4), 1547–1621.
- Cespa, G., Foucault, T., 2014. Illiquidity contagion and liquidity crashes. *Rev. Financ. Stud.* 27 (6), 1615–1660.
- Charoenwong, B., Morck, R., Wiwattanakantang, Y., 2021. Bank of Japan equity purchases: The (non-)effects of extreme quantitative easing. *Rev. Financ.* 25 (3), 713–743.
- Colliard, J.-E., Hoffmann, P., 2017. Financial transaction taxes, market composition, and liquidity. *J. Financ.* 72 (6), 2685–2716.
- Dannhauser, C.D., 2017. The impact of innovation: evidence from corporate bond Exchange-Traded Funds (ETFs). *J. Financ. Econ.* 125 (3), 537–560.
- Donatoni, E., Paterlini, S., Bazzana, F., 2022. Market making with inventory control and order book information. *Quant. Finance* 22 (3), 597–610.
- Evans, R.B., Moussawi, R., Pagano, M.S., Sedunov, J., 2022. Operational shorting and ETF liquidity provision. Darden Business School Working Paper.
- Foucault, T., Hombert, J., Roşu, I., 2016. News trading and speed. *J. Financ.* 71 (1), 335–382.
- Foucault, T., Kozhan, R., Tham, W.W., 2017. Toxic arbitrage. *Rev. Financ. Stud.* 30 (4), 1053–1094.
- Fremault, A., 1991. Stock index futures and index arbitrage in a rational expectations model. *J. Bus.* 64 (4), 523–547.
- Fulkerson, J.A., Jordan, S.D., Travis, D.H., 2022. Equity ETF arbitrage and daily cash flow. *J. Invest.* 31 (3), 98–121.
- Gromb, D., Vayanos, D., 2010. Limits of arbitrage. *Annu. Rev. Financ. Econ.* 2 (1), 251–275.
- Grossman, S., 1992. The informational role of upstairs and downstairs trading. *J. Bus.* 65 (4), 509–528.
- Harada, K., Okimoto, T., 2021. The BOJ's ETF purchases and its effects on Nikkei 225 stocks. *Int. Rev. Financ. Anal.* 77, 101826.
- Kempf, A., Mayston, D., Gehde, M., Yadav, P.K., 2009. Resiliency: a dynamic view of liquidity. p. 64, AFA 2009 San Francisco Meetings Paper.
- Kumar, P., Seppi, D.J., 1994. Information and index arbitrage. *J. Bus.* 67 (4), 481–509.
- Lettau, M., Madhavan, A., 2018. Exchange-traded funds 101 for economists. *J. Econ. Perspect.* 32 (1), 135–154.
- Madhavan, A., 1996. Security prices and market transparency. *J. Financ. Intermediation* 5 (3), 255–283.
- Malamud, S., 2015. A dynamic equilibrium model of ETFs. (15–37), Swiss Finance Institute Research Paper.
- Métais, C., Roger, T., 2022. Are retail investors less aggressive on small price stocks? *J. Financ. Mark.* 59, 100685.
- Pan, K., Zeng, Y., 2019. ETF arbitrage under liquidity mismatch. Jacobs Levy Equity Management Center for Quantitative Financial Research Paper.
- Pasquariello, P., 2017. Government intervention and arbitrage. *Rev. Financ. Stud.* 31 (9), 3344–3408.
- Saglam, M., Tuzun, T., Wermers, R., 2019. Do ETFs increase liquidity? p. 63, CFR Working Paper.
- Shleifer, A., Vishny, R.W., 1997. The limits of arbitrage. *J. Financ.* 52 (1), 35–55.