Financial Intermediaries and Contagion in Market Efficiency: The Case of ETFs[*]

Claire Hong[†] Frank Weikai Li[‡] Jiang Luo[§] Avanidhar Subrahmanyam[¶]

This Draft: June 2023

[*]We thank Zhiqi Cao (discussant), Zhuo Chen (discussant), Tinghua Duan, Allaudeen Hameed, Byoung Kang, Charles Lee, Dan Li, Roger Loh, Tao Shu (discussant), Johan Sulaeman, Sheridan Titman, Giorgio Valente, Xue Wang (discussant), Xiaoyan Zhang, and seminar participants at the National University of Singapore, Stockholm University, IESEG School of Management, ABFER 2022, CIRF 2022, SFS Cavalcade Asia-Pacific 2022, 2023 SMU Summer Camp, and the Hong Kong Institute for Monetary and Financial Research for their helpful comments. We appreciate Shuoge Qian's outstanding research assistance. Frank Weikai Li gratefully acknowledges financial support from the Hong Kong Institute for Monetary and Financial Research. This paper represents the views of the author(s), which are not necessarily the views of the Hong Kong Monetary Authority, the Hong Kong Institute for Monetary and Financial Research, or its Board of Directors or Council of Advisers. The above-mentioned entities except the author(s) take no responsibility for any inaccuracies or omissions contained in the paper.

[†]Shanghai Advanced Institute of Finance, Shanghai Jiao Tong University, Shanghai, P.R. China (Email: yrhong@saif.sjtu.edu.cn).

[‡]Corresponding author, Singapore Management University, Lee Kong Chian School of Business, Singapore (Email: wkli@smu.edu.sg).

[§]Nanyang Technological University, Singapore (Email:luojiang@ntu.edu.sg).

[¶]Anderson School, University of California at Los Angeles, USA (Email: asubrahm@anderson.ucla.edu).

Financial Intermediaries and Contagion in Market Efficiency: The Case of ETFs

Abstract

We propose that intermediaries' capital constraints cause contagion in the pricing efficiency for assets they manage. We use a simple model to demonstrate this idea for ETFs and their lead market makers (LMMs). Empirically, we show significant comovement in premia for ETFs with the same LMM, which exceeds that in ETFs without a common LMM. The comovement is not due to style or region effects, and is stronger for more capital-constrained LMMs. Around the debt-market disruptions of COVID-19, the non-fixed-income ETFs of LMMs more active in fixed income experience greater premia. Overall, intermediaries' constraints indeed influence comovements in market efficiency.

JEL classification: G12, G14, G21, G23

Keywords: ETFs, financial intermediaries, capital constraints

How do financial intermediaries affect asset prices? This issue has gathered considerable momentum in recent years. An important development has been to link asset prices and risk premia to frictions in financial intermediation. Empirical work that supports intermediary-based asset pricing includes Adrian, Etula, and Muir (2014) and He, Kelly, and Manela (2017), who construct a proxy for an intermediary stochastic discount factor (SDF) that explains cross-sectional variation in asset returns. There is ongoing debate, however, on the reasons for the connection between intermediary balance sheet capacity and asset returns. One challenge is the issue of confounding effects. For example, some argue that the relation between intermediary balance sheet capacity and asset prices might in part be driven by macroeconomic factors, time-varying sentiment or risk aversion (Baron and Xiong (2017); Gomes, Grotteria, and Wachter (2019); Santos and Veronesi (2022)).

We pursue an alternative rationale for how financial intermediaries affect prices by proposing that they can cause contagion in market efficiency. Specifically, we consider the idea that capital constraints of intermediaries influence commonalities in pricing efficiencies across markets where the intermediaries have a key presence. We first develop a simple model of such contagion, and then use exchange-traded funds (ETFs) as a laboratory to test whether sharing a common intermediary enhances comovements in pricing efficiency (as measured by the pricing discrepancy between ETFs and their underlying portfolios).

Why study the ETF market? There are three reasons. First, since each ETF is assigned an intermediary (lead market maker, or LMM), this market allows us to differentiate intermediary-specific capital constraints from aggregate funding constraints (both observed and unobserved). We can test a sharper prediction of intermediary asset pricing theories that intermediary-specific constraints have a larger impact on prices when intermediaries are more likely to be the "marginal" investor (Baron and Muir (2022)). Second, when compared to other financial assets, pricing efficiency in ETFs is cleanly defined — the price of ETFs should perfectly replicate the value of their underlying assets in a frictionless world. Indeed, an ETF premium is an accurate measure of arbitrageurs' expected return, a key object in asset pricing. Thus, ETF arbitrage offers a powerful setting for understanding how intermediary-induced frictions affect asset prices, in contrast to realized returns, which are noisy proxies for expected returns (Merton (1980)). Third, ETFs have grown quickly in both size and scope. As of the end of 2021, there were 2,570 ETFs in the U.S. with total assets under management of around $7.2 trillion.[1] The sheer size and economic importance of the ETF market suggest that understanding ETF pricing efficiency is important.

ETF LMMs, along with other authorized participants (APs), are responsible for ensuring that ETF prices do not deviate significantly from their net asset value (NAV). If LMMs observe any significant premium or discount between an ETF's price and its NAV, they conduct arbitrage activities by taking long (short) positions on the relatively undervalued (overvalued) side.[2] However, arbitrage is capital-intensive and LMMs are subject to capital constraints. Given that an LMM typically needs to maintain the law of one price in many ETFs, a natural prediction is that the pricing gap between ETFs and their constituents should comove across the different ETFs served by the same LMM. The rationale is that if one ETF experiences a higher level of (absolute) pricing gaps due to an exogenous demand shock, the LMM will direct more capital towards that ETF to exploit the arbitrage opportunity.[3] Consequently, less capital will be available to maintain the law of one price for the other ETFs the LMM manages. The model we develop formalizes this economic intuition, and also indicates that a decrease in LMM capital constraints and the costs of arbitraging ETF pricing gaps reduce comovements in ETFs managed by the same LMM.[4] We test these implications utilizing ETF LMM data for January 2012 to December 2020.

We measure ETF pricing efficiency using the ETF premium, defined as the absolute value of the percentage deviation of the price from its NAV. We regress each ETF's daily premium on the average counterpart across all ETFs sharing the same LMM, excluding the focal ETF, in our baseline model. We identify a strong comovement in premium among ETFs sharing the same LMM. The coefficient estimate on the premium is 1.59 (t-stat. = 17.04), and the estimate is eight times higher than the corresponding estimate for ETFs not served by the same LMM. This suggests that a one-standard deviation decrease in the average premium of non-focal ETFs managed by the same LMM leads to a 1.59 bps decrease in the focal ETFs' premium, equivalent to 7.0% of its standard deviation. Since the

[1] Source: 2022 Investment Company Factbook.

[2] The ETF premium can be driven by both demand-side and supply-side factors. Demand-side factors include non-fundamental shocks or shifts in noise traders' sentiment. Supply-side factors include aggregate funding liquidity or intermediary capital constraints, which is the focus of our study.

[3] LMMs rationally allocate capital to correct ETF prices until the marginal benefit of arbitrage per unit of capital is equalized across different ETFs. When one ETF's price moves away from its intrinsic value due to exogenous reasons, the marginal benefit of arbitrage for that ETF becomes greater.

[4] One potential explanation for the comovement in premia for assets managed by the same LMM could be the algorithms used for arbitrage are common for all assets managed by an LMM, and different across different LMMs. However, unlike violations of other parity conditions (such as put-call parity), ETF premium is a model-free measure for the violation of the law of one price. Hence, the algorithms used to arbitrage on ETF premia should be similar across different intermediaries. This means that we can reliability attribute the commonality to LMM capital constraints.

annual dollar trading volume in all ETFs managed by an average LMM is around $528 billion during our sample period, a one standard deviation decrease in the non-focal average LMM premium results in an annual dollar savings of $84.0 million to investors who trade ETFs managed by the LMM on opportune days, and vice versa.[5]

In our tests, we control for a list of ETF characteristics that may affect the ETF premium. We include style-day fixed effects, where "style" refers to the detailed style category to which the focal ETF belongs. The inclusion of this fixed effect helps alleviate the concern that the LMM-level comovement in ETF premia might be driven by investors' correlated (time-varying) demand for ETFs belonging to the same investment style. To further ensure that LMM-level premium comovement is not driven by commonality in the premium across all ETFs, we orthogonalize the premium with respect to its non-LMM counterpart, and use the residual premium as our variable of interest in most of our empirical tests.[6] We also investigate whether the ETF premium exhibits excess correlations in the right tails of the distributions. Intuitively, when the average premium of ETFs managed by the LMM is large, the LMM likely faces severe capital constraints, which increases the chance of observing high premia for the focal ETFs. Using a quantile regression approach following Boyson, Stahel, and Stulz (2010), we find that the focal ETF's premium has a 16.6% probability of being above the 90th quantile when the non-focal average LMM premium is also above the 90th quantile, compared to an unconditional probability of 10% if there were no dependence.

To show that LMMs play a causal role in driving ETF premium comovement, we conduct an event study around the days when an ETF changes its LMM. A significant fraction of these LMM change events are due to mergers between two LMMs.[7] It can be reasonably assumed that a change of LMM for an *individual* ETF is relatively exogenous to the ETF's unobserved characteristics that may drive comovement in premia. Using an event-study approach, we find that the focal ETF's premium comovement with that of its old LMM reduces significantly from a pre-level of 1.17 bps to a magnitude of close-to-zero after the ETF switches to a new LMM. The pre- and post-difference in comovement is 1.03 bps (t-stat. = 4.28). Further, the focal ETF exhibits stronger comovement with

[5]Our estimate provides a lower bound for the effect of LMM-specific capital constraints on the pricing of ETFs. The effect of common LMM constraints on ETF premia is removed via the inclusion of style-day fixed effects and the construction of the residual premium.

[6]Since our regression models include style-day fixed effects, they already account for comovement in ETF premia due to systematic factors. However, the difference between style-day fixed effects and using the regression residual is that the latter allows each ETF to have differential exposure to the aggregate ETF premium.

[7]For example, Virtu Financial acquired KCG Holdings in July 2017, which accounts for 37% of the LMM switching events in our sample.

that of its new LMM after the switch. The comovement with the new LMM increases by a magnitude of 1.47 bps (t-stat. = 2.95) from its pre-level of 0.24 bps (t-stat. = 1.24). The absence of comovement in the premium between the focal ETF and the new LMM before the switch and the presence of strong comovement after the switch, together suggest that our finding is unlikely explained by market-wide funding constraints that unanimously affect all ETFs.

Next, we examine heterogeneity in ETF premium comovement due to LMM capital constraints. Intuitively, ETFs with higher return volatility, lower liquidity, and smaller market cap should require more costly liquidity provision from their LMMs to maintain the law of one price. We interact these ETF characteristics with the non-focal average LMM premium. We find that the interaction terms are significantly positive for ETF volatility and illiquidity, and significantly negative for ETF size, consistent with our conjecture that the premium comovement is stronger for ETFs that are more costly to arbitrage. We also predict that the comovement should be more pronounced when the *underlying assets* of the ETF are more costly to arbitrage. To test this idea, we restrict our sample to ETFs with U.S. equity as the underlying asset. Aggregating stock-level bid-ask spreads, return volatility, and lendable supply at the ETF level as proxies of arbitrage costs, we find that the LMM-level comovement effect is stronger for ETFs with underlying assets in which arbitrage is costlier.

To further examine the role of LMMs, we construct two measures to capture LMM-specific capital constraints: the total market capitalization of ETFs managed by the LMM, and the number of active APs for each ETF in a year. We then regress the focal ETF's premium on the interaction between the LMM-specific capital constraints and the non-focal average LMM premium. Consistent with our conjecture, we find stronger ETF premium comovement when the LMM faces more binding constraints. Moreover, for a smaller sample of LMMs, we construct more direct measures of capital constraints, utilizing LMMs' capital position information from the Commodity Futures Trading Commission (CFTC), as well as changes in individual LMMs' net worth during a short window around earnings announcements. Consistently, in periods with tightening of LMM-specific capital constraints, we observe a stronger comovement in the premia of ETFs intermediated by these LMMs.

To provide causal evidence that LMM-specific capital constraints drive comovement in ETF pricing inefficiencies, we conduct a difference-in-differences (DiD) test around the onset of COVID-19, using the fact that fixed income ETFs experienced unprecedented large discounts during the COVID-19 market sell-off (Falato, Goldstein, and Hortaçsu (2021); Haddad, Moreira, and Muir

(2021)). The advantage of this setting is that COVID-19 pandemic is an exogenous shock that originates outside the financial sector. The idea is that LMMs who manage relatively more fixed income ETFs likely experience more binding capital constraints during the COVID-19 pandemic. We hypothesize that non-fixed income ETFs managed by more constrained LMMs should experience greater pricing gaps, compared to ETFs that are managed by less constrained LMMs. Our results are consistent with our prediction. The DiD analysis provides evidence that negative shocks to LMMs' capital constraints led to greater ETF pricing gaps. The results also have policy implications, as they show that inefficiencies can spill over across ETF segments via the sharing of common intermediaries.

We conduct several supplementary tests. First, we perform subsample analyses conditional on aggregate funding constraints. We use four proxies for constraints: the VIX, the credit spread, the intermediary capital ratio of He, Kelly, and Manela (2017), and the prior month's stock market return (Hameed, Kang, and Viswanathan (2010)). Subsample tests reveal that LMM-induced ETF premium comovement is similarly strong and significant during both periods of tightened and loosening aggregate funding constraints. This suggests that the role of LMM-specific capital constraints is independent from the impacts of aggregate funding constraints. Second, we find the premium comovement effect is larger for ETFs with a style that is different from that of the focal ETF, suggesting that our finding is unlikely driven by investors' correlated demands for ETFs tracking the same investment style. Third, we include additional fixed effects to rule out the possibility that our result is driven by investors' correlated demand for ETFs sharing similar characteristics other than investment style.[8] Fourth, we conduct our baseline analysis for ETFs that track different assets, and find that our central result holds for all types of ETFs except the highly liquid currency ETFs.[9] Fifth, we find the premium comovement is pervasive across ETFs with different geographical coverage, with the estimated effect ranging from a low of 0.38 for North America to a high of 3.56 for Asia-Pacific ETFs. Lastly, the results are similar when we control for the average LMM premium in the prior month and past ETF returns, for both ETF premiums and discounts, for more mature ETFs, and when we construct value-weighted LMM premia.

One important caveat of our setting is that we focus mainly on LMMs of ETFs. Although an LMM is an important financial intermediary with a principal obligation to provide liquidity and maintain the law of one price, an ETF can potentially have other authorized participants (APs) that

[8] Specifically, we include region-day, exchange-day, issuer-day, and distributor-day fixed effects.
[9] The LMM-level comovement in ETF premia also exists for a sample of leveraged/inverse ETFs.

also contribute to liquidity provision. However, the presence of other APs should only weaken the role of the LMM, and bias us against finding any LMM-level comovement effect. Using information on APs reported in SEC N-CEN filings, we indeed find evidence that the presence of active APs mitigates the impact of LMMs' capital constraints.[10]

1 Institutional Background and Literature

In Section 1.1, we describe the institutional background of ETF arbitrage and the role of LMMs in maintaining the law of one price for ETFs. In Section 1.2, we review the related literature and highlight the paper's contributions.

1.1 *Institutional background of ETF arbitrage*

ETFs are passive investment vehicles that seek to mimic the returns of baskets of securities. They are traded on both primary and secondary markets. On the latter, ETFs are actively traded by both institutional investors and retail investors, with the price of an ETF determined by supply and demand. As a result, the price of an ETF can diverge from the NAV of its underlying assets. To minimize the divergence between the price of the ETF and its NAV, the ETF sponsor reports the NAV of the ETF's underlying assets every 15 seconds during the trading day. By doing so, the ETF sponsor helps facilitate arbitrage across primary and secondary markets.

On the primary market, LMMs, along with other authorized participants (APs), play a critical role in facilitating the functioning of the ETF ecosystem. They create and redeem ETF units to ensure that an ETF's market price and NAV are closely linked. For example, RBC Capital Markets, one of the LMMs in our sample, mentions that LMMs "fulfill other important roles in addition to providing liquidity and maintaining market equilibrium – they also help to ensure the market price of each ETF unit reflects the value of its underlying securities intraday."[11] While other APs can typically trade as they please, firms acting as LMMs must consistently offer competitive buy-and-sell quotes for their assigned ETFs, and they receive rebates on exchange fees. As Figure 1 illustrates, when an ETF trades at a premium relative to the price of the underlying basket of assets, APs buy the underlying assets, exchange them for "creation units" from the ETF sponsor, and sell those units on the secondary

[10] Arora et al. (2020) show that the market of authorized participants is highly concentrated, with 8 (3) APs accounting for around 80% of total gross creations and redemptions of equity (fixed income) ETFs in 2019. The high concentration suggests that capital constraints may also be an issue for APs if they need to simultaneously manage a large number of ETFs.

[11] https://www.rbcgam.com/documents/en/articles/what-is-the-role-of-the-market-maker-for-etfs.pdf.

market, thereby harvesting the spread between the price of the ETF and that of the underlying assets. In practice, an AP could buy the underlying basket and simultaneously short the ETF, and reverse the positions at the end of the trading day. Such arbitrage activity reduces the ETF premium. Conversely, when the ETF trades at a discount relative to the price of the underlying basket of assets, APs buy the ETF on the secondary market, redeem them through the ETF sponsor for baskets of underlying securities, and offload the underlying securities in the market, and such arbitrage activity narrows the ETF price discount.

On the secondary market, arbitrageurs such as hedge funds and high-frequency traders can take advantage of the price differential between the ETF and the underlying basket of securities without accessing the primary market. When the price of the ETF exceeds (falls below) that of the underlying assets, the arbitrageur can take a long (short) position in the underlying basket of assets, short (go long in) the more expensive ETF, and wait for prices to converge to realize an arbitrage profit. ETF prices can also be arbitraged against other ETFs (Petajisto (2017)) or against futures contracts (Richie, Daigler, and Gleason (2008)). However, such arbitrage trades are exposed to holding costs and idiosyncratic risk for as long as the arbitrage trade is kept open (Pontiff (2006)). Moreover, short sales constraints may prevent arbitrageurs from conducting such activities in the first place. For these reasons, even though arbitrageurs can engage in ETF arbitrage in the secondary market, LMMs and APs can do so with much lower arbitrage risk.

1.2 *Related literature and our contribution*

First, our paper is related to the important literature on intermediary-based asset pricing. One of the key predictions from these studies is that liquidity provision by financially constrained intermediaries is a main driver of comovement in the pricing efficiency of intermediated assets (e.g., Adrian, Etula, and Muir (2014); He, Kelly, and Manela (2017)).[12] Although prior studies find supportive evidence for intermediary-based asset pricing, the relationship between intermediary balance sheet capacity and asset prices could at least partially be driven by macroeconomic factors or time-varying sentiment or risk aversion (Baron and Xiong (2017); Gomes, Grotteria, and Wachter (2019); Santos and Veronesi (2022)). Some papers reverse the idea, using the common component of market inefficiencies as a measure of financial market dislocation and linking it to aggregate funding

[12]Other related studies on intermediary-based asset pricing include He, Khorrami, and Song (2019); Baron and Muir (2022); Goldberg and Nozawa (2021); Haddad and Muir (2021); and Macchiavelli and Zhou (2022).

constraints (Pasquariello (2014); Rösch, Subrahmanyam, and Van Dijk (2017)).

Recent studies emphasize the role of individual intermediaries' capital constraints on the pricing efficiency of certain assets. For example, in the foreign exchange market, Du, Tepper, and Verdelhan (2018) show that deviations from covered interest rate parity are particularly strong for contracts that appear on banks' balance sheets at the end of the quarter.[13] Utilizing a regulation reform in the United Kingdom on the leverage ratio of dealers, Cenedese, Della Corte, and Wang (2021) provide similar evidence. Lewis, Longstaff, and Petrasek (2021) find strong commonality in the mispricing of corporate bonds guaranteed by the full faith and credit of the U.S. government, which can be explained by dealer funding costs.

Different from the above studies that focus on mispricing within an asset class, we investigate pricing efficiency comovement across ETFs tracking all major asset classes, including U.S. equities, global equities, fixed income securities, commodities, currencies, and real estate. Indeed, a disaggregated analysis shows that our central result holds in virtually all asset classes. Further, using the debt market disruption of COVID-19 as an exogenous shock to LMMs' capital constraints, we show inefficiency contagion across ETFs tracking different assets via the common LMM link. A key differentiating factor in our paper is that since the price of an ETF should perfectly replicate the value of its underlying assets, pricing efficiency in ETFs is cleanly defined. We thus complement previous work by offering more direct evidence on the causal relationship between financial intermediaries' capital constraints and the pricing efficiency of intermediated assets.

Our paper also contributes to the burgeoning literature that examines the impact of rising ETFs on financial markets. Although the introduction of ETFs substantially lowered management fees and introduced greater intraday trading flexibility for investors, practitioners and academics alike have expressed concerns about the potential negative effects of ETFs. Some recent evidence suggests that ETFs can increase systemic risk, induce non-fundamental volatility and excess comovement, and impede price discovery for individual constituent stocks (Israeli, Lee, and Sridharan (2017); Ben-David, Franzoni, and Moussawi (2018); Da and Shive (2018)). On the other hand, some studies document that ETFs can improve the price efficiency of the underlying stocks, by allowing investors to exploit stock mispricing through hedging (Huang, O'Hara, and Zhong (2021)) and facilitate the

[13] In untabulated analysis, we find comovement in ETF premia is not significantly higher at quarter ends. This may be due to the fact that most LMMs in our sample are non-bank intermediaries that are regulated less stringently relative to banks.

transmission of systematic information into the underlying stocks' prices (Bhojraj, Mohanram, and Zhang (2020); Glosten, Nallareddy, and Zou (2021)). While previous studies mostly focus on the impact of ETFs on the underlying constituent securities, few examine whether the price formation process is efficient at the ETF level and what factors may improve or impede the efficient pricing of ETFs. This is an important question as, like all other assets, ETFs may be subject to non-fundamental demand shocks that drive prices temporarily away from their fundamental value.

Among the few studies that examine price efficiencies at the ETF level, Petajisto (2017) finds that ETF prices significantly deviate from their NAVs, particularly for ETFs holding illiquid securities. Similarly, Bae and Kim (2020) document that illiquid ETFs have large tracking errors. Brown, Davies, and Ringgenberg (2021) show, theoretically and empirically, that creation and redemption activities (ETF flows) provide signals of non-fundamental demand shocks and negatively predict future ETF returns. Pan and Zeng (2019) and Gorbatikov and Sikorskaya (2021) provide evidence that ETF arbitrage is limited by the balance sheet space constraints of authorized participants. Our paper focuses on the comovement in ETF premia, instead of focusing on the level of the premium.[14] We find that this comovement is higher for LMMs with more severe capital constraints, which confirms a link between the efficacy of intermediation and market efficiency.

Finally, our study is also related to the literature on the excess return comovement among firms or funds sharing similar characteristics, such as firms headquartered in the same state (Pirinsky and Wang (2006)); stocks belonging to the same indices (Barberis, Shleifer, and Wurgler (2005); Greenwood (2008); Boyer (2011)); stocks priced at similar levels (Green and Hwang (2009)); firms covered by similar sets of analysts (Israelsen (2016)); stocks held by a common set of mutual funds (Anton and Polk (2014)); stocks that pay dividends (Hameed and Xie (2019)); and hedge funds sharing the same prime broker (Chung and Kang (2016)). Boyson, Stahel, and Stulz (2010) find strong evidence of negative return contagion across hedge fund styles. One issue in interpreting these excess return comovement studies is that it is often challenging to establish whether the return comovement is indeed excessive (Grieser, Lee, and Zekhnini (2020)). The advantage of our setting is that we can directly observe a model-free measure of pricing efficiency, which allows us to rule out fundamental or information-based explanations for the comovement in ETF premia. Moreover,

[14] Broman (2016) documents comovement in ETF premia across ETFs with similar investment styles, which can be attributed to investors' correlated demand for ETFs tracking the same investment styles. Our paper instead focuses on the liquidity provision role of LMMs in driving comovement.

differing from previous studies that focus on the demand-side factors in driving return comovement, we focus on the supply-side by examining the liquidity provision role of LMMs.[15]

2 The Model

In this section, we present a simple model that serves as the basis for our empirical tests. All proofs, unless otherwise stated, are in Appendix A.

2.1 *The setting*

We use a setting with two dates, denoted as 0 and 1. Investors trade at Date 0, and consume at Date 1. There are two markets: the underlying asset market (termed the "stock market") and the ETF market.

The stock market: There are $K \geq 2$ stocks; each stock is indexed by κ. At Date 1, stock κ pays a liquidating cash flow:

$$V_\kappa = \theta_\kappa,$$

where θ_κ follow independent and identical (i.i.d.) normal distributions with mean zero and variance ν. The supply of each stock is normalized to be zero.

A mass M of fundamental investors trades in the stock market. The mth such investor has a standard exponential utility function given by:

$$U(W_{m1}) = -\exp(-AW_{m1}),$$

where W_{m1} denotes the wealth at Date 1 and A is a positive constant representing the absolute risk-aversion coefficient. There is also a group of liquidity traders in the stock market. At Date 0, they have a liquidity demand z_κ for stock κ, where z_κ are i.i.d with mean zero and variance μ.

The ETF market: There are two synthetic securities, each representing an ETF indexed by $i = 1, 2$. ETF i is constructed based on a subset S_i of $\hat{K}$ stocks. At Date 1, ETF i pays a liquidating cash flow

$$\hat{V}_i = \sum_{\kappa \in S_i} \theta_\kappa.$$

Since the ETF is effectively a derivative based on the existing stocks, its supply equals zero. For

[15] The LMM-level comovement in ETF premia implies that the benefits of diversification may be circumscribed for investors using ETFs for portfolio construction, especially during periods when financial intermediaries have more severe funding constraints.

simplicity, we assume that S_1 and S_2 have no overlap (i.e., $S_1 \cap S_2 = \emptyset$).

A mass N of fundamental investors trades only in the ETF market. The nth such investor has a standard exponential utility function given by:

$$U(W_{n1}) = -\exp(-AW_{n1}),$$

where W_{n1} denotes the wealth at Date 1. There is also a group of liquidity traders in the ETF market. At Date 0, they have a liquidity demand $\hat{z}_i$ for ETF i, where $\hat{z}_i$ are i.i.d. normal with mean zero and variance $\hat{\mu}$.

In our model, fundamental investors and liquidity traders trade in either the stock market or the ETF market, but not both. This trading restriction can be motivated by investment mandates (Almazan, Brown, Carlson, and Chapman (2004)), or limited attention (Peng and Xiong (2006)).

Market maker: In our main analysis, we consider a scenario where a single arbitrageur has access to both the stock and the ETF markets. We refer to this arbitrageur as the "LMM" throughout our analysis. Furthermore, for the moment, we assume that both ETFs 1 and 2 are served by the same LMM.

For each ETF i (where $i = 1, 2$), the LMM chooses to create Y_i shares of the ETF and correspondingly buys Y_i shares of each constituent stock in the subset S_i, where the value of Y_i is determined endogenously. The LMM incurs a cost for this creation, which is given by $C_i = cY_i^2$. The parameter c denotes the cost of arbitrage; it is a reduced form way to capture trading frictions, such as illiquidity in the underlying stocks and/or the ETF, which require capital commitments. Further, we assume that the LMM has a limited amount of capital, leading to the capital constraint:

$$\sum_{i=1}^{2} C_i \leq \Phi. \tag{1}$$

Here, $\Phi > 0$ is a constant parameter representing the LMM's constraint. A lower value of Φ indicates that the LMM is not well-funded and has a more limited capacity for trading.

In addition, there exists a risk-free asset with a constant gross rate of return and a price set to unity. The risk-free asset is accessible to all participants in the economy. Finally, for simplicity, we assume that all the random variables, θ_κs, z_κs, and $\hat{z}_i$s, are independent of each other.

2.2 *The equilibrium*

Let Q_i denote the price of ETF i (where $i = 1, 2$), and P_κ denote the price of stock κ (where $\kappa = 1, ..., K$). The ETF's NAV is expressed as:

$$NAV_i = \sum_{\kappa \in S_i} P_\kappa.$$

The ETF's signed pricing gap is given by $Q_i - NAV_i$.

We follow Kyle (1989) to consider the symmetric equilibrium with strategic trading by the LMM. The starting conjecture for the equilibrium is as follows:

- In the stock market, each of the mass M of fundamental investors submits a demand X_κ for stock κ according to the following trading strategy:

$$X_\kappa = -\lambda P_\kappa, \quad \text{where} \quad \lambda = 1/(A\nu).$$

- In the ETF market, each of the mass N of fundamental investors submits a demand $\hat{X}_i$ for ETF i according to the following trading strategy:

$$\hat{X}_i = -\hat{\lambda} Q_i, \quad \text{where} \quad \hat{\lambda} = \lambda/\hat{K}.$$

- The LMM creates Y_i shares of ETF i and buys Y_i shares of each constituent stock $\kappa \in S_i$ according to the following strategy:

$$Y_i = G\left(Q_1 - NAV_1, Q_2 - NAV_2\right),$$

  where $G(.,.)$ is a function of the pricing gaps in the two ETFs.

We verify and derive the conjectured equilibrium strategies in the proof of Proposition 1. We define two random variables $\bar{z}_{\kappa,i}$ and Z_i as follows:

$$\bar{z}_{\kappa,i} \equiv \sum_{\kappa \in S_i} z_\kappa \Big/ \hat{K}, \quad \text{and} \quad Z_i \equiv \frac{\hat{z}_i}{N} - \frac{\bar{z}_{\kappa,i}}{M}.$$

Thus, $\bar{z}_{\kappa,i}$ denotes the cross-sectional average liquidity demand for ETF i's constituent stock $\kappa \in S_i$, and Z_i denotes the difference between ETF i's normalized liquidity demand and the average liquidity demand of its constituent stocks. Henceforth, we refer to Z_i simply as the demand shock in ETF i. Further, to minimize notation, we define a parameter η and two random variables δ_i and Δ_i as follows:

$$\begin{aligned} \eta &\equiv \frac{1}{c + (\hat{\lambda})^{-1}(N^{-1} + M^{-1})}, \\ \delta_i &\equiv \frac{Z_i}{\hat{\lambda} + \eta(N^{-1} + M^{-1})}, \quad \text{and} \quad \Delta_i \equiv \frac{Z_i}{\hat{\lambda}} \left[1 - \left(N^{-1} + M^{-1}\right) \sqrt{\frac{\Phi/c}{Z_1^2 + Z_2^2}}\right]. \end{aligned} \tag{2}$$

Proposition 1 provides the equilibrium prices for the ETF and the underlying stocks.

Proposition 1. *If both ETFs are served by the same LMM, the equilibrium prices of ETF i (where $i = 1, 2$) and each constituent stock $\kappa \in S_i$ are given as follows:*

(i) If $\delta_1^2 + \delta_2^2 < \Phi/(c\eta^2)$, then

$$Q_i = \frac{-\eta\delta_i + \hat{z}_i}{N\hat{\lambda}} \quad \text{and} \quad P_\kappa = \frac{\eta\delta_i + z_\kappa}{M\lambda};$$

the ETF's pricing gap $Q_i - NAV_i = \delta_i$.

(ii) If $\delta_1^2 + \delta_2^2 \geq \Phi/(c\eta^2)$, then

$$Q_i = \frac{1}{N\hat{\lambda}} \left(-\Delta_i \sqrt{\frac{\Phi/c}{\Delta_1^2 + \Delta_2^2}} + \hat{z}_i\right) \quad \text{and} \quad P_\kappa = \frac{1}{M\lambda} \left(\Delta_i \sqrt{\frac{\Phi/c}{\Delta_1^2 + \Delta_2^2}} + z_\kappa\right);$$

the ETF's pricing gap $Q_i - NAV_i = \Delta_i$.

The price of stock $\kappa \notin S_1 \cup S_2$ is given by: $P_\kappa = z_\kappa/(M\lambda)$.

In the proof of Proposition 1, we show that when the absolute scales of the liquidity shocks, $|Z_1|$ and $|Z_2|$, are sufficiently small such that $\delta_1^2 + \delta_2^2 < \Phi/(c\eta^2)$, the capital constraint for the LMM (i.e., Condition (1)) does not bind. In this case, the pricing gap in ETF i is given by $Q_i - NAV_i = \delta_i$. Intuitively, if δ_i is higher, the LMM has more incentives to create additional shares of the ETF and purchase more shares of each constituent stock κ (i.e., a higher Y_i). Consequently, the price Q_i of the ETF decreases, while the price P_κ of each constituent stock κ increases. Note that the prices Q_i and P_κ do not depend on the premium δ_j associated with the other ETF j ($j \neq i$).

When $|Z_1|$ and $|Z_2|$ are sufficiently large such that $\delta_1^2 + \delta_2^2 \geq \Phi/(c\eta^2)$, the capital constraint for the LMM (i.e., Condition (1)) binds. In this case, ETF i's pricing gap is given by $Q_i - NAV_i = \Delta_i$. Now, the ETF's price Q_i and each constituent stock κ's price P_κ depend not only on the ETF's pricing gap Δ_i, but also on the pricing gap Δ_j associated with the other ETF j. This indicates the presence of a spillover effect between the two ETFs due to the capital constraint.

We use Proposition 1 to express ETF i's premium as:[16]

$$|Q_i - NAV_i| = \begin{cases} |\delta_i| & \text{if } \delta_1^2 + \delta_2^2 < \Phi/(c\eta^2), \\ \\ |\Delta_i| & \text{if } \delta_1^2 + \delta_2^2 \geq \Phi/(c\eta^2). \end{cases} \tag{3}$$

We obtain the following results:

Proposition 2. *Consider the scenario where both ETFs share the same LMM.*

(i) If $\delta_1^2 + \delta_2^2 < \Phi/(c\eta^2)$, then the premium of ETF i, $|Q_i - NAV_i| = |\delta_i|$, increases with $|Z_i|$, but does not depend on $|Z_j|$ where $j \neq i$.

(ii) If $\delta_1^2 + \delta_2^2 \geq \Phi/(c\eta^2)$, then the premium of ETF i, $|Q_i - NAV_i| = |\Delta_i|$, increases with both $|Z_i|$ and $|Z_j|$.

The premium in ETF i increases with the absolute scale of its own liquidity shock, $|Z_i|$. This is because a higher $|Z_i|$ results in a greater divergence between the price Q_i of the ETF and the prices P_κs of its constituent stocks. The LMM aims to correct this divergence through arbitrage activities, such as creating more shares of the ETF (i.e., a higher Y_i). However, these arbitrage activities come at a cost, preventing the LMM from completely eliminating the divergence.

When $\delta_1^2 + \delta_2^2 \geq \Phi/(c\eta^2)$, so that the capital constraint for the LMM (Condition (1)) binds, the premium of ETF i is also influenced by the liquidity shock in the other ETF j, $|Z_j|$. The reason for this spillover effect is that a higher $|Z_j|$ leads to a higher absolute premium $|\Delta_j|$ for ETF j. The LMM recognizes this arbitrage opportunity and directs more capital towards ETF j to exploit it. Consequently, less capital is available to correct the pricing gap in ETF i. In equilibrium, the LMM

[16] Following Petajisto (2017), we call this measure an ETF premium even though it could be either a premium or a discount.

allocates capital rationally to correct the pricing gap in ETFs until the marginal benefit of arbitrage per unit of capital is equalized across different ETFs.

We quantify the comovement between ETF i's premium, $|Q_i - NAV_i|$, and that of the other ETF j (where $j \neq i$), $|Q_j - NAV_j|$, using the following parameter:

$$\beta = \frac{\mathrm{Cov}(|Q_i - NAV_i|, |Q_j - NAV_j|)}{\mathrm{Var}(|Q_j - NAV_j|)},$$

which represents the slope coefficient in the regression: $|Q_i - NAV_i| = Intercept + Slope \times |Q_j - NAV_j| + \epsilon$ for $i \neq j$. We obtain the following result:

Proposition 3. *When ETFs share the same LMM, there is a positive comovement between the premium of ETF i, $|Q_i - NAV_i|$, and the premium of the other ETF j (where $j \neq i$), $|Q_j - NAV_j|$; that is, the comovement parameter $\beta > 0$.*

The intuition for this proposition is as follows. According to Proposition 2, the premium in ETF i increases in the absolute scale of its own liquidity shock, $|Z_i|$. In addition, when $\delta_1^2 + \delta_2^2 \geq \Phi/(c\eta^2)$ so that the capital constraint for the LMM (Condition (1)) binds, the liquidity shock to ETF j depletes the arbitrage capital available for ETF i; consequently, the premium of ETF i is also influenced by the liquidity shock to ETF j. This spillover effect from ETF j to ETF i results in the comovement in premia across the two ETFs.

The comovement in ETF premia depends on the cost of arbitrage (represented by the parameter c), return volatility (represented by the parameter ν), and the LMM's funding condition (represented by the parameter Φ). We obtain the following result:

Corollary 1. *When ETFs share the same LMM, the comovement parameter β converges to zero either as the cost of arbitrage $c \to 0$, or return volatility $\nu \to 0$, or as the LMM becomes increasingly well-funded, i.e., $\Phi \to \infty$.*

Intuitively, if it costs little to arbitrage (i.e., $c \to 0$) and/or the LMM is well-funded (i.e., $\Phi \to \infty$), then the liquidity shock in ETF j does not materially deplete the arbitrage capital available for ETF i; consequently, the premium of ETF i remains largely unaffected by the shock. In this case, there is no spillover effect from ETF j to ETF i, implying no comovement in premia across the two ETFs. If the return volatility is sufficiently low (i.e., $\nu \to 0$), then fundamental investors in the stock and

ETF markets price stocks and ETFs at their expected payoffs. As there is no pricing gap between an ETF and its constituent stocks (i.e., $|Q_i - NAV_i| = 0$ for $i = 1, 2$), it holds trivially that there is no comovement in ETF premia. The setting precludes further analyzing the parameter β other than for the limiting cases above, necessitating numerical analysis that we present in Section 2.3 below.

The comovement in ETF premia is also influenced by whether ETFs are served by the same or different LMMs. To see this, consider a scenario where two LMMs have access to both the stock and the ETF markets, and ETFs 1 and 2 are each served by one of the two LMMs. In this case, the LMM serving ETF i creates Y_i shares of the ETF, and each LMM is subject to its own capital constraint (i.e., $C_i = cY_i^2 \leq \Phi$ where $i = 1, 2$). In this scenario, there is no spillover effect; consequently, there is no comovement in ETF premia. Proposition 4 presents this result formally.

Proposition 4. *If ETFs 1 and 2 are served by different LMMs, then there is no comovement between the premium of ETF i, $|Q_i - NAV_i|$, and that of ETF j (where $j \neq i$), $|Q_j - NAV_j|$; that is, the comovement parameter $\beta = 0$.*

2.3 *A numerical analysis of comovement in ETF premia*

We now use numerical analysis to investigate the comovement between ETF premia. Note that if ETFs 1 and 2 are served by different LMMs, then there is no comovement between ETF premia, as indicated by Proposition 4. Therefore, we focus on the scenario where both ETFs are served by the same LMM.

Figure 2 shows the comovement parameter β as a function of c (the parameter representing the cost of arbitrage). The other parameter values are $A = 2$, $\nu = 1$, $M = 1$, $\mu = 0.2$, $\hat{K} = 25$, $N = 0.5$, $\hat{\mu} = 0.1$, and $\Phi = 0.05$.[17] The figure provides further support for the findings discussed earlier (Proposition 3 and Corollary 1). Specifically, if it costs little to arbitrage, indicated by a sufficiently low value of c, the comovement tends to approach zero (i.e., $\beta \to 0$). Further, as it becomes costlier to arbitrage (i.e., a higher c), the comovement between ETF premia becomes more pronounced (i.e., a higher β). The reason for this is that in this case, even a minor liquidity shock to ETF j significantly depletes the available capital that could be used to attenuate the premium of ETF i. This amplifies

[17] Our value for risk aversion, $A = 2$, is the same as that used in Leland (1992) and Holden and Subrahmanyam (2002). The values of the payoff volatility and the mass of stock investors (i.e., $\nu = 1$ and $M = 1$) are subjective. We set $N = 0.5$ to represent a significant mass of ETF traders, and $\mu = 0.2$ and $\hat{\mu} = 0.1$ to represent a substantial scale of noise trades in both the stock market and in the ETF market. We let $\hat{K} = 25$ so each ETF tracks a sufficiently large number of stocks, and $\Phi = 0.05$ so the LMM is not well-funded and has a limited capacity for trading. We have verified that our results are robust to a range of parameter choices.

the comovement in ETF premia.

In Figure 3, we consider the comovement parameter β as a function of ν (the parameter representing return volatility), fixing $c = 1$. It is evident that as return volatility increases, the comovement between ETF premia becomes more pronounced (i.e., a higher β). The reason for this is as follows. In our model, fundamental investors assist the LMM in attenuating the premium. As return volatility increases, risk-averse investors in a particular ETF (say j) become more cautious and absorb the liquidity shock in j less effectively. This amplifies the spillover effect from ETF j to ETF i, resulting in a higher degree of comovement between the premia.

Finally, in Figure 4 we show β as a function of Φ (the parameter representing the LMM's capital constraint), fixing $c = 1$ and $\nu = 1$. We find that as the LMM becomes more capital-constrained (i.e., as Φ decreases), the comovement between ETF premia becomes more pronounced. As in Figure 2, the reason is that liquidity shocks to ETF j materially affects the capital to attenuate the premium of ETF i. Consequently, comovement gets amplified.

2.4 *Empirical implications*

Based on our theoretical analyses, we formulate and test four main hypotheses. These, along with references to the supporting corollaries and/or figures, are provided below:

(i) There is positive comovement between the premia of ETFs that are served by the same LMM. By contrast, there is little comovement between the premia of ETFs that are served by different LMMs (Propositions 3 and 4).

Further, the comovement between the premia of ETFs that are served by the same LMM is more pronounced if

(ii) arbitrage is more costly (Corollary 1 and Figure 2),

(iii) constituent stocks have higher return volatilities ((Corollary 1 and Figure 3), and

(iv) the LMM has stronger capital constraints (Corollary 1 and Figure 4).

3 Data and Summary Statistics

In this section, we provide detailed descriptions of the data and the main variables. We also consider time series patterns in ETF premia.

3.1 *Data*

Our ETF LMM data are from ETF Global, which covers all ETFs listed in the U.S. with no survivorship bias. ETF Global offers detailed ETF data including the NAV, share price, shares outstanding, flows, bid/ask prices, volume, inception date, and LMMs of ETFs. Our sample period is from January 1, 2012 to December 31, 2020. We verify the data (and correct any data errors) on ETF prices, shares outstanding, and bid-ask spreads using data from CRSP security files. We confirm the ETF NAV information using CRSP mutual fund data. We exclude leveraged and inverse ETFs from our main sample, but later show that our main finding also obtains for these ETF types. After excluding ETFs with missing LMM information,[18] our final sample includes 3,848 ETFs with broad geographic coverage including Emerging Markets, Developed Markets, Asia-Pacific, Europe, Global Ex-U.S., Global, and North America. In terms of asset class coverage, around 70% of the ETFs are equity ETFs, with the remaining being commodity, currency, fixed income, real estate, and multi-assets.

Table IA.1 in the Internet Appendix lists the LMMs in our sample. There are 18 LMMs and their names match those provided by NYSE Arca.[19] Some LMMs in our sample are broker-dealers, such as Goldman Sachs and Credit Suisse, while others are market makers affiliated with hedge funds, such as Citadel Securities and Jane Street.

3.2 *Construction of variables*

Our main variable of interest is the premium of an ETF, calculated as (ETF Price − ETF NAV)/ETF NAV.[20] Since the absolute deviation of ETF price from its NAV, regardless of the direction, determines an LMM's arbitrage opportunities, we consider the absolute value of the ETF premium, raw $|Premium|$ in our analysis. To make sure the comovement in ETF raw $|Premium|$ is not simply driven by aggregate funding constraints, we orthogonalize each ETF's raw $|Premium|$ with respect to its non-LMM raw $|Premium|$, by estimating the following regression[21]:

$$\text{raw } |Premium|_{i,t} = b_0 + b_1 \text{ non-LMM raw } |Premium|_{i,t} + \epsilon_{i,t}, \tag{4}$$

[18] Observations with missing LMMs account for around 10% of the whole sample and are more prevalent after 2019 and for ETFs with a shorter history. Our robustness tests show that our key result is unlikely driven by missing LMMs in the ETF Global dataset.

[19] https://www.nyse.com/products/nyse-arca-market-making.

[20] We use end-of-day closing prices and NAVs of each ETF to calculate the premium.

[21] All our results continue to hold when we use the ETF raw premium as the variable of interest. These results are available upon request.

where non-LMM raw $|Premium|_{i,t}$ is the average raw $|Premium|$ across all ETFs managed by LMMs that are different from that of the focal ETF i. For each ETF, we use the full sample to estimate Equation (4) and take the regression residual $\epsilon_{i,t}$ as the main variable of interest, $|Premium|_{i,t}$. Essentially, we allow each ETF to have a differential exposure to market-wide factors that influence ETF premia.

In our empirical analyses, we control for several ETF characteristics.[22] $Log(Size)$ is the natural logarithm of an ETF's market capitalization. $Turnover$ is the daily dollar trading volume of an ETF scaled by its market capitalization (in percent), estimated using data from the prior 30 days. $BidAsk$ is the difference between ask and bid quotes scaled by the average of bid and ask quotes (in bps), estimated using data from the prior 30 days. STD is the standard deviation of daily ETF returns estimated using data from prior 30 days. Summary statistics for our main variables are in Table 1. The mean and standard deviations of ETF raw $|Premium|$ are 25.5 and 32 bps, respectively. By construction, ETF $|Premium|$ has a mean close to zero. The standard deviation of $|Premium|$ is large with a magnitude of 22.8 bps, suggesting that a large degree of variation in $|Premium|$ is not explained by market-wide factors that influence the premium. The last columns in Table IA.1 show the average ETF raw $|Premium|$ by LMMs. There is considerable cross-sectional variation in the average raw $|Premium|$ across different LMMs, ranging from the tightest 5 bps to 48.4 bps.

We also construct two measures to capture time-varying LMM-specific capital constraints. $Log(Mktcap\ of\ ETFs)$ is the natural logarithm of the total market capitalization of ETFs managed by the LMM. $\#Active\ AP$ is the number of active APs for the ETF in a given year, as reported in form N-CEN. To control for aggregate funding constraints, we include several macroeconomic variables. VIX is the CBOE volatility index; the credit spread (CS) is the difference between Moody's BAA yield and the yield on 10-year constant maturity Treasury bond; and HKM is the intermediary capital ratio of He, Kelly, and Manela (2017).

3.3 *Time series patterns*

We first investigate aggregate patterns in ETF premia. Panel A of Figure 5 shows the number and total size of ETFs (in billions USD) managed by an average LMM in our sample. The figure shows that on average, the total assets under management (AUM) of ETFs managed by an average LMM increased

[22] Appendix B provides a detailed description of main variables used in the paper.

from \$106 billion to \$226 billion from 2012 to 2020 (as indicated by the blue line). This aggregate trend suggests that LMM-managed ETF market cap is growing over time, and if LMM capital does not grow at the same pace, LMMs face tightening capital constraints. In Panel B of Figure 5, we plot the average raw and residual $|Premium|$ along with the CBOE Volatility Index (VIX). We find a strong comovement between VIX and the average raw $|Premium|$, with a correlation coefficient of around 0.6. Since the raw $|Premium|$ is a proxy for the expected returns from ETF arbitrage, the pattern is consistent with the notion that expected returns from liquidity provision and arbitrage opportunity increase with aggregate uncertainty (VIX). The time-series variation in raw $|Premium|$ is consistent with Nagel (2012), in which he shows that expected returns from liquidity provision in equity markets are highly predictable from the VIX. In contrast to the time-varying pattern of raw $|Premium|$, the green line plots the average residual ETF $|Premium|$ over time. It is clear from the figure that the residual $|Premium|$ is quite stable throughout the sample period, suggesting that the measure of residual $|Premium|$ mostly captures the idiosyncratic components of ETF premia.

4 Main Results on Comovements in ETF Premia

In Section 4.1, we conduct a baseline analysis investigating comovement in pricing efficiency for ETFs sharing the same LMM (implication (i) of Section 2.4). We supplement the baseline panel regression results with quantile regression analysis in Section 4.2 and event studies based on ETFs switching LMMs in Section 4.3.

4.1 *Baseline regression*

Our first test is to run panel regressions of each ETF's daily $|Premium|$ on the equally weighted average $|Premium|$ of all ETFs sharing the same LMM, controlling for a set of ETF characteristics that may affect the ETF $|Premium|$. This regression framework has been used to test excess return comovement among stocks sharing similar characteristics (Pirinsky and Wang (2006); Green and Hwang (2009)). The regression specification is as follows:

$$\begin{aligned} |Premium|_{i,j,t} &= \beta_0 + \beta_1 \text{ LMM } |Premium|_{i,t} + \beta_2 \text{ non-LMM } |Premium|_{i,t} \\ &+ \beta_3 X_{i,t} + \alpha_i + \gamma_{j,t} + \epsilon_{i,t}, \end{aligned} \tag{5}$$

where LMM $|Premium|_{i,t}$ is the average daily $|Premium|$ across all ETFs (excluding the focal ETF i itself) that share the same LMM as the focal ETF. In some specifications without time-fixed effects, in order to absorb any residual comovement due to market-wide factors, we also control for non-LMM $|Premium|_{i,t}$, which is the average $|Premium|$ of all ETFs served by an LMM that is different from that of the focal ETF. $X_{i,t}$ denotes a set of control variables, including ETF size ($Log(Size)$), ETF turnover ($Turnover$), ETF bid-ask spread ($BidAsk$), and ETF return volatility (STD). To facilitate comparison, we standardize all independent variables to have a mean of zero and a standard deviation of one. The observations are at the ETF-day levels. In most specifications, we also include ETF fixed effects (α_i) and $Style \times Day$ fixed effects ($\gamma_{j,t}$), where $Style$ denotes the detailed style category to which the ETF belongs. Note that the inclusion of $Style \times Day$ fixed effects absorbs any time-varying change in ETF premia at the style level, and hence also absorbs the non-LMM $|Premium|_{i,t}$. This helps address two endogeneity concerns. First, the LMM-level comovement in ETF premium could be driven by investors' correlated (time-varying) demand for ETFs belonging to the same investment style (Broman (2016)). Second, ETFs whose underlying constituents are in the same style might become more costly to arbitrage in certain periods.

Table 2 reports the results. T-statistics are based on standard errors double clustered at the ETF and Day levels. Columns (1) to (4) consider the raw $|Premium|$, while columns (5) to (8) the residual $|Premium|$. Across different specifications, the coefficients on LMM $|Premium|_{i,t}$ are significantly positive. For example, column (1) shows that the coefficient on LMM raw $|Premium|_{i,t}$ is 8.34 bps (t-stat. = 20.21), when estimated without any fixed effects. This suggests that a one-standard deviation increase in LMM raw $|Premium|_{i,t}$ is associated with an 8.34 bps increase in the focal ETF's raw $|Premium|$. By contrast, the coefficient on Non-LMM raw $|Premium|_{i,t}$ is less than one-eighth of that on LMM raw $|Premium|_{i,t}$ at 0.95 bps (t-stat. = 2.69). The last two rows of Table 2 show a significant difference between the coefficients of LMM raw $|Premium|_{i,t}$ and non-LMM raw $|Premium|_{i,t}$. We next add a set of ETF characteristics and the ETF fixed effects. The results in column (2) show that the coefficient of LMM raw $|Premium|_{i,t}$ decreases slightly to 6.85 bps (t-stat. = 29.52). When we include both the $Style \times Day$ and ETF fixed effects, the results in column (4) show that the coefficient of LMM raw $|Premium|_{i,t}$ is 1.94 (t-stat. = 10.65). The declining pattern in the coefficient estimates of LMM raw $|Premium|_{i,t}$ suggests that a non-trivial part of the comovement in ETF premia is driven by market-wide factors. One caveat is that part of the market-wide factors could be driven by LMMs' systematic capital constraints. Hence, our estimate in column (4) provides a lower bound for the effect of LMM-specific capital constraints on the comovement in ETF premia.

Focusing on the coefficients on the control variables, we find that the estimates are consistent with theories of limits to arbitrage (Shleifer and Vishny (1997)). For example, the negative coefficient on $Log(Size)$ in column (4) suggests that larger ETFs have lower premia, potentially because there are more arbitrageurs in the secondary market for larger ETFs. The positive coefficient on the bid-ask spread ($BidAsk$) indicates that ETF premia are greater for ETFs with lower liquidity, consistent with the evidence in Bae and Kim (2020). Similarly, the positive coefficient on STD is consistent with the notion that, when the ETF return is more volatile, it is more costly for arbitrageurs to take large arbitrage positions. As a result, the premia are higher for such ETFs.

We next run the same regressions using the residual $|Premium|$, which further accounts for each ETF's differential exposure to market-wide influences on ETF premia. Across all specifications, we find in columns (5) to (8) that the coefficients on LMM $|Premium|_{i,t}$ are positive and highly significant, while that on non-LMM $|Premium|_{i,t}$ becomes insignificant. Importantly, since the residual $|Premium|_{i,t}$ already removes the effects of market-wide influences, the coefficient estimates of LMM $|Premium|_{i,t}$ are quite stable across different specifications, with estimated coefficients ranging from 1.58 to 2.12 bps. In terms of economic magnitude, when both the ETF and $Style{\times}Day$ fixed effects are included, the coefficient on LMM $|Premium|_{i,t}$ in column (8) suggests that a one standard deviation increase in LMM $|Premium|_{i,t}$ is associated with a 1.59 bps increase in the focal ETF's residual $|Premium|$. Since the standard deviation of residual $|Premium|$ is 22.8 bps, a 1.59 bps is equivalent to 7.0% of its standard deviation. Given that for our sample, the annual dollar trading volume of all ETFs managed by an average LMM is around $528 billion, a one-standard-deviation decrease in LMM $|Premium|$ results in an annual dollar savings of $84.0 million for investors who trade ETFs managed by the LMM on opportune days, and vice versa.

Overall, the baseline results are consistent with our hypothesis that there is a strong comovement in the premia of ETFs served by the same LMM. Since the residual $|Premium|$ mainly captures the idiosyncratic component of ETF premia, we focus on the residual $|Premium|$ as the variable of interest in the subsequent analyses to provide insight on the importance of LMM-specific capital constraints. All the empirical results are robust when estimated using the raw $|Premium|$, and are sometimes even stronger than the results based on the residual $|Premium|$.

4.2 *Tests of contagion in the right tails of ETF premia*

The results in the previous subsection indicate that there is excess comovement in the $|Premium|$ of ETFs serviced by the same LMM. We next investigate whether ETF $|Premium|$ exhibit excess correlations in the right tails of their distributions. The rationale for this test is that when the average $|Premium|$ of ETFs managed by the LMM is very large, the LMM likely faces severe capital constraints, which increases the chance of observing very high premia for the focal ETFs. We use two different approaches for the investigation. The first approach uses quantile regressions, which make it possible to estimate the probability of a given ETF $|Premium|$ falling in a particular range, conditional on the $|Premium|$ of all other ETFs sharing the same LMM. The second approach uses a linear probability estimation. The two approaches provide consistent results.

4.2.1 *Co-dependence in ETF and LMM premia – a quantile regression approach*

We first visually show the existence of contagion in ETF premia using a "comovement box" approach as in Boyson, Stahel, and Stulz (2010). Their quantile regression estimates the conditional probability that a random variable y falls below (or above) a given quantile conditional on a different random variable x also falling below (or above) the same quantile. The estimated co-dependence is plotted in a box, which is a square of unit side with the conditional probabilities plotted against the quantiles. When the plot of the conditional probability lies above (below) the 45-degree line, which represents the unconditional probability of no dependence between the variables, there is evidence of positive (negative) conditional comovement between x and y. In our analysis, y represents the residual $|Premium|$ of a focal ETF, while x is the equally-weighted average residual $|Premium|$ on all other ETFs managed by the same LMM (LMM $|Premium|$). Results from the analysis are presented in Figure 6. To construct the figure, we calculate the probabilities in 5% increments between the 5th and 95th quantiles. The blue line denotes the 45-degree line, and the red line denotes the actual conditional probability distribution estimated using the residual ETF $|Premium|$.

The figure shows that the plot of the conditional probability always lies above the 45-degree line. It provides strong evidence that the conditional probability of an ETF having a $|Premium|$ below (or above) any quantile is increased significantly when the LMM $|Premium|$ is also below (or above) the same quantile. Importantly, the difference between the two plots of probability estimates is larger at the right tails of the distribution. For example, at the 90th quantile, the focal ETF's $|Premium|$ has a

16.6% probability of being above the 90th quantile when the LMM $|Premium|$ is also above the 90th quantile, compared to an unconditional probability of 10% if there were no dependence.

4.2.2 *Tests of contagion in the right tails of ETF premia using a linear probability model*

The previous results provide evidence of contagion in the ETF $|Premium|$, especially at the right tails of the distributions. In this subsection, we use a linear model for the probability of observing an extremely large ETF $|Premium|$ conditional on whether its LMM $|Premium|$ also experiences an extremely large $|Premium|$.

To conduct this test, we first use a 90% cutoff of the overall distribution of $|Premium|$ to identify extreme levels of ETF premia. The dependent variable is a dummy variable that equals one if the focal ETF's $|Premium|$ is higher than the 90th percentile cutoff, and zero otherwise. To measure the extent of clustering of extreme ETF premium at the LMM level, we add an indicator variable, $D(LMM\ |Premium| > 90th\ percentile)$, that equals one if the LMM $|Premium|$ is above its 90th percentile cutoff, and zero otherwise. We include the same set of control variables as in Equation (5), and use a linear probability model (instead of a logit model) for estimation, in order to accommodate various fixed effects. Results are reported in Table 3. The coefficients on $D(LMM\ |Premium| > 90th\ percentile)$ are always positive and statistically significant, providing strong evidence that extreme premia cluster in ETFs sharing the same LMM. Economically, the coefficient estimates suggest that the probability of a focal ETF's $|Premium|$ being in the top decile increases by an additional 2.4% to 5.7% when its LMM $|Premium|$ is also in the top decile. These results are consistent with the comovement box in Figure 6.

4.3 *Identification based on ETFs switching LMM*

Our panel regression results show a strong comovement in the idiosyncratic component of ETF $|Premium|$ among ETFs sharing the same LMM. One might be concerned, however, that the comovement in ETF premia is driven by self-selection of LMMs. That is, LMMs select the list of ETFs to make markets based on some unobservable (to an econometrician) ETF characteristics, and these ETF characteristics may lead to comovement in ETF premia due to correlated investor demand or time-varying arbitrage costs. To show that LMMs play a causal role in the comovement of ETF premia, we conduct event studies around the days when ETFs change their LMMs. A significant fraction of these LMM change events are due to mergers between two LMMs. For example, Virtu

Financial acquired KCG Holdings in July 2017, which accounts for 37% of the LMM switching events in our sample. Thus, we can reasonably assume that a change of LMM for an *individual* ETF is relatively exogenous to the ETF's unobserved characteristics that drive comovement in ETF premia.

We identify 1,264 events where an ETF changed its LMM. We choose a window of $[-120, 120]$ trading days, with day 0 as the date on which the ETF changed its LMM. We then regress the residual $|Premium|$ on the average residual $|Premium|$ of ETFs that are managed by the focal ETF's previous and new LMMs. In Figure 7, we plot the regression coefficients of LMM (raw) $|Premium|_{i,t}$ around the event days, where the coefficient for each event day is estimated using the $[-3, 3]$ trading day window surrounding it. The upper graph in the figure shows the regression coefficients estimated using the raw $|Premium|$ and the lower graph shows the estimations for the residual $|Premium|$. The red line indicates the coefficient of the previous LMM (raw) $|Premium|_{i,t}$ while the blue line indicates the coefficient of the new LMM (raw) $|Premium|_{i,t}$. The figure clearly shows that, after an ETF changes its LMM, its premium comoves to a lesser extent with that of ETFs managed by the previous LMM, while the premium becomes more correlated with those of ETFs managed by the new LMM.

Next, we confirm this pattern in formal regressions using the specification below:

$$\begin{aligned} |Premium|_{i,j,t} &= \beta_0 + \beta_1\, LMM_{Old}\, |Premium|_{i,t} + \beta_2\, Post_t \times LMM_{Old}\, |Premium|_{i,t} \\ &+ \beta_3\, LMM_{New}\, |Premium|_{i,t} + \beta_4\, Post_t \times LMM_{New}\, |Premium|_{i,t} \\ &+ \beta_5\, Post_t + \beta_6\, X_{i,t} + \epsilon_{i,t}, \end{aligned} \tag{6}$$

where $LMM_{Old}\ |Premium|_{i,t}$ ($LMM_{New}\ |Premium|_{i,t}$) is the average residual $|Premium|$ of ETFs managed by the old (new) LMM before (after) switching. $Post_t$ is a dummy variable that equals one for the days after an ETF changes its LMM. $X_{i,t}$ denotes the same set of ETF-level controls as those in Equation (5). In some specifications, we also include macroeconomic factors, returns on the Fama-French five factors (Fama and French (2015)) and the ten Fama-French industry portfolios to control for correlated demand shocks to ETFs belonging to the same style or sector (Wahal and Yavuz (2013)).

Table 4 reports the results. Consistent with our predictions, we find that the coefficients on $Post \times LMM_{Old}\ |Premium|_{i,t}$ are negative and significant across all specifications. Column (1) shows that the coefficient on $LMM_{Old}\ |Premium|_{i,t}$ is 1.20, while the coefficient on $Post \times LMM_{Old}\ |Premium|_{i,t}$

is -0.92 (t-stat. = -3.48). The economic magnitude suggests that the comovement in the premium with those of other ETFs served by the previous LMM reduces by around 80% after the ETF switches to a new LMM. On the other hand, we find the coefficients on $Post \times LMM_{New}\ |Premium|_{i,t}$ are positive and significant across all specifications, suggesting that the ETF premium becomes more closely correlated with those of other ETFs served by the new LMM after switching. Importantly, we find that the coefficients on $LMM_{New}\ |Premium|_{i,t}$ are statistically insignificant across all specifications, suggesting that the comovement in premia is unlikely driven by the self-selection effects of LMMs.[23] Column (1) of Table 4 shows that the coefficient on $LMM_{New}\ |Premium|_{i,t}$ is 0.288 (t-stat. = 1.42), while the coefficient on $Post \times LMM_{New}\ |Premium|_{i,t}$ is 1.44 (t-stat. = 2.84). Table IA.2 in the Internet Appendix shows that the results are similar when we exclude LMM switching events due to the acquisition of KCG by Virtu.

Overall, the absence of comovement *before* the LMM switch and the presence of strong comovement *after* the switch between the focal ETF's premium and that of the new LMM show that the excess comovement in ETF premia is indeed driven by these ETFs sharing the same LMM.

5 Cross-Sectional Heterogeneity

In this section, we examine the hypothesis that LMM capital constraints have a greater impact on comovement for ETFs that are more costly to arbitrage (implication (ii) in Section 2.4). To test this hypothesis, we use specific characteristics of ETFs and, in turn, their constituents as proxies for arbitrage costs in Sections 5.1 and 5.2, respectively.

5.1 *Arbitrage costs of ETFs*

When an LMM acts to attenuate the premium in an ETF that is more costly to arbitrage, the ETF demands a greater capital commitment per unit of attenuation in the premium, resulting in a larger capital withdrawal from other ETFs. As a result, we should expect to find a stronger comovement effect for ETFs that are more costly to arbitrage. We use the market capitalization ($Log(size)$) and liquidity ($BidAsk$) of the ETF to capture arbitrage costs (Pontiff (1996); Gromb and Vayanos (2018)). Our theory suggests that when return volatility (STD) is higher, risk-averse fundamental investors absorb less liquidity shocks; this increases the comovement in the ETF premia that the LMM needs

[23] If the results are driven by the selection of LMMs based on unobservable ETF characteristics, we should observe significant comovement between the premium of the focal ETF and those of other ETFs served by the new LMM prior to the actual switching date.

to correct (implication (iii) in Section 2.4). We therefore include return volatility in our analysis.

Columns (1) to (3) of Table 5 report the results when we interact LMM $|Premium|$ with the relevant ETF characteristics. The results are consistent with our conjecture. Column (1) shows that the comovement is weaker for ETFs with larger market capitalization. Columns (2) and (3) show that the effect is more pronounced for ETFs with higher return volatility and higher bid-ask spread, respectively. The economic effect is also non-trivial. Taking the bid-ask spread as an example, column (3) shows that for a one-standard-deviation increase in an ETF's bid-ask spread, the impact of LMM $|Premium|$ on its own $|Premium|$ is 16.6% greater.

5.2 *Arbitrage costs of ETFs' underlying constituents*

Since ETF arbitrage requires LMMs (and other arbitrageurs) to take positions in both the ETF and its underlying basket securities,[24] another cross-sectional prediction is that the comovement effect should be stronger when the ETFs' underlying assets are, on average, more costly to arbitrage. To test this hypothesis, we focus on ETFs with US equity as underlying assets, for which we can access standard metrics for arbitrage. We use three measures of such costs: the bid-ask spread ($Spread\ CS$), stock return volatility ($Volatility$), and lendable supply ($Supply$). We construct stock-level bid-ask spreads following the approach of Corwin and Schultz (2012).[25] We obtain stock lendable supply (lendable shares divided by total shares outstanding) from the Markit Securities Finance (formerly Data Explorer) database. Both a higher bid-ask spread and higher return volatility indicate more severe arbitrage frictions, while a greater lendable supply in the securities lending market indicates less-constrained short selling. We first aggregate the stock-level arbitrage cost measures to the ETF level, and then interact these measures with LMM $|Premium|_{i,t}$, to test the incremental effect of arbitrage frictions on comovement.

Columns (4) to (6) of Table 5 reports the results. Consistent with our hypothesis, columns (4) and (5) show that the interactions between LMM $|Premium|_{i,t}$ and $Spread\ CS$ and $Volatility$ are significantly positive, and column (6) indicates that the interaction between LMM $|Premium|_{i,t}$ and

[24] LMMs need to create (redeem) shares of an ETF and simultaneously enter into an opposite direction of trades for the underlying constituents when the ETF is traded at a premium (discount).

[25] The Corwin and Schultz (2012) spread estimate is based on two reasonable assumptions. First, daily high-prices are almost always buyer-initiated trades and daily low-prices are almost always seller-initiated trades. The ratio of high to low prices for a day therefore reflects both the fundamental volatility of the asset and its bid-ask spread. Second, the component of the high-to-low price ratio that is due to volatility increases proportionately with the length of the trading interval while the component due to bid-ask spreads do not. Corwin and Schultz (2012) show via simulations that, under realistic conditions, the correlation between their spread estimates and true spreads is about 0.9.

$Supply$ is significantly negative. The economic effect is also meaningful. Taking bid-ask spread as an example, column (4) shows that for a one-standard deviation increase in the average bid-ask spread of an ETF's underlying stocks, the impact of LMM $|Premium|$ on ETF $|Premium|$ is 43% greater. These results support our hypothesis that the comovement in pricing efficiency is more pronounced when the LMM faces higher costs in taking arbitrage positions in an ETF's underlying assets.

6 The Role of LMM-Specific Capital Constraints

In this section, we test implication (iv) of Section 2.4 by examining comovement conditional on LMM-specific capital constraints. Intuitively, when an LMM faces limited arbitrage capital, the pricing gap in one ETF managed by an LMM can spill over to pricing gaps in other ETFs for which the LMM is responsible. Hence, we expect LMM-specific capital constraints to impact comovement in ETF premia. In Sections 6.1 and 6.2, we consider capital constraints based on specific ETF and LMM characteristics, respectively. In Section 6.3, we conduct a DiD analysis of ETF premium comovements around the COVID-19 market sell-off, on the basis that due to the debt market disruptions around the pandemic, debt-based ETFs suffered increased capital constraints, which would imply greater spillovers for such ETFs.

6.1 *LMM-specific capital constraints based on ETF characteristics*

We first use two variables to capture LMM-specific capital constraints based on the characteristics of ETFs managed by the LMM. Our first measure, $Log(Mktcap\ of\ ETFs)$, is the natural logarithm of the total market capitalization of all ETFs managed by the LMM. The idea is intuitive: if the LMM needs to simultaneously arbitrage ETFs with larger total market capitalization, then it has less capital devoted to attenuating pricing gaps for each individual ETF.[26]

Our second measure is the number of active APs for each ETF in a year. In addition to LMMs, APs also play an important role in maintaining the law of one price for ETFs. We would expect fewer APs to imply more stringent capital constraints faced by the LMM. To construct this measure, we collect information on ETFs' active APs from SEC N-CEN filings. We create a variable, $Log(1/\#Active\ APs)$, calculated as the natural logarithm of one divided by the number of active APs, constructed using filings data from the last fiscal year.

[26]The total market capitalization of ETFs served by the LMM can also be viewed as a proxy for the LMM's (in)attention. However, most ETFs are traded on electronic exchanges, such as NYSE Arca, with LMMs adopting algorithmic trading for ETFs. For these reasons, attention constraint is unlikely the major reason for the comovement effect we find.

We interact each of the above measures with LMM $|Premium|_{i,t}$ to estimate the incremental effect of LMM-specific capital constraints on comovement in ETF pricing efficiency. ETF fixed effects and $Style{\times}Day$ fixed effects are included in all the regression specifications. Table 6 reports the results. Consistent with our hypothesis, we find that the interaction terms are significantly positive for both measures of LMM capital constraints. For example, column (1) reports that the estimated coefficient on the interaction between $Log(Mktcap\ of\ ETFs)$ and LMM $|Premium|_{i,t}$ is 0.319 (t-stat. = 8.62). The economic magnitude indicates that, for a one-standard-deviation increase in the $Log(Mktcap\ of\ ETFs)$ of an LMM, the impact of LMM $|Premium|$ on the focal ETF's $|Premium|$ is 18.5% greater.

6.2 *LMM-specific capital constraints based on LMM characteristics*

Our previous analysis uses ETF characteristics to capture LMM-specific capital constraints, because most LMMs are private companies with little information about their capital positions. For a small sample of visible, publicly-listed LMMs, however, we are able to measure their capital constraints directly using publicly available information, in two ways. First, we use changes in individual LMMs' net worth during a short window around earnings announcements to capture negative shocks to their capital. As argued by Ottonello and Song (2022), negative earnings news leads to a discontinuity in the equity value of intermediaries that is difficult to anticipate and is less confounded by omitted variables. Specifically, we define the LMM-specific capital constraint as a dummy variable that equals one for the three-day window [-1, 1] around the LMM's quarterly earnings announcements when the market-adjusted announcement abnormal return is less than -3%, and zero otherwise. The sample is restricted to the 21-day window [-10, 10] around quarterly earnings announcements for all publicly traded LMMs, including Credit Suisse, Deutsche Bank, Goldman Sachs, and Virtu Financial.

Second, we measure an LMM's capital constraint by computing its net capital ratio (*Net Capital Ratio = Net Capital Required/Adjusted Net Capital*). We are able to gather capital positions for five LMMs in our sample, using data from the Commodity Futures Trading Commission (CFTC).[27] This ratio is closely monitored by the CFTC because it measures the amount of capital a market maker holds (the denominator) relative to the CFTC's net capital requirement (the numerator). We expect the ratio to be larger for more constrained LMMs.

[27] The five LMMs are Cantor Fitzgerald, Credit Suisse, Deutsche Bank, Goldman Sachs, and RBC Capital Markets.

We interact these two measures of capital constraints with LMM $|Premium|_{i,t}$ and report the results in Table 7. In columns (1) to (3), the LMM's capital constraint is measured by the dummy indicating its earnings announcement return is less than -3%. We find that the interaction terms are all significantly positive, indicating a stronger comovement effect for ETFs managed by more constrained LMMs. The economic effect is also meaningful. For example, column (3) reports that the estimated coefficient on the interaction between LMM $|Premium|_{i,t}$ and $Constraint$ is 0.487 (t-stat. = 1.95). This suggests that the impact of LMM $|Premium|$ on the focal ETF's $|Premium|$ is 35.7% greater when the LMM experiences negative shocks to its capital.[28] In columns (4) to (6), we use the *Net Capital Ratio* to capture LMM-specific capital constraints and find similar results. Overall, the results support our premise; specifically, the comovement is more pronounced when LMM-specific capital constraints are more binding.

6.3 *DiD analysis of the ETF premium during the COVID-19 pandemic*

We next conduct a difference-in-differences (DiD) estimation around the COVID-19 market sell-off in 2020, in order to examine whether intermediary capital constraints amplify comovement in ETF premia. During that period, the ETF market experienced unprecedented pricing gaps, especially for fixed income ETFs. In Panel A of Figure 8, we plot the average raw $|Premium|$ for ETFs tracking different asset classes from January 2020 to June 2020. The shaded area indicates the period when COVID-19 caused significant financial market turmoil, which runs from February 20, 2020, to April 30, 2020, following Pástor and Vorsatz (2020). As the figure shows, the average absolute premium for all types of ETFs widened dramatically during the crisis period, with the effect being most pronounced for fixed income ETFs. The premium for fixed income ETFs increases from 14.9 bps on February 1, 2020 to 156.7 bps at the peak of the crisis on March 20, 2020. This is consistent with recent studies documenting a significant disruption to the fixed income market during the COVID-19 pandemic (Falato, Goldstein, and Hortaçsu (2021); Haddad and Muir (2021)). In Panel B of Figure 8, we show that the widening pricing gap mainly manifested as a discount (i.e., the prices of ETFs traded below their NAV), potentially because ETFs are the asset type that investors chose to liquidate first in the cash crunch, due to their superior liquidity and trading convenience.

Our DiD test exploits the fact that fixed income ETFs experienced the largest $|Premium|$ during

[28] We also conduct a test using positive earnings announcement returns as placebo events. Untabulated results show no significant increase in the LMM-level comovement in ETF premia when an LMM experiences a positive change in its capital.

the COVID-19 pandemic. The idea is that LMMs who need to manage a larger fraction of fixed income ETFs likely face more binding capital constraints during the pandemic. Our arguments predict that non-fixed income ETFs managed by more constrained LMMs should experience greater pricing gaps, compared to non-fixed income ETFs that are managed by less constrained LMMs.[29]

The advantage of the DiD setting, of course, is that the COVID-19 pandemic is an exogenous shock that originates outside of the financial sector. As a result, LMMs are unlikely to anticipate the widening ETF premium during this period, which ensures a close-to-random assignment across the more and less constrained LMMs. We conduct the DiD estimation using the following specification:

$$\text{raw } |Premium|_{i,j,t} = \beta_0 + \beta_1 COVID_t + \beta_2 X_i \times COVID_t + \beta_3 Control_{i,t} + \alpha_i + \epsilon_{i,t}, \tag{7}$$

where $COVID_t$ is a dummy variable indicating the post-treatment period, which equals one for the period from February 21, 2020 to April 30, 2020, and zero otherwise. The variable X_i denotes two proxies capturing the LMM's fixed income exposure. The first proxy, $FI\ Weight_i$, is a continuous treatment variable defined at the ETF level, and is calculated as the market capitalization of fixed income ETFs managed by the focal ETF's LMM scaled by the total market capitalization of all ETFs managed by the LMM. Importantly, we measure $FI\ Weight_i$ at the end of 2019 (i.e., before the start of the COVID-19 pandemic). Our second proxy is a dummy variable, $D(FI\ Weight > Median)$, that defines the treatment sample. $D(FI\ Weight > Median)$ is an indicator that equals one if the LMM's $FI\ Weight_i$ is above the sample median, and zero otherwise. The coefficient of interest is the interaction between $FI\ Weight_i$ (or $D(FI\ Weight > Median)$) and $COVID_t$, which captures the spillover effect of LMMs on the $|Premium|$ of non-fixed income ETFs, due to LMMs managing fixed income ETFs as part of their portfolios. $Control_{i,t}$ denotes the same set of control variables as in the baseline regression in Equation (5). We also control for ETF fixed effects (α_i) in all specifications, which subsume the effect of $FI\ Weight_i$.

Our sample period spans January 1, 2020, to June 30, 2020, and Table 8 reports the DiD results. In the regression for column (1), we only include the $COVID_t$ dummy, which has a coefficient of 21.01 (t-stat. = 7.92). This is consistent with Figure 8, where ETFs on average experienced widening pricing

[29] It is possible that the large premium for fixed-income ETFs partially reflect significant differences in liquidity between the ETF and underlying bonds. However, our prediction for the commonality in ETF premium should still hold, since Coughenour and Saad (2004) provide evidence of commonality in liquidity among stocks handled by the same NYSE specialist firm.

gaps during the market sell-off. We next add the treatment variable $FI\ Weight_i$ and its interaction with $COVID_t$. Column (2) shows that coefficient of $FI\ Weight_i$*$COVID_t$ is significantly positive, consistent with our prediction. In column (3), we find similar results after including control variables and $Style{\times}Day$ fixed effects in the regression, with the latter absorbing the $COVID_t$ dummy. The economic effect is also meaningful. For example, the estimated coefficient of $FI\ Weight_i$*$COVID_t$ in Column (3) is 4.88 (t-stat. = 1.80). The economic magnitude suggests that for a non-fixed income ETF managed by an LMM with a 75% weight in fixed income ETFs, the increase in its $|Premium|$ during the sample period is 2.44 bps higher than ETFs managed by an LMM with only 25% in fixed income ETFs.

We next conduct a dynamic effect analysis to assess the parallel trend assumption required for DiD estimation. Specifically, we create time dummies indicating half-month periods around the COVID-19 outbreak. For example, $COVID_{t-15,t-1}$ is a dummy variable that equals one from 15 days before February 20, 2020 to the start of the outbreak on February 20, 2020, and zero otherwise. $COVID_{t+1,t+15}$ and $COVID_{t+16,t+31}$ are defined similarly. $COVID_{t+32,End}$ equals one for the period from 32 days after February 20, 2020 to the end of the first acute phase of the COVID-19 pandemic on April 30, 2020. The variables of interest are the interactions of fixed income exposure with these COVID time dummies. The results in column (4) show that the effect of being managed by an LMM highly exposed to fixed-income ETFs on $|Premium|$ only becomes significant in the post-COVID-onset period, and is statistically insignificant before the onset of COVID. Columns (5) to (7) report similar results when we use the dummy variable $D(FI\ Weight > Median)$ to define the treatment sample, as the coefficients of the interaction between $D(FI\ Weight > Median)$ and $COVID_t$ are positive and statistically significant and the effect only manifests in the period following the onset of COVID.

In sum, the DiD test indicates that negative shocks to LMMs' capital constraints causally lead to increased ETF pricing gaps. From a policy perspective, the result suggests that inefficiencies in one segment of the ETF market can potentially spillover to other segments through the common LMM linkage.

7 Alternative Explanations and Robustness tests

In this section, we conduct tests to rule out alternative explanations for the comovement effect. Section 7.1 provides further evidence that the impact of LMM-specific capital constraints is not explained by aggregate funding constraints. In Section 7.2, we test the premium comovement effect separately for ETFs with the same or different benchmark/style as the focal ETF. In Section 7.3, we re-estimate the baseline regression by inserting alternative sets of fixed effects, which help absorb premium comovement due to ETFs sharing other similar characteristics. In Sections 7.4 and 7.5, we conduct our baseline tests separately for ETFs covering different regions and tracking different asset classes, respectively. In Section 7.6, we conduct some additional miscellaneous tests to rule out confounding effects.

7.1 *The impacts of aggregate funding constraints*

Our evidence in Section 4 indicates that LMMs play a key role in driving the comovement in premia for ETFs under their umbrella. However, it is possible that LMMs face more severe capital constraints when aggregate funding constraints tighten. Accordingly, in this subsection we conduct subperiod analysis controlling for measures of aggregate funding constraints.

We use the VIX index, the credit spread (CS), the intermediary capital ratio of He, Kelly, and Manela (2017) (HKM), and $Mktret$, the stock market return in the prior month (Hameed, Kang, and Viswanathan (2010)) as proxies for aggregate funding constraints. Higher values of VIX and CS, and lower values of HKM and $Mktret$, indicate tightened constraints. We divide the sample into halves based on each of these four measures. We then conduct the baseline regression in Equation (5) for the two subperiods, with High (Low) indicating periods with tightened (loosened) aggregate funding constraints and Up (Down) denoting market states with monthly returns above (below) the sample median.

Table 9 shows that the coefficients on LMM $|Premium|_{i,t}$ are positive and significant with similar economic magnitudes in both periods. For example, columns (1) and (2) show that the coefficients on LMM $|Premium|_{i,t}$ are 1.517 (t-stat. = 15.51) and 1.618 (t-stat. = 16.48) in subperiods with a low and high VIX, respectively. The pattern is similar when we use the credit spread (CS), the intermediary capital ratio (HKM), and the last month's market return ($Mktret$) as proxies for aggregate funding constraints. Overall, the results suggest that the role of LMM-specific capital constraints in driving

ETF premia comovement is independent from that of aggregate funding constraints.

7.2 *Correlated demand shocks among ETFs tracking the same benchmark/style*

Our baseline specification includes $Style \times Day$ fixed effects, which help absorb premium comovement due to investors' correlated demand for ETFs tracking the same investment styles. To further address this issue, we construct the non-LMM $|Premium|$ for ETFs following the same benchmark as the focal ETF. Column (1) of Table 10 reports an insignificant coefficient on Non-LMM $|Premium|$ constructed using ETFs tracking the same benchmark as the focal ETF.[30] The lack of premium comovement among ETFs tracking the same benchmark help identify the unique role played by LMMs in driving the comovement effect. Next, we further decompose the LMM $|Premium|$ into two parts: one constructed using ETFs with the same style and another using ETFs with a different style as the focal ETF. The results in column (2) show that the LMM premium comovement effect is larger for ETFs with a style different that is from that of the focal ETF. Overall, the evidence indicates that our results are not due to investors' correlated (time-varying) demand for ETFs belonging to the investment style, or tracking the same benchmark.

7.3 *Correlated demand shocks for ETFs similar in other dimensions*

Investors may have correlated demand for ETFs sharing attributes beyond style or tracking benchmarks. To address this concern, we add alternative sets of fixed effects in the baseline specification of Equation (5) and report the results in Table 11. In column (1), we control for $Region \times Day$ fixed effects, where $Region$ refers to the geographical focus of the ETF. This also helps rule out an alternative explanation that the premium comovement effect arises from the time zone differences between ETFs' trading venues and those of their underlying assets. In the regression for column (2), we include $Exchange \times Day$ fixed effects, where $Exchange$ denotes the stock exchange in which the ETF is listed. In the regression for columns (3) and (4), we include $Issuer \times Day$ and $Distributor \times Day$ fixed effects, respectively, where $Issuer$ and $Distributor$ refer to the issuer and distributor of the ETF. We also control for ETF fixed effects and the same set of ETF-level characteristics as in Equation (5). Across all specifications, the coefficients of LMM $|Premium|$ remain positive and highly significant. Thus, overall, the LMM-level comovement in premia is robust to alternative fixed effects.

[30]Since there are not many ETFs that track exactly the same benchmark, the sample size reduces dramatically for this test.

7.4 *ETFs with different regional coverage*

We further examine ETF premium comovement within regions. As reported in Table 12, comovement in ETF premia is pervasive across regions, with the estimated coefficients on LMM $|Premium|$ ranging from 0.38 for North America to 3.56 for the Asia-Pacific region. The economic magnitude of the estimated coefficients is consistent with the notion that pricing efficiency comovement is higher for ETFs that are more costly to arbitrage. In particular, columns (1) and (2) show that the comovement for emerging markets ETFs is 19% higher than that for the developed markets ETFs. Columns (3) to (6) show that ETFs with the highest level of comovement cover the Asia-Pacific region, followed by those covering Europe and those excluding the U.S. Not surprisingly, North American ETFs have the lowest degree of premium comovement. In untabulated results, we find that the average level of the (absolute) premium is also the highest for Asia-Pacific and Emerging Markets ETFs, and is the lowest for the North America. The lower comovement in the premium for North American ETFs may be due to the existence of many (non-LMM) arbitrageurs in this ETF segment, with the attenuation of the premium being less reliant on LMMs.

7.5 *ETFs tracking different asset classes*

We next consider premium comovement for ETFs tracking different asset categories, including equities, fixed income securities, real estate, commodities, currencies, and multiple asset classes. In Table 13, we find that the LMM-level comovement in ETF residual $|Premium|$ is significant for all asset categories except currencies. The coefficients of LMM $|Premium|$ range from 0.687 for currencies to 2.16 for multi-assets. We also conduct the comovement test for the subsample of leveraged and inverse ETFs.[31] Column (7) shows that the coefficient of the LMM $|Premium|$ is 1.16 (t-stat. = 2.16), which is slightly smaller than that observed for our main sample of ETFs. Overall, however, the findings suggest that the notion of LMMs' constraints influencing comovement in ETF pricing efficiencies holds across asset classes.

7.6 *Other robustness tests*

We present several additional robustness tests in the Internet Appendix. First, in all of our previous analyses, we exclude ETFs with missing LMMs at the very beginning of the data cleaning stage.

[31] Lu and Qin (2021) infer the market-wide shadow cost of leverage from the return shortfall between leveraged funds' daily gross NAV returns and the target multiple of the underlying index returns. Our paper focuses on the difference between ETF prices and NAVs, which mainly captures pricing efficiency.

Observations with missing LMMs account for around 10% of the whole sample and are much more prevalent after 2019 and for ETFs with a shorter history (less than two years old). To show that excluding ETFs with missing LMMs does not systematically bias our results, in Table IA.3, we conduct subsample tests as in Table 2, using observations up to the year 2018 and for ETFs with a history longer than 24 months. Our results are unaltered.

Second, we use the signed ETF premium to examine whether our comovement result is symmetric across positive and negative premia. Specifically, we define *Positive (Negative) Premium* as the absolute value of the ETF premium when the sign of the premium is positive (negative), and zero otherwise. We then regress the *Positive (Negative) Premium* of the focal ETF on the corresponding *LMM Positive (Negative) Premium*. *LMM Positive (Negative) Premium* is calculated as the equally-weighted average *Positive (Negative) Premium* of all ETFs sharing the same lead market maker, excluding the focal ETF. Table IA.4 shows that the LMM-level comovement is similar in magnitude for both positive and negative premiums. For example, column (4) shows that with both ETF and $Style \times Day$ fixed effects included, the coefficient of *LMM Positive Premium* is 1.099, while the corresponding coefficient of *LMM Negative Premium* in column (8) is 0.823.

Third, we control for the average daily LMM (raw or residual) $|Premium|$ in the prior month, denoted as $|Premium|_{m-1}$. Because naïve traders could chase past returns and drive ETF premia, we also include past ETF returns in the regressions as a control. Specifically, Ret_{m-1} and $Ret_{m-2,m-12}$ represent the ETF's return in the prior month and the cumulative return from the past two to 12 months, respectively. Table IA.5 shows that the coefficients of LMM $|Premium|$ are still positive and highly significant, for both the raw and residual $|Premium|$.

Finally, we conduct the baseline test of Equation (5) using the value-weighted versions of LMM and non-LMM $|Premium|$, where the weights are given by the total NAV of ETFs as of the end of the previous month. Table IA.6 shows that the coefficients of LMM $|Premium|$ are still positive and highly significant, for both the raw and residual $|Premium|$. Furthermore, the economic magnitudes of the coefficients on the value-weighted LMM premia are similar to the estimates in Table 2, indicating that our results are robust to using the value-weighted $|Premium|$.

8 Conclusion

How do financial intermediaries affect market efficiency in the assets they manage? In this paper, we develop a simple model to show that if ETFs share a common lead market maker (LMM), their pricing differentials (relative to their constituents) exhibit higher comovement than if they do not. The economic reasoning is that LMM capital constraints cause contagion in the premia across assets managed by a common intermediary. Empirically, we find strong comovement in the absolute levels of pricing gaps between ETFs and their underlying assets, among ETFs served by the same LMM. Additional tests based on changes in ETFs' LMMs provide causal evidence that the excess comovement in ETF premium is indeed due to these ETFs sharing the same LMM. Specifically, for ETFs that change their LMMs, we find that their pricing gaps comove less with those of ETFs served by their previous LMMs, and more with that of ETFs served by their new LMMs. We also conduct a difference-in-differences test around the onset of COVID-19, driven by the observation that fixed income ETFs had large pricing gaps around this event. We find that LMMs that manage relatively more fixed income ETFs (and thus are likely more constrained) experience greater pricing gaps in their non-fixed income ETFs.

Our evidence suggests that LMMs play an important role in the pricing efficiencies of ETFs. Consistent with theories of intermediary-based asset pricing, the comovement in pricing efficiency among ETFs is less pronounced when the ETF and its underlying constituents are less costly to arbitrage, and for LMMs with less constrained capital. Our results validate the role of intermediaries and their capital constraints in the efficiency of financial market prices.

References

Adrian, T., Etula, E., and Muir, T. (2014). Financial intermediaries and the cross-section of asset returns. Journal of Finance 69(6), 2557–2596.

Almazan, A., Brown, K.C., Carlson, M., and Chapman, D.A. (2004). Why constrain your mutual fund manager? Journal of Financial Economics 73(2), 289–321.

Anton, M. and Polk, C. (2014). Connected stocks. Journal of Finance 69(3), 1099–1127.

Arora, R., Betermier, S., Ouellet Leblanc, G., Palumbo, A., and Shotlander, R. (2020). Concentration in the market of authorized participants of US fixed-income exchange-traded funds. Working Paper, Bank of Canada.

Bae, K. and Kim, D. (2020). Liquidity risk and exchange-traded fund returns, variances, and tracking errors. Journal of Financial Economics 138(1), 222–253.

Barberis, N., Shleifer, A., and Wurgler, J. (2005). Comovement. Journal of Financial Economics 75(2), 283–317.

Baron, M. and Muir, T. (2022). Intermediaries and asset prices: International evidence since 1870. Review of Financial Studies 35(5), 2144–2189.

Baron, M. and Xiong, W. (2017). Credit expansion and neglected crash risk. Quarterly Journal of Economics 132(2), 713–764.

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

Bhojraj, S., Mohanram, P., and Zhang, S. (2020). ETFs and information transfer across firms. Journal of Accounting and Economics 70(2-3), 101336.

Boyer, B.H. (2011). Style-related comovement: Fundamentals or labels? Journal of Finance 66(1), 307–332.

Boyson, N.M., Stahel, C.W., and Stulz, R.M. (2010). Hedge fund contagion and liquidity shocks. Journal of Finance 65(5), 1789–1816.

Broman, M.S. (2016). Liquidity, style investing and excess comovement of exchange-traded fund returns. Journal of Financial Markets *30*, 27–53.

Brown, D.C., Davies, S.W., and Ringgenberg, M.C. (2021). ETF arbitrage, non-fundamental demand, and return predictability. Review of Finance 25(4), 937–972.

Cenedese, G., Della Corte, P., and Wang, T. (2021). Currency mispricing and dealer balance sheets. Journal of Finance 76(6), 2763–2803.

Chung, J.W. and Kang, B.U. (2016). Prime broker-level comovement in hedge fund returns: Information or contagion? Review of Financial Studies 29(12), 3321–3353.

Corwin, S.A. and Schultz, P. (2012). A simple way to estimate bid-ask spreads from daily high and low prices. Journal of Finance 67(2), 719–760.

Coughenour, J. and Saad, M. (2004). Common market makers and commonality in liquidity. Journal of Financial Economics 73(1), 37–69.

Da, Z. and Shive, S. (2018). Exchange traded funds and asset return correlations. European Financial Management 24(1), 136–168.

Du, W., Tepper, A., and Verdelhan, A. (2018). Deviations from covered interest rate parity. Journal of Finance 73(3), 915–957.

Falato, A., Goldstein, I., and Hortaçsu, A. (2021). Financial fragility in the COVID-19 crisis: The case of investment funds in corporate bond markets. Journal of Monetary Economics 123, 35–52.

Fama, E.F. and French, K.R. (2015). A five-factor asset pricing model. Journal of Financial Economics 116(1), 1–22.

Glosten, L., Nallareddy, S., and Zou, Y. (2021). ETF activity and informational efficiency of underlying securities. Management Science 67(1), 22–47.

Goldberg, J. and Nozawa, Y. (2021). Liquidity supply in the corporate bond market. Journal of Finance 76(2), 755–796.

Gomes, J.F., Grotteria, M., and Wachter, J.A. (2019). Cyclical dispersion in expected defaults. Review of Financial Studies 32(4), 1275–1308.

Gorbatikov, E. and Sikorskaya, T. (2021). Two APs are better than one: ETF mispricing and primary market participation. Available at SSRN 3923503.

Green, T.C. and Hwang, B.H. (2009). Price-based return comovement. Journal of Financial Economics 93(1), 37–50.

Greenwood, R. (2008). Excess comovement of stock returns: Evidence from cross-sectional variation in Nikkei 225 weights. Review of Financial Studies 21(3), 1153–1186.

Grieser, W., Lee, J.H., and Zekhnini, M. (2020). Ubiquitous comovement. Available at SSRN 3385841.

Gromb, D. and Vayanos, D. (2018). The dynamics of financially constrained arbitrage. Journal of Finance 73(4), 1713–1750.

Haddad, V., Moreira, A., and Muir, T. (2021). When selling becomes viral: Disruptions in debt markets in the COVID-19 crisis and the Fed's response. Review of Financial Studies 34(11), 5309–5351.

Haddad, V. and Muir, T. (2021). Do intermediaries matter for aggregate asset prices? Journal of Finance 76(6), 2719–2761.

Hameed, A., Kang, W., and Viswanathan, S. (2010). Stock market declines and liquidity. Journal of Finance 65(1), 257–293.

Hameed, A. and Xie, J. (2019). Preference for dividends and return comovement. Journal of Financial Economics 132(1), 103–125.

He, Z., Kelly, B., and Manela, A. (2017). Intermediary asset pricing: New evidence from many asset classes. Journal of Financial Economics 126(1), 1–35.

He, Z., Khorrami, P., and Song, Z. (2019). Commonality in credit spread changes: Dealer inventory and intermediary distress. Working Paper, National Bureau of Economic Research.

Holden, C. and Subrahmanyam, A. (2002). News events, information acquisition, and serial correlation. Journal of Business *75*, 1–32.

Huang, S., O'Hara, M., and Zhong, Z. (2021). Innovation and informed trading: Evidence from industry ETFs. Review of Financial Studies 34(3), 1280–1316.

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.

Israelsen, R.D. (2016). Does common analyst coverage explain excess comovement? Journal of Financial and Quantitative Analysis 51(4), 1193–1229.

Kyle, A.S. (1989). Informed speculation with imperfect competition. Review of Economic Studies 56(3), 317–355.

Leland, H. (1992). Insider trading: Should it be prohibited? Journal of Political Economy 100(4), 859–887.

Lewis, K.F., Longstaff, F.A., and Petrasek, L. (2021). Asset mispricing. Journal of Financial Economics 141(3), 981–1006.

Lu, Z. and Qin, Z. (2021). Leveraged funds and the shadow cost of leverage constraints. Journal of Finance 76(3), 1295–1338.

Macchiavelli, M. and Zhou, X. (2022). Funding liquidity and market liquidity: the broker-dealer perspective. Management Science 68(5), 3379–3398.

Merton, R.C. (1980). On estimating the expected return on the market: An exploratory investigation. Journal of Financial Economics 8(4), 323–361.

Nagel, S. (2012). Evaporating liquidity. Review of Financial Studies 25(7), 2005–2039.

Ottonello, P. and Song, W. (2022). Financial intermediaries and the macroeconomy: Evidence from a high-frequency identification. Working Paper, National Bureau of Economic Research.

Pan, K. and Zeng, Y. (2019). ETF arbitrage under liquidity mismatch. Available at SSRN 3723406.

Pasquariello, P. (2014). Financial market dislocations. Review of Financial Studies 27(6), 1868–1914.

Pástor, L. and Vorsatz, M.B. (2020). Mutual fund performance and flows during the COVID-19 crisis. Review of Asset Pricing Studies 10(4), 791–833.

Peng, L. and Xiong, W. (2006). Investor attention, overconfidence and category learning. Journal of Financial Economics 80(3), 563–602.

Petajisto, A. (2017). Inefficiencies in the pricing of exchange-traded funds. Financial Analysts Journal 73(1), 24–54.

Pirinsky, C. and Wang, Q. (2006). Does corporate headquarters location matter for stock returns? Journal of Finance 61(4), 1991–2015.

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

Pontiff, J. (2006). Costly arbitrage and the myth of idiosyncratic risk. Journal of Accounting and Economics 42(1), 35–52.

Richie, N., Daigler, R.T., and Gleason, K.C. (2008). The limits to stock index arbitrage: Examining S&P 500 futures and SPDRS. Journal of Futures Markets: Futures, Options, and Other Derivative Products 28(12), 1182–1205.

Rösch, D.M., Subrahmanyam, A., and Van Dijk, M.A. (2017). The dynamics of market efficiency. Review of Financial Studies 30(4), 1151–1187.

Santos, T. and Veronesi, P. (2022). Leverage. Journal of Financial Economics 145(2), 362–386.

Shleifer, A. and Vishny, R.W. (1997). The limits of arbitrage. Journal of Finance 52(1), 35–55.

Wahal, S. and Yavuz, M.D. (2013). Style investing, comovement and return predictability. Journal of Financial Economics 107(1), 136–154.

Appendix A

Proof of Proposition 1: (a) Consider the mass M of fundamental investors who trade the underlying stocks. Write the mth such investor's wealth at Date 1 as:

$$W_{m1} = W_{m0} + \sum_{\kappa=1}^{K} \left[X_{m\kappa}(V_\kappa - P_\kappa) \right].$$

The demand $X_{m\kappa}$ maximizes:

$$\begin{aligned} E\left[U(W_{m1})\right] &= E\left[-\exp\left[-AW_{m0} - A\sum_{\kappa=1}^{K}\left[X_{m\kappa}(V_\kappa - P_\kappa)\right]\right]\right] \\ &= -\exp\left[-AW_{m0} - A\sum_{\kappa=1}^{K}\left[X_{m\kappa}(-P_\kappa)\right] + 0.5A^2\nu\sum_{\kappa=1}^{K}X_{m\kappa}^2\right]. \end{aligned}$$

The first order condition (f.o.c.) with respect to (w.r.t.) $X_{m\kappa}$ implies that the demand is:

$$X_\kappa = -\lambda P_\kappa, \quad \text{where} \quad \lambda = 1/(A\nu). \tag{A.1}$$

The second order condition holds in the above case, and in all other cases below, so let us omit the reference to it in the rest of the proofs.

For stock $\kappa \notin S_1 \cup S_2$, the market clearing condition $MX_\kappa + z_\kappa = 0$, where $X_\kappa = -\lambda P_\kappa$ from Equation (A.1), implies that $P_\kappa = z_\kappa/(M\lambda)$. In what follows, we focus on stocks in $S_1 \cup S_2$.

(b) Consider the mass N of fundamental investors who trade the ETFs. Write the nth such investor's wealth at Date 1 as

$$W_{n1} = W_{n0} + \sum_{i=1}^{2} \left[\hat{X}_{ni}(\hat{V}_i - Q_i) \right].$$

The demand $\hat{X}_{ni}$ maximizes

$$\begin{aligned} E\left[U(W_{n1})\right] &= E\left[-\exp\left[-AW_{n0} - A\sum_{i=1}^{2}\left[\hat{X}_{ni}(\hat{V}_i - Q_i)\right]\right]\right] \\ &= -\exp\left[-AW_{n0} - A\sum_{i=1}^{2}\left[\hat{X}_{ni}(-Q_i)\right] + 0.5A^2\hat{K}\nu\sum_{i=1}^{2}\hat{X}_{ni}^2\right]. \end{aligned}$$

The f.o.c. w.r.t. $\hat{X}_{ni}$ implies that the demand is:

$$\hat{X}_i = -\hat{\lambda} Q_i, \quad \text{where} \quad \hat{\lambda} = \lambda/\hat{K}. \tag{A.2}$$

(c) Consider the LMM trading ETF i. The LMM recognizes that with a demand $-Y_i$ (which is equivalent to supplying Y_i shares of the ETF) and given that all other investors follow their trading strategies, the market clearing condition is: $N\hat{X}_i - Y_i + \hat{z}_i = 0$, where $\hat{X}_i = -\hat{\lambda} Q_i$ from Equation (A.2), specifies a price:

$$Q_i = \hat{\ell} Y_i + Q_i^-, \quad \text{where} \quad \hat{\ell} = -1/(N\hat{\lambda}). \tag{A.3}$$

$\hat{\ell}$ represents the slope of the supply curve for the LMM's supply Y_i specifically, and Q_i^- represents the residual price.

For stock $\kappa \in S_i$, the LMM recognizes that with a demand Y_i and given that all other investors follow their trading strategies, the market clearing condition $MX_\kappa + Y_i + z_\kappa = 0$, where $X_\kappa = -\lambda P_\kappa$ from Equation (A.1), specifies the price

$$P_\kappa \quad = \quad \ell Y_i + P_\kappa^-, \quad \text{where} \quad \ell = 1/(M\lambda). \tag{A.4}$$

ℓ represents the slope of the demand curve for the LMM's demand Y_i specifically, and P_κ^- represents the residual price.

Using Equations (A.3) and (A.4), we express the LMM's payoff, net of the cost, as:

$$\begin{aligned} \mathit{\Pi} &= \sum_{i=1}^{2} \left[Y_i \left(Q_i - \sum_{\kappa \in S_i} P_\kappa \right) - 0.5cY_i^2 \right] \\ &= \sum_{i=1}^{2} \left[Y_i \left[(\hat{\ell} Y_i + Q_i^-) - \sum_{\kappa \in S_i} (\ell Y_i + P_\kappa^-) \right] - 0.5cY_i^2 \right], \end{aligned} \tag{A.5}$$

with the constraint

$$\sum_{i=1}^{2} C_i = c \sum_{i=1}^{2} Y_i^2 \leq \mathit{\Phi}. \tag{A.6}$$

In what follows, we consider two cases: In the first case (d1), Constraint (A.6) is not binding; in the second case (d2), Constraint (A.6) binds.

(d1) Suppose that Constraint (A.6) is not binding. Consider Equation (A.5), and denote

$$\delta_i \equiv Q_i - \sum_{\kappa \in S_i} P_\kappa. \tag{A.7}$$

The f.o.c. w.r.t. Y_i is expressed as:

$$\delta_i + Y_i(\hat{\ell} - \hat{K}\ell) - cY_i = 0.$$

It follows from the previously derived $\ell = (M\lambda)^{-1}$ and $\hat{\ell} = -(N\hat{\lambda})^{-1}$ and the fact $\hat{\lambda} = \lambda/\hat{K}$ that $\hat{K}\ell - \hat{\ell} = (\hat{\lambda})^{-1}(N^{-1} + M^{-1})$. We express the demand as:

$$Y_i \;=\; \eta\delta_i, \quad \text{where} \;\; \eta = \frac{1}{c + (\hat{\lambda})^{-1}(N^{-1} + M^{-1})}. \tag{A.8}$$

The market clearing conditions for the ETF and stock κ, respectively, are:

$$N\hat{X}_i - Y_i + \hat{z}_i = 0 \;\; \text{and} \;\; MX_\kappa + Y_i + z_\kappa = 0.$$

From the expressions of X_κ, $\hat{X}_i$, and Y_i in Equations (A.1), (A.2), and (A.8), we obtain the prices: and imply the prices:

$$Q_i = \frac{-\eta\delta_i + \hat{z}_i}{N\hat{\lambda}} \;\; \text{and} \;\; P_\kappa = \frac{\eta\delta_i + z_\kappa}{M\lambda}.$$

Denote

$$Z_i = \frac{\hat{z}_i}{N} - \frac{\bar{z}_{\kappa,i}}{M}, \;\; \text{where} \;\; \bar{z}_{\kappa,i} = \sum_{\kappa \in S_i} z_\kappa \Big/ \hat{K}.$$

It then follows from Equation (A.7) that:

$$\delta_i = \frac{Z_i}{\hat{\lambda} + \eta(N^{-1} + M^{-1})}.$$

Note from Equation (A.8) that $Y_i = \eta\delta_i$. For Constraint (A.6) to not be binding, the requirement is that $\delta_1^2 + \delta_2^2 < \Phi/(c\eta^2)$.

(d2) Now suppose that Constraint (A.6) binds; this requires $\delta_1^2 + \delta_2^2 \geq \Phi/(c\eta^2)$.

In this case, $Y_1^2 + Y_2^2 = \Phi/c$. Treating Y_2 as a function of Y_1, it follows that $dY_2/dY_1 = -Y_1/Y_2$. Also from $Y_1^2 + Y_2^2 = \Phi/c$, Π in Equation (A.5) is expressed as:

$$\Pi \quad = \quad \sum_{i=1}^{2} \left[Y_i \left[(\hat{\ell} Y_i + Q_i^-) - \sum_{\kappa \in S_i} (\ell Y_i + P_\kappa^-) \right] \right] - \Phi.$$

Treat Π as a function of Y_1, and denote

$$\Delta_i \equiv Q_i - \sum_{\kappa \in S_i} P_\kappa. \tag{A.9}$$

The f.o.c. w.r.t. Y_1 is given by:

$$0 \quad = \quad \Delta_1 + Y_1(\hat{\ell} - \hat{K}\ell) - \left[\Delta_2 + Y_2(\hat{\ell} - \hat{K}\ell) \right] \frac{Y_1}{Y_2} = \Delta_1 - \Delta_2 \frac{Y_1}{Y_2};$$

it follows that $Y_1 = (\Delta_1/\Delta_2)Y_2$. Using $Y_1^2 + Y_2^2 = \Phi/c$, we obtain

$$Y_i = \Delta_i \sqrt{\frac{\Phi/c}{\Delta_1^2 + \Delta_2^2}}. \tag{A.10}$$

The market clearing conditions for ETF i and stock κ, respectively, are:

$$N\hat{X}_i - Y_i + \hat{z}_i = 0 \;\; \text{and} \;\; MX_\kappa + Y_i + z_\kappa = 0.$$

From Equations (A.1), (A.2), and (A.10), we obtain the prices:

$$Q_i = \frac{1}{N\hat{\lambda}} \left(-\Delta_i \sqrt{\frac{\Phi/c}{\Delta_1^2 + \Delta_2^2}} + \hat{z}_i \right) \;\; \text{and} \;\; P_\kappa = \frac{1}{M\lambda} \left(\Delta_i \sqrt{\frac{\Phi/c}{\Delta_1^2 + \Delta_2^2}} + z_\kappa \right).$$

We can use Equation (A.9) to show that:

$$\Delta_i \left[\hat{\lambda} + \left(N^{-1} + M^{-1} \right) \sqrt{\frac{\Phi/c}{\Delta_1^2 + \Delta_2^2}} \right] = \frac{\hat{z}_i}{N} - \frac{\bar{z}_{\kappa,i}}{M} = Z_i,$$

which holds for both $i = 1$ and $i = 2$; then,

$$\mathit{\Delta}_i = \frac{Z_i}{\hat{\lambda}} \left[1 - \left(N^{-1} + M^{-1} \right) \sqrt{\frac{\Phi/c}{Z_1^2 + Z_2^2}} \right].$$

Finally, it is straightforward to show from $\Phi/(c\eta^2) \leq \delta_1^2 + \delta_2^2$ that

$$1 - \left(N^{-1} + M^{-1} \right) \sqrt{\frac{\Phi/c}{Z_1^2 + Z_2^2}} > 0.$$

□

Proof of Proposition 2: Without loss of generality, let $i = 1$. Note that if $\delta_1^2 + \delta_2^2 < \Phi/(c\eta^2)$, then the premium of ETF 1 is given by:

$$|\delta_1| \quad = \quad \frac{|Z_1|}{\hat{\lambda} + \eta(N^{-1} + M^{-1})},$$

which increases in $|Z_1|$, but does not depend on $|Z_2|$.

If $\delta_1^2 + \delta_2^2 \geq \Phi/(c\eta^2)$, then the premium of ETF 1 is given by:

$$|\mathit{\Delta}_1| = \frac{|Z_1|}{\hat{\lambda}} \left[1 - \left(N^{-1} + M^{-1} \right) \sqrt{\frac{\Phi/c}{Z_1^2 + Z_2^2}} \right],$$

where from the proof of Proposition 1,

$$1 - \left(N^{-1} + M^{-1} \right) \sqrt{\frac{\Phi/c}{Z_1^2 + Z_2^2}} > 0.$$

It is straightforward to show, after taking derivatives, that $\mathit{\Delta}_1$ increases in both $|Z_1|$ and $|Z_2|$. This completes the proof of the proposition. □

Proof of Proposition 3: We use three steps in this proof. In Steps 1 and 2, we prove some general statements regarding the expectations of monotonic functions of random variables. In Step 3, we apply the statements to prove this proposition.

Step 1: We prove the following statement: If $f(x)$ and $g(x)$ are increasing functions of a random variable x, then $E[f(x)g(x)] > E[f(x)]\,E[g(x)]$ or, equivalently, $\mathrm{Cov}[f(x), g(x)] > 0$.

Suppose that x_1 and x_2 are drawn independently from the distribution of x. Since both $f(\cdot)$ and

$g(\cdot)$ are increasing functions,

$$[f(x_1) - f(x_2)]\,[g(x_1) - g(x_2)] \geq 0,$$

where the inequality holds strictly in a non-zero measure within the set $\{x_1, x_2\}$. Taking an expectation yields:

$$E\,[[f(x_1) - f(x_2)]\,[g(x_1) - g(x_2)]] > 0,$$

where the left-hand side, after expansion, becomes:

$$E\,[f(x_1)g(x_1) - f(x_1)g(x_2) - f(x_2)g(x_1) + f(x_2)g(x_2)] = 2E\,[f(x)g(x)] - 2E\,[f(x)]\,E\,[g(x)]\,.$$

It follows immediately that $E\,[f(x)g(x)] > E\,[f(x)]\,E\,[g(x)]$.

Step 2: We prove the following statement: If $F(x, y)$ and $G(x, y)$ are increasing functions of two independent random variables x and y, then $E\,[F(x, y)G(x, y)] \; > \; E\,[F(x, y)]\,E\,[G(x, y)]$ or, equivalently, $\mathrm{Cov}\,[F(x, y), G(x, y)] > 0$.

Let $E_y(.)$ denote taking an expectation over y. Note that conditional on x, both $F(x, y)$ and $G(x, y)$ increase with y; it follows from the Step-1 statement that

$$E_y\,[F(x, y)G(x, y)|x] \quad > \quad E_y\,[F(x, y)|x]\,E_y\,[G(x, y)|x]\,. \tag{A.11}$$

Also note that both $E_y[F(x, y)|x]$ and $E_y[G(x, y)|x]$ increase with x; it follows from the Step-1 statement that:

$$\begin{aligned} E_x\,[E_y\,[F(x, y)|x]\,E_y\,[G(x, y)|x]] \quad &> \quad E_x\,[E_y\,[F(x, y)|x]]\,E_x\,[E_y\,[G(x, y)|x]] \\ &= \quad E\,[F(x, y)]\,E\,[G(x, y)]\,. \end{aligned} \tag{A.12}$$

Therefore,

$$\begin{aligned} E\,[F(x, y)G(x, y)] = E_x\,[E_y\,[F(x, y)G(x, y)|x]] \quad &> \quad E_x\,[E_y\,[F(x, y)|x]\,E_y\,[G(x, y)|x]] \\ &> \quad E\,[F(x, y)]\,E\,[G(x, y)]\,, \end{aligned}$$

where the first inequality follows from Equation (A.11), and the second inequality follows from Equation (A.12).

Step 3: From Proposition 2, we can treat the premium of ETF i, $|Q_i - NAV_i|$, as a function of $|Z_i|$ and $|Z_j|$. This premium is increasing in $|Z_i|$; it also is increasing in $|Z_j|$ when $\delta_1^2 + \delta_2^2 \geq \Phi/(c\eta^2)$. A symmetric result holds for the premium of ETF j, $|Q_j - NAV_j|$. Note that $|Z_i|$ and $|Z_i|$ are independent random variables; it follows from the Step-2 statement that $\text{Cov}(|Q_i - NAV_i|, |Q_j - NAV_j|) > 0$. Therefore,

$$\beta \propto \text{Cov}(|Q_i - NAV_i|, |Q_j - NAV_j|) > 0.$$

□

Proof of Corollary 1: It follows from Equation (3) that if $c \to 0$ and/or $\Phi \to \infty$, then almost surely, $\delta_1^2 + \delta_2^2 < \Phi/(c\eta^2)$; in this case, the premium of ETF i, $|Q_i - NAV_i| = |\delta_i|$. It follows that

$$\beta \propto \text{Cov}(|Q_i - NAV_i|, |Q_j - NAV_j|) = 0$$

because for $i \neq j$, $Z_i \perp Z_j$ and thus $\delta_i \perp \delta_j$ from Equation (2).

If $\nu \to 0$, then almost surely, $V_\kappa = 0\ \forall \kappa$ and $V_i = 0$ for $i = 1, 2$. We can follow the derivation in the proof of Proposition 1 to show that in this case, $P_\kappa = V_\kappa = 0\ \forall \kappa$ and $P_i = V_i = 0$ for $i = 1, 2$; thus, $|Q_i - NAV_i| = 0$ for $i = 1, 2$. It holds trivially that the comovement parameter $\beta \to 0$. □

Proof of Proposition 4: (a) Consider the mass M of fundamental investors who trade the underlying stocks. We use a similar derivation as in the Proof of Proposition 1 to show that each such investor's demand for stock κ is given by:

$$X_\kappa = -\lambda P_\kappa. \tag{A.13}$$

Further, for stock $\kappa \notin S_1 \cup S_2$, this price $P_\kappa = z_\kappa/(M\lambda)$. In what follows, we focus on stocks in S_i, $i = \{1, 2\}$.

(b) Consider the mass N of fundamental investors who trade the ETFs. We use a similar derivation as in the Proof of Proposition 1 to show that each such investor's demand for ETF i is given by:

$$\hat{X}_i = -\hat{\lambda} Q_i. \tag{A.14}$$

(c) Consider the LMM trading ETF i. We use a similar derivation as in the Proof of Proposition 1 to show that the prices of ETF i and stock $\kappa \in S_i$ take the following form:

$$Q_i = \hat{\ell} Y_i + Q_i^- \text{ and } P_\kappa = \ell Y_i + P_\kappa^-.$$

The LMM's payoff, net of the cost, is expressed as:

$$\Pi_i = Y_i \left[(\hat{\ell} Y_i + Q_i^-) - \sum_{\kappa \in S_i} (\ell Y_i + P_\kappa^-) \right] - 0.5 c Y_i^2 \tag{A.15}$$

with the constraint

$$C_i = c Y_i^2 \leq \Phi. \tag{A.16}$$

(d1) Suppose that Constraint (A.16) is not binding. We use a similar derivation as in the Proof of Proposition 1 to show that the demand $Y_i = \eta \delta_i$, and the prices:

$$Q_i = \frac{-\eta \delta_i + \hat{z}_i}{N \hat{\lambda}} \text{ and } P_\kappa = \frac{\eta \delta_i + z_\kappa}{M \lambda},$$

where

$$\delta_i = Q_i - \sum_{\kappa \in S_i} P_\kappa = \frac{Z_i}{\hat{\lambda} + \eta (N^{-1} + M^{-1})}.$$

Note that $Y_i = \eta \delta_i$. For Constraint (A.16) not bind, we require $\delta_i^2 < \Phi / (c \eta^2)$.

(d2) Now suppose that Constraint (A.16) binds; this requires $\delta_i^2 \geq \Phi / (c \eta^2)$.

Since Constraint (A.16) is binding, $Y_i^2 = \Phi / c$. Express Π_i in Equation (A.15) as: $\Pi_i = Y_i \Delta_i^* - \Phi$, where

$$\varDelta_i^* = Q_i - \sum_{\kappa \in S_i} P_\kappa. \tag{A.17}$$

It follows that:

$$Y_i = \text{sign}(\varDelta_i^*) \times \sqrt{\frac{\varPhi}{c}}. \tag{A.18}$$

The market clearing conditions for ETF i and stock κ, respectively, are:

$$N\hat{X}_i - Y_i + \hat{z}_i = 0 \;\text{ and }\; MX_\kappa + Y_i + z_\kappa = 0.$$

From Equations (A.13), (A.14), and (A.18), we obtain the prices:

$$Q_i = \frac{1}{N\hat{\lambda}}\left[-\text{sign}(\varDelta_i^*) \times \sqrt{\frac{\varPhi}{c}} + \hat{z}_i\right] \;\text{ and }\; P_\kappa = \frac{1}{M\lambda}\left[\text{sign}(\varDelta_i^*) \times \sqrt{\frac{\varPhi}{c}} + z_\kappa\right].$$

We use Equation (A.17) to show that:

$$\varDelta_i^* = \frac{1}{\hat{\lambda}}\left[Z_i - \text{sign}(\varDelta_i^*) \times \sqrt{\frac{\varPhi}{c}} \times \left(\frac{1}{N} + \frac{1}{M}\right)\right]. \tag{A.19}$$

Note from the binding Constraint (A.16), $\delta_i^2 \geq \varPhi/(c\eta^2)$. If $\delta_i \geq \eta^{-1}\sqrt{\varPhi/c}$, then it is straightforward to show that:

$$Z_i \geq \sqrt{\frac{\varPhi}{c}} \times \left(\frac{\hat{\lambda}}{\eta} + N^{-1} + M^{-1}\right) > \sqrt{\frac{\varPhi}{c}} \times \left(N^{-1} + M^{-1}\right);$$

it follows from Equation (A.19) that in this case,

$$\varDelta_i^* = \frac{1}{\hat{\lambda}}\left[Z_i - \sqrt{\frac{\varPhi}{c}} \times \left(N^{-1} + M^{-1}\right)\right].$$

If $\delta_i \leq -\eta^{-1}\sqrt{\varPhi/c}$, then it is straightforward to show that:

$$Z_i \leq -\sqrt{\frac{\varPhi}{c}} \times \left(\frac{\hat{\lambda}}{\eta} + N^{-1} + M^{-1}\right) < -\sqrt{\frac{\varPhi}{c}} \times \left(N^{-1} + M^{-1}\right);$$

it follows from Equation (A.19) that in this case,

$$\Delta_i^* = \frac{1}{\hat{\lambda}}\left[Z_i + \sqrt{\frac{\Phi}{c}} \times \left(N^{-1} + M^{-1}\right)\right].$$

(e) It follows from the previous derivations (d1) and (d2) that the premium of ETF i is expressed as:

$$|Q_i - NAV_i| = \begin{cases} |\delta_i| & \text{if } |\delta_i| < \eta^{-1}\sqrt{\Phi/c}, \\ \frac{1}{\hat{\lambda}}\left[Z_i - \sqrt{\frac{\Phi}{c}} \times \left(N^{-1} + M^{-1}\right)\right] & \text{if } \delta_i \geq \eta^{-1}\sqrt{\Phi/c}, \\ \frac{1}{\hat{\lambda}}\left[Z_i + \sqrt{\frac{\Phi}{c}} \times \left(N^{-1} + M^{-1}\right)\right] & \text{if } \delta_i \leq -\eta^{-1}\sqrt{\Phi/c}. \end{cases}$$

Note that here $|Q_i - NAV_i|$ depends only on Z_i; thus, $|Q_i - NAV_i|$ is independent of $|Q_j - NAV_j|$ where $j \neq i$. It follows that the comovement parameter

$$\beta \propto \text{Cov}(|Q_i - NAV_i|, |Q_j - NAV_j|) = 0.$$

□

Appendix B: Variable Definitions

This table reports the definitions of the main variables used in the empirical analysis.

Variables	Definition
raw $\vert Premium\vert_{i,t}$	The absolute value of an ETF's raw premium, defined as |(ETF Price − ETF NAV)/ETF NAV|.
LMM raw $\vert Premium\vert_{i,t}$	Equally-weighted average raw |*Premium*|$_{i,t}$ of all ETFs sharing the same lead market maker as the ETF i, excluding the focal ETF i.
non-LMM raw $\vert Premium\vert_{i,t}$	Equally-weighted average $\vert Premium\vert_{j,t}$ of all ETFs managed by a lead market maker that is different from that of the focal ETF i.
$\vert Premium\vert_{i,t}$	We orthogonalize each ETF's *raw* $\vert Premium\vert$ with respect to its *non-LMM raw* $\vert Premium\vert$, by estimating the following regression: *raw* $\vert Premium\vert_{i,t} = a + b \times$ *non-LMM raw* $\vert Premium\vert_{i,t} + \epsilon_{i,t}$. An ETF's $\vert Premium\vert_{i,t}$ is captured by the residual terms $\epsilon_{i,t}$.
LMM $\vert Premium\vert_{i,t}$	Equally-weighted average $\vert Premium\vert_{i,t}$ of all ETFs sharing the same lead market maker as the ETF i, excluding the focal ETF i.
non-LMM $\vert Premium\vert_{i,t}$	Equally-weighted average $\vert Premium\vert_{j,t}$ of all ETFs managed by a lead market maker that is different from that of the focal ETF i.
LMM$_{Old}$ $\vert Premium\vert_{i,t}$	Equally-weighted average $\vert Premium\vert$ of all ETFs managed by the ETF's old LMM, excluding the focal ETF i itself. Old LMM is the lead market maker just before the ETF changes its lead market maker.
LMM$_{New}$ $\vert Premium\vert_{i,t}$	Equally-weighted average $\vert Premium\vert$ of all ETFs managed by the ETF's new LMM, excluding the focal ETF i itself. New LMM is the lead market maker just after the ETF changes its lead market maker.
$Log(Size)$	The natural logarithm of an ETF's market capitalization.

STD	The standard deviation of daily ETF returns estimated using data from $t-30$ to $t-1$.
$BidAsk$	The difference between ask and bid quotes scaled by the average of bid and ask quotes (in bp), estimated using daily data from $t-30$ to $t-1$.
$Turnover$	Daily dollar trading volume of an ETF scaled by its market capitalization (in percent), estimated using daily data from $t-30$ to $t-1$.
$Log(Mktcap\ of\ ETFs)$	Natural logarithm of the total market capitalization of ETFs managed by the LMM.
#Active AP	The number of active authorized participants for the ETF, as reported in form N-CEN.
D (Earnings Announcement Return <-3%)	Dummy variable that equals one for the [-1,1] trading day window of earnings announcement, with an announcement-day CAR more negative than -3%.
Net Capital Ratio	Net Capital Required/Adjusted Net Capital, which measures the amount of capital an LMM holds relative to the CFTC's net capital requirement.
VIX	The CBOE volatility index.
CS	Credit spread = Moody's BAA Yield − Yield on Treasury 10-year constant maturity.
HKM	The intermediary capital ratio of He, Kelly, and Manela (2017).
$Mktret$	Stock market return (VWRETD) in the prior month.

Figure 1. Illustration of arbitrage on ETF price and net asset value deviation

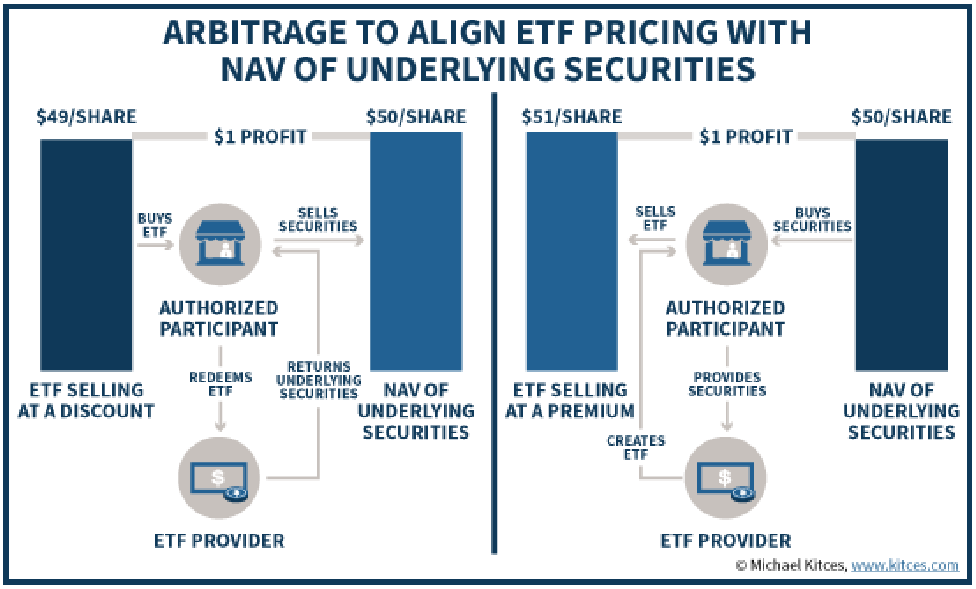


Figure 2. Comovement in ETF premia as a function of the cost of arbitrage

This graph plots β (the parameter representing the comovement between the premium of ETF i, $|Q_i - NAV_i|$, and the premium of ETF j (where $j \neq i$), $|Q_j - NAV_j|$) as a function of c (the parameter representing the cost of arbitrage). The other parameter values are $A = 2$, $\nu = 1$, $M = 1$, $\mu = 0.2$, $\hat{K} = 25$, $N = 0.5$, $\hat{\mu} = 0.1$, and $\Phi = 0.05$.

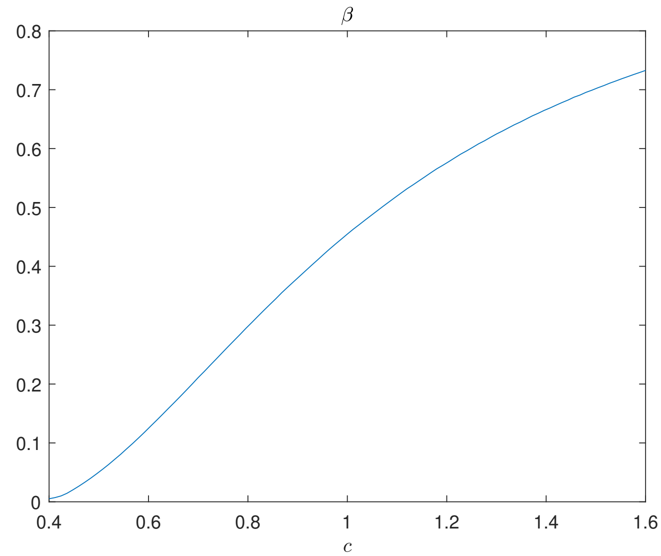


Figure 3. Comovement in ETF premia as a function of return volatility

This graph plots β (the parameter representing the comovement between the premium of ETF i, $|Q_i - NAV_i|$, and the premium of ETF j (where $j \neq i$), $|Q_j - NAV_j|$) as a function of ν (the parameter representing return volatility). The other parameter values are $A = 2$, $M = 1$, $\mu = 0.2$, $\hat{K} = 25$, $N = 0.5$, $\hat{\mu} = 0.1$, $c = 1$, and $\Phi = 0.05$.

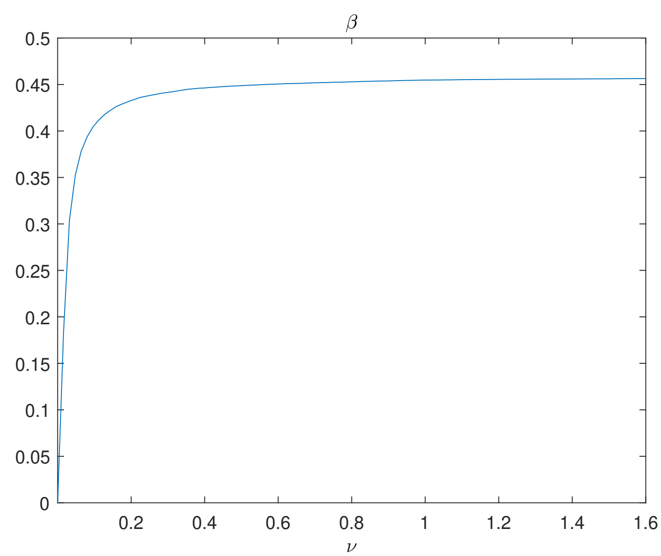


Figure 4. Comovement in ETF premium as a function of the LMM capital constraint

This graph plots β (the parameter representing the comovement between the premium of ETF i, $|Q_i - NAV_i|$, and the premium of ETF j (where $j \neq i$), $|Q_j - NAV_j|$) as a function of Φ (the parameter representing the LMM's capital constraint). The other parameter values are $A = 2$, $\nu = 1$, $M = 1$, $\mu = 0.2$, $\hat{K} = 25$, $N = 0.5$, $\hat{\mu} = 0.1$, and $c = 1$.

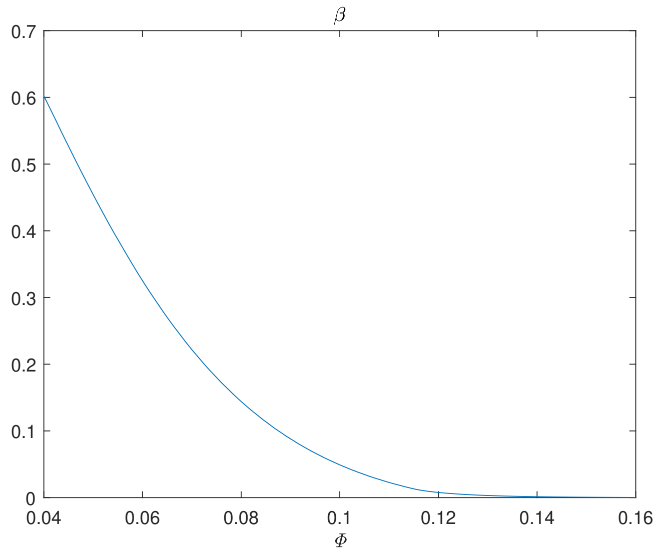
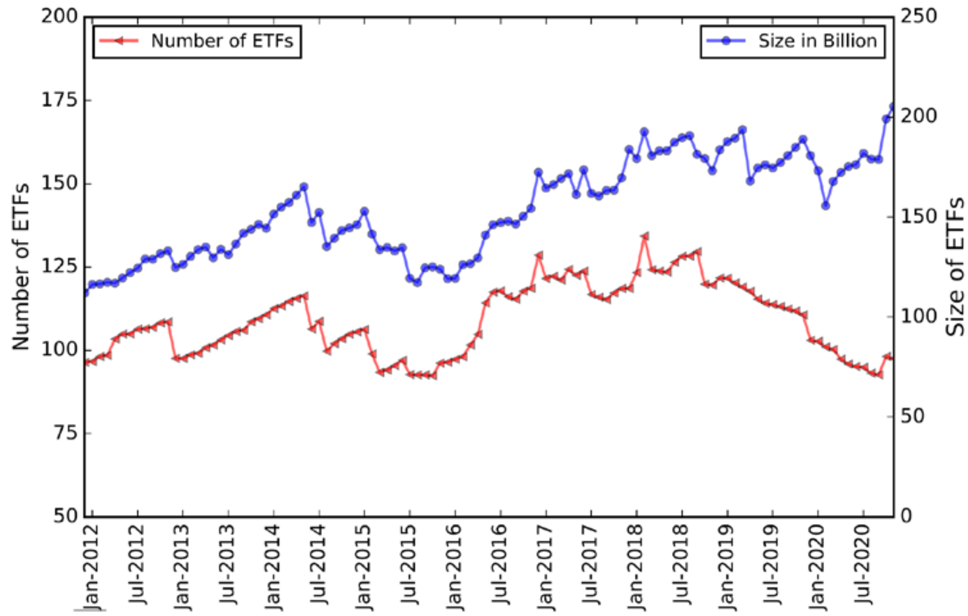
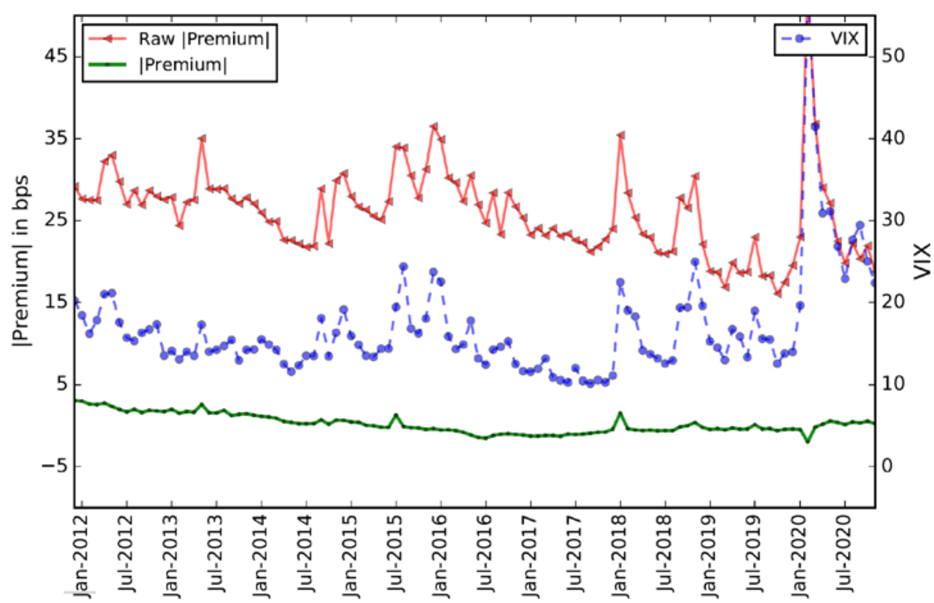


Figure 5. Number and total size of ETFs managed by an average LMM, and ETF premium over Time
Panel A shows the number and total size of ETFs (in billions USD) managed by an average lead market maker in our sample. Panel B shows the equally-weighted average raw |*Premium*| and residual |*Premium*| for each month. On the right axis, the blue dotted line shows the level of the VIX. The sample period runs from January 2012 to December 2020.



Panel A: Coverage of ETFs by an average LMM



Panel B: Average absolute ETF premium

Figure 6. Comovement box: Relationship between ETF |*Premium*| and LMM |*Premium*|
We estimate co-dependence between the ETF residual |*Premium*| and LMM residual |*Premium*| using a quantile regression approach, and plot the results in a co-movement box. This box is a square of unit side that plots the conditional probability that an ETF has a residual |*Premium*| below or above a certain percentile conditional on the same event occurring in its LMM residual |*Premium*|. To construct the figure, we calculate the probabilities at the 5% increments between the 5th and 95th quantiles. The blue line denotes the 45-degree line, which represents the unconditional probability of no dependence between the variables. The red line denotes the actual conditional probability distribution estimated using the ETF residual |Premium|.

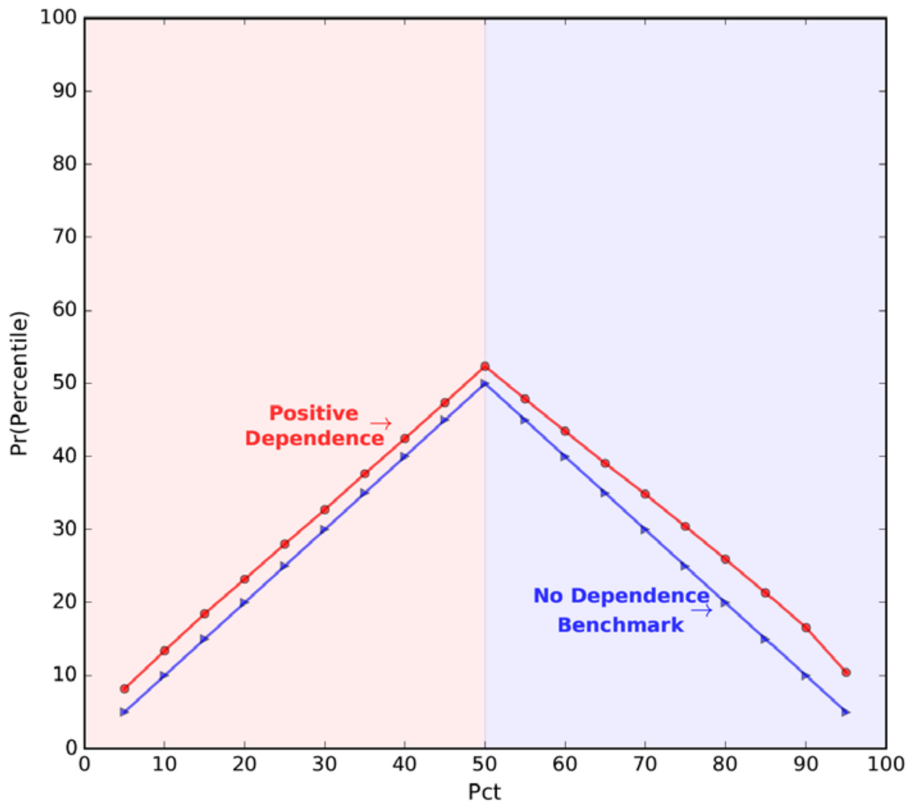


Figure 7. Comovement in |*Premium*| when ETFs change the lead market makers

This figure reports the comovement in the (absolute) ETF premium with those of other ETFs served by the old and new lead market makers (LMMs) over the [-120, 120] trading days around the change in LMM. $LMM_{Old}\ |Premium|$ ($LMM_{New}\ |Premium|$) is the average absolute premium of ETF i's previous (new) LMM, excluding ETF i. For each trading day around the event, we use a [-3, +3] trading day window to estimate the following regression:

$$|Premium|_{i,t} = a + bLMM_{Old}|Premium|_{i,t} \quad + cLMM_{New}|Premium|_{i,t} + \varepsilon,$$

and plot the estimates of b and c. The shaded area is the 95% confidence interval. The upper graph shows the regression coefficients estimated using the raw |*Premium*| and the lower graph shows the estimates for the residual |*Premium*|.

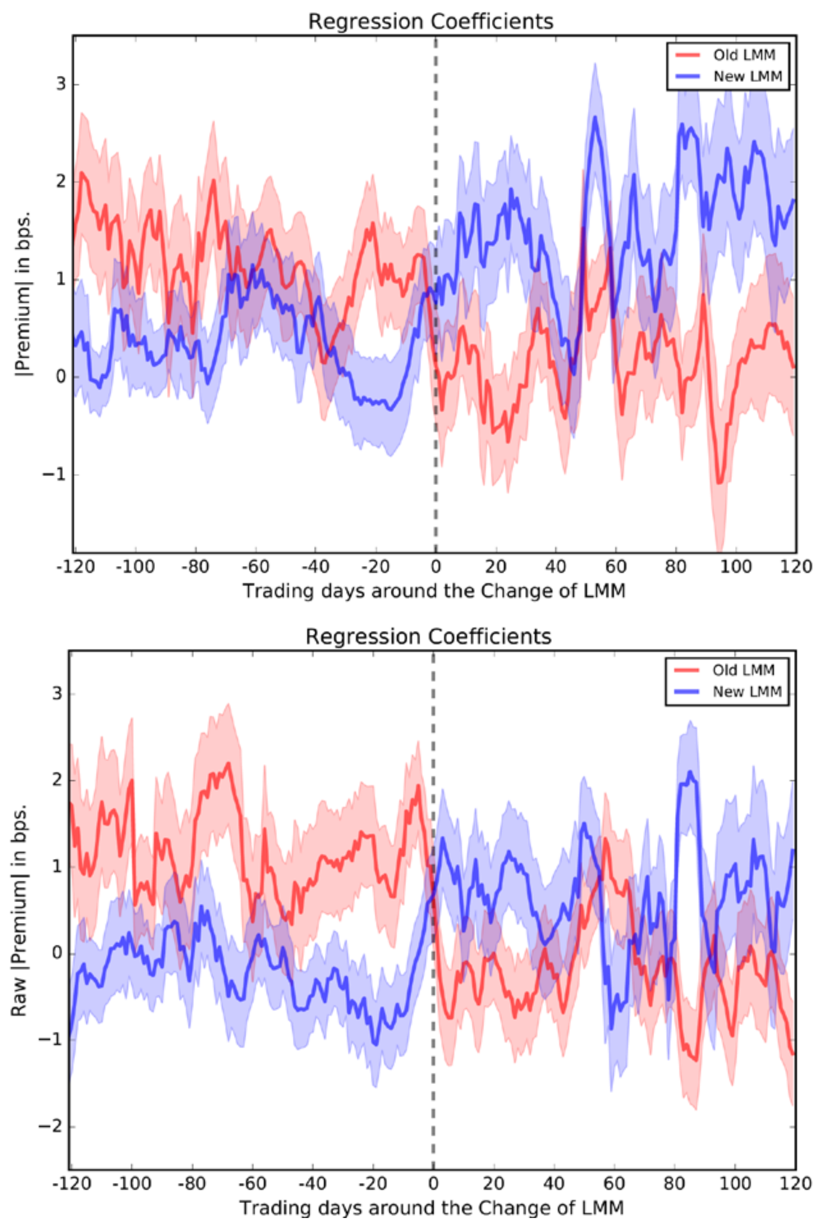
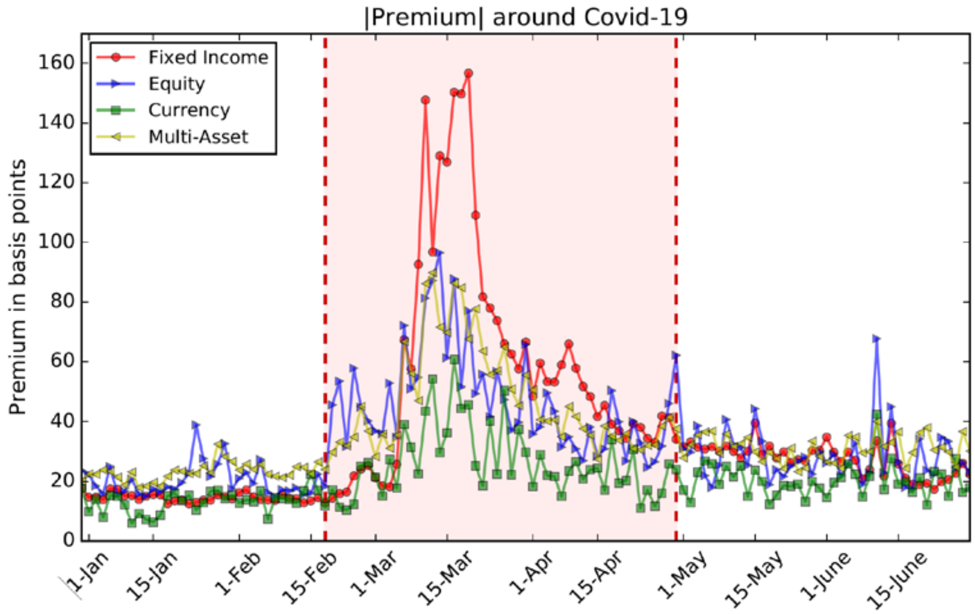
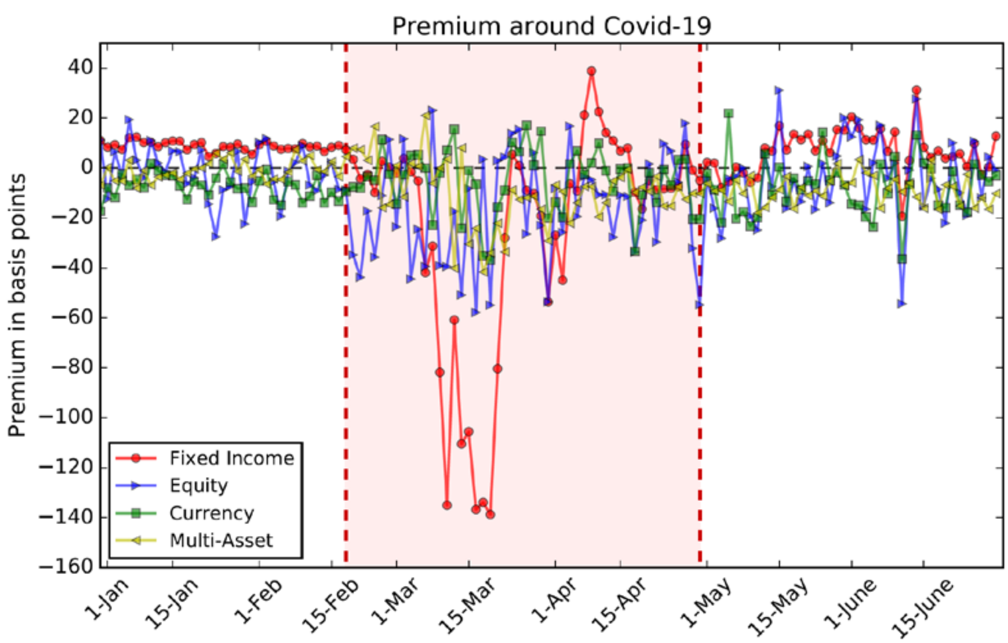


Figure 8. ETF premium during the COVID-19 pandemic

Panel A (Panel B) of this figure shows the average absolute (signed) premium for ETFs tracking different assets from January 1, 2020 to June 30, 2020. The shaded area denotes the period around the onset of COVID-19 pandemic period, and runs from February 20, 2020 to April 30, 2020.



Panel A: Average Absolute Premium for Different Types of ETFs



Panel B: Average Signed Premium for Different Types of ETFs

Table 1. Summary Statistics

This table shows summary statistics for the main variables. Raw |*Premium*| is the absolute value of the ETF premium in bps. |*Premium*| is the residual |*Premium*|, which is the residual from a regression of |*Premium*| on the non-LMM |*Premium*|. LMM (raw) |*Premium*| is the equally-weighted average (raw) |*Premium*| of all ETFs sharing the same LMM (in bps.). *Log (Size)* is the natural logarithm of an ETF's market capitalization. *Turnover* is an ETF's daily dollar trading volume scaled by its market capitalization (in percent). *STD* is the standard deviation of daily ETF return. *BidAsk* is the difference between ask and bid quotes scaled by the average of bid and ask quotes (in bps). *Turnover, STD,* and *BidAsk* are estimated using daily observations from last 30 days. We winsorize the continuous variables at the 1% and 99% levels. See Appendix B for variable definitions.

Summary Statistics of Main Variables						
Variable	*N*	Mean	Std	Q1	Median	Q3
raw |*Premium*| (bps)	2,946,280	25.48	31.99	4.44	12.59	33.21
|*Premium*| (bps)	2,946,280	-0.02	22.80	-10.46	-1.90	5.81
LMM raw |*Premium*| (bps)	2,946,280	25.74	9.51	19.20	24.81	31.04
LMM |*Premium*| (bps)	2,946,280	-0.02	3.33	-1.96	-0.28	1.62
Log (Size)	2,946,280	18.81	2.27	17.16	18.77	20.38
STD (percent)	2,946,280	0.83	0.60	0.44	0.72	1.09
BidAsk (bps)	2,946,280	0.24	0.27	0.06	0.14	0.30
Turnover (precent)	2,946,280	1.36	2.41	0.38	0.68	1.28

Table 2. LMM-level comovement in ETFs' raw and residual |Premium|

This table reports the regression estimates of ETFs' daily raw |*Premium*| and residual |*Premium*| on the equally weighted average raw |*Premium*| and residual |*Premium*| for all ETFs sharing the same LMM (excluding the focal ETF). Columns (1) - (4) show the results for ETFs' raw |*Premium*|, while columns (5) - (8) show the results for ETF's residual |*Premium*|. Controls include the contemporaneous non-LMM (raw) |*Premium*|, ETF size, ETF turnover, ETF bid-ask spread, and ETF return volatility. All independent variables are standardized with a mean of zero and a standard deviation of one. We include ETF and *Style* × *Day* fixed effects as indicated. Standard errors are double clustered at the ETF and Day level. The *t*-statistics are in parentheses. See Appendix B for variable definitions. The sample period is from January 1, 2012, to December 31, 2020.

	Dep.Var. = *Raw* |*Premium*|			
	(1)	(2)	(3)	(4)
non-LMM raw |*Premium*| (a)	0.954			
	(2.69)			
LMM raw |*Premium*| (b)	8.343	6.847	4.039	1.939
	(20.21)	(29.52)	(10.40)	(10.65)
Log (Size)		-3.320	-0.867	-3.171
		(-9.94)	(-2.60)	(-9.08)
STD		1.281	3.305	2.168
		(7.01)	(6.96)	(9.74)
BidAsk		6.696	10.379	6.848
		(25.22)	(28.87)	(25.72)
Turnover		0.206	2.432	0.500
		(1.32)	(6.63)	(3.68)
Style×*Day* FE	N	N	Y	Y
ETF FE	N	Y	N	Y
Observations	2,946,280	2,946,280	2,916,565	2,916,565
R-squared	7.80%	43.30%	29.4%	47.2%
(b)-(a)	7.388			
F-stat.	(10.13)			

	Dep.Var. = |*Premium*|			
	(5)	(6)	(7)	(8)
non-LMM |*Premium*| (a)	-0.060			
	(-0.69)			
LMM |*Premium*| (b)	2.123	1.999	1.582	1.585
	(20.53)	(19.58)	(17.00)	(17.04)
Log (Size)		-3.027	0.768	-3.506
		(-9.85)	(7.41)	(-10.05)
STD		-0.350	0.751	2.058
		(-2.46)	(6.03)	(10.36)
BidAsk		6.209	2.713	6.515
		(24.02)	(18.19)	(24.76)
Turnover		0.292	0.389	0.352
		(1.98)	(5.43)	(2.37)
Style×*Day* FE	N	N	Y	Y
ETF FE	N	Y	N	Y
Observations	2,946,280	2,946,280	2,916,567	2,916,567
R-squared	0.8%	3.6%	6.0%	8.1%
(b)-(a)	2.184			
F-stat.	(19.83)			

Table 3. Tests of contagion in the extreme level of ETF |Premium| using a linear probability model

This table reports the probability of observing a high ETF |*Premium*| conditional on whether its *LMM* |*Premium*| also takes on a high value. The dependent variable is an indicator that equals one if the focal ETF's |*Premium*| is higher than the 90th percentile cutoff, and zero otherwise. To capture extreme values of *LMM* |*Premium*|, we create an indicator variable, *D(LMM* |*Premium*| *> 90th percentile)*, that equals one if the *LMM* |*Premium*| is above its 90th percentile cutoff, and zero otherwise. Other controls include ETF size, ETF turnover, ETF bid-ask spread, and ETF return volatility. All continuous independent variables are standardized with a mean of zero and a standard deviation of one. We include ETF and *Style* × *Day* fixed effects as indicated. Standard errors are double clustered at the ETF and Day level. The *t*-statistics are in parentheses. See Appendix B for variable definitions. The sample period is from January 1, 2012 to December 31, 2020.

	Dep. Var. = *Dummy (*|*Premium*| *>90th percentile)*			
	(1)	(2)	(3)	(4)
D (LMM |*Premium*| *> 90th percentile)*	0.057	0.050	0.027	0.024
	(11.05)	(13.79)	(6.40)	(7.07)
Log (Size)		-0.021	-0.002	-0.017
		(-6.94)	(-1.31)	(-5.29)
STD		0.010	0.019	0.017
		(10.85)	(8.39)	(9.52)
BidAsk		0.049	0.046	0.050
		(18.78)	(19.80)	(18.76)
Turnover		0.003	0.006	0.003
		(2.05)	(4.05)	(2.03)
Constant	0.047	0.048	0.049	0.049
	(26.70)	(123.38)	(40.22)	(117.57)
Style×*Day* FE	N	N	Y	Y
ETF FE	N	Y	N	Y
Observations	2,946,280	2,946,280	2,946,280	2,946,280
R-squared	0.3%	9.9%	8.1%	12.2%

Table 4. Comovement in |Premium| when ETFs change lead market makers

This table shows the comovement of the ETF |*Premium*| with the equally-weighted average |*Premium*| of ETFs served by their old and new LMMs. The sample includes ETF-Day observations within the [-120, 120] trading days around the 1,264 events when an ETF changes its LMM. We include the same set of controls as in Table 2. LMM_{Old} $|Premium|$ (LMM_{New} $|Premium|$) is the equally-weighted average |*Premium*| of ETF i's previous (new) LMM, excluding ETF i. *Post* is a dummy variable that equals one for the period after an ETF changes its LMM. We also control for aggregate funding constraints, including *VIX*, *CS*, and *HKM*, the returns on the five Fama-French factors, and the ten Fama-French industries as indicated. All independent variables are standardized with a mean of zero and a standard deviation of one. The standard errors are clustered at the event levels. The *t*-statistics are in parentheses. See Appendix B for variable definitions.

	Dep. Var. = |*Premium*|				
	(1)	(2)	(3)	(4)	(5)
LMM_{Old} $\|Premium\|$	1.196	1.229	1.195	1.169	1.171
	(4.26)	(4.91)	(4.47)	(4.03)	(3.95)
$Post$×LMM_{Old} $\|Premium\|$	-0.923	-0.986	-0.977	-1.030	-1.032
	(-3.48)	(-3.94)	(-4.16)	(-4.29)	(-4.28)
LMM_{New} $\|Premium\|$	0.288	0.26	0.245	0.244	0.242
	(1.42)	(1.33)	(1.28)	(1.26)	(1.24)
$Post$×LMM_{New} $\|Premium\|$	1.444	1.550	1.547	1.479	1.467
	(2.84)	(2.88)	(2.86)	(2.99)	(2.95)
Post	0.351	0.788	0.478	0.532	0.551
	(0.98)	(1.28)	(0.63)	(0.73)	(0.77)
Log (Size)		1.546	1.514	1.511	1.511
		(4.27)	(4.14)	(4.12)	(4.12)
STD		0.346	0.323	0.331	0.332
		(1.39)	(1.39)	(1.38)	(1.38)
BidAsk		3.236	3.230	3.233	3.233
		(7.58)	(7.65)	(7.68)	(7.69)
Turnover		-0.145	-0.146	-0.148	-0.149
		(-0.83)	(-0.89)	(-0.91)	(-0.91)
Controls of Aggregate Funding Constraints			Y	Y	Y
FF 5 factors				Y	Y
FF 10 Industries					Y
Observations	189,471	189,432	189,432	189,432	189,432
R-squared	0.5%	2.0%	2.0%	2.0%	2.0%

Table 5. Cross-sectional heterogeneity: Arbitrage costs of ETFs and ETFs' constituents

This table reports the comovement in ETF premia conditional on proxies measuring the arbitrage costs of ETFs and ETFs' constituents. In columns (1) to (3), the coefficients of interest are the interaction between *LMM* |*Premium*| and ETF characteristics, including the natural logarithm of ETF market capitalization (column (1)), ETF return volatility (column (2)), and ETF bid-ask spread (column (3)). Columns (4) to (6) restrict the sample to US equity ETFs, for which we can measure arbitrage costs of the ETF's constituent stocks. *Spread CS* is the bid-ask spread calculated following the method of Corwin and Schultz (2012). *Volatility* is the daily stock return volatility within a month. Lendable supply (*Supply*) is the lendable shares from Markit divided by total shares outstanding. All independent variables are standardized with a mean of zero and a standard deviation of one. ETF and *Style* × *Day* fixed effects are included. Standard errors are double clustered at the ETF and Day level. The *t*-statistics are in parentheses. See Appendix B for variable definitions. The sample is from January 1, 2012 to December 31, 2020.

	Dep. Var. = |*Premium*|					
	(1)	(2)	(3)	(4)	(5)	(6)
LMM |*Premium*|	1.706	1.684	1.721	0.316	0.317	0.340
	(17.96)	(18.42)	(18.20)	(4.94)	(4.96)	(5.02)
Log (Size)×*LMM* |*Premium*|	-0.285					
	(-3.23)					
STD×*LMM* |*Premium*|		0.233				
		(3.24)				
BidAsk×*LMM* |*Premium*|			0.286			
			(3.16)			
Spread CS×*LMM* |*Premium*|				0.136		
				(2.46)		
Volatility×*LMM* |*Premium*|					0.122	
					(2.35)	
Supply×*LMM* |*Premium*|						-0.252
						(-2.72)
Controls	Y	Y	Y	Y	Y	Y
Style×*Day* FE	Y	Y	Y	Y	Y	Y
ETF FE	Y	Y	Y	Y	Y	Y
Observations	2,946,280	2,946,280	2,946,280	842,002	842,002	806,818
R-squared	5.9%	5.9%	5.9%	8.8%	8.8%	8.8%

Table 6. LMM-specific capital constraints based on ETF characteristics

This table reports the effects of LMM-specific capital constraints on comovement in ETF premia. In the regression for column (1), the LMM-specific capital constraint is measured by the natural logarithm of total market capitalization of all ETFs managed by the LMM (*Log(Mktcap of ETFs)*). In the regression for column (2), *Log(1/#Active APs)* inversely measures the number of other APs available to alleviate LMM constraints. It is the natural logarithm of one over the number of active APs, estimated using SEC N-CEN filings data from last fiscal year. Active APs are those APs that create or redeem shares for the ETF at any point in time. Other controls are the same as those in Table 2. ETF and *Style* × *Day* fixed effects are included. All independent variables are standardized with a mean of zero and a standard deviation of one. The *t*-statistics are in parentheses. See Appendix B for variable definitions. In column (1), the sample period is from January 1, 2012 to December 31, 2020. In column (2), the sample period is from July 1, 2017 to December 31, 2020.

Dep. Var.= |*Premium*|		
	(1)	(2)
LMM |Premium|	1.724	1.576
	(17.58)	(11.34)
Log (Mktcap of ETFs)	0.158	
	(1.07)	
Log (Mktcap of ETFs) ×LMM |*Premium*|	0.319	
	(8.62)	
Log (1/#Active APs)		0.545
		(1.48)
Log (1/#Active APs) ×LMM |*Premium*|		0.365
		(3.15)
Controls	Y	Y
Style×*Day* FE	Y	Y
ETF FE	Y	Y
Observations	2,916,567	664,972
R-squared	8.1%	13.5%

Table 7. LMM-specific capital constraints based on LMM characteristics

This table reports the effects of time-series variation in LMM-specific capital constraints on comovement in ETF premia. In columns (1) to (3), we use the change in LMMs' net worth during a window around earnings announcements to capture negative shocks to capital constraints. Specifically, we define the LMM-specific constraint as a dummy that equals to one for the three-day window [-1, 1] around LMM's quarterly earnings announcements when the market-adjusted announcement abnormal return is less than -3%, and zero otherwise. The sample is restricted to the 21-day window [-10, 10] around quarterly earnings announcements for all publicly traded LMMs, including Credit Suisse, Deutsche Bank, Goldman Sachs, and Virtu Financial. In columns (4) to (6), the LMM-specific capital constraint is captured by the *net capital ratio*, which measures the amount of capital an LMM holds relative to CFTC's net capital requirement. A larger value of *net capital ratio* indicates that the LMM is more capital constrained. Information on net capital ratios is available for five LMMs including Cantor Fitzgerald, Credit Suisse, Deutsche Bank, Goldman Sachs, and RBC Capital Markets. We include the same set of controls as in the regressions for Table 4. The standard errors are double clustered at the ETF and day levels. The *t*-statistics are in parentheses. See Appendix B for variable definitions.

	Dep. Var. = |*Premium*|					
	D (Earnings Announcement Returns <-3%)			*Net Capital Ratio*		
	(1)	(2)	(3)	(4)	(5)	(6)
LMM |Premium|×Constraint	0.528	0.519	0.487	0.574	0.564	0.559
	(2.00)	(2.02)	(1.95)	(4.34)	(4.30)	(4.25)
Constraint	0.224	0.201	0.243	0.230	0.225	0.222
	(1.50)	(1.30)	(1.56)	(1.87)	(1.83)	(1.79)
LMM |Premium|	1.423	1.400	1.364	1.670	1.663	1.658
	(10.22)	(9.82)	(10.11)	(13.09)	(13.09)	(13.03)
Log (Size)	1.020	1.020	1.022	1.003	1.004	1.005
	(8.19)	(8.17)	(8.18)	(5.07)	(5.08)	(5.08)
STD	0.518	0.521	0.524	0.256	0.261	0.263
	(4.18)	(4.21)	(4.23)	(1.86)	(1.89)	(1.91)
BidAsk	2.417	2.419	2.425	2.855	2.857	2.859
	(15.38)	(15.40)	(15.43)	(9.90)	(9.90)	(9.90)
Turnover	0.252	0.251	0.251	0.208	0.208	0.207
	(3.35)	(3.34)	(3.33)	(2.06)	(2.05)	(2.04)
Controls of aggregate funding constraints	Y	Y	Y	Y	Y	Y
FF5F		Y	Y		Y	Y
FF10 industries			Y			Y
Observations	104,596	104,596	104,596	798,152	798,152	798,152
R-squared	1.30%	1.30%	1.30%	1.70%	1.70%	1.70%

Table 8. DiD analysis of ETF |Premium| during COVID-19 pandemic

This table reports the results from a difference-in-differences estimation of ETF |*Premium*| around the COVID-19 pandemic. The sample includes only non-fixed income ETFs. The sample period is from January 1, 2020, to June 30, 2020. *COVID* is a dummy variable that equals one for period from February 21, 2020, to April 30, 2020. We use two proxies to capture LMM's fixed income exposure: *FI Weight* is calculated as the market capitalization of fixed income ETFs managed by the focal ETF's LMM scaled by the total market capitalization of all ETFs managed by the LMM. We calculate *FI Weight* for each LMM based on the observations in December 2019. *D(FI Weight>Median)* is a dummy variable that equals one if the *FI Weight* is above sample median, and zero otherwise. The coefficient of interest is the interaction of *COVID* and *FI Weight* (or *D(FI Weight>Median)*), which identifies the mispricing of ETFs managed by LMMs with large exposure to fixed income ETFs with that of ETFs managed by LMMs with smaller exposure to fixed income ETFs during the COVID-19 pandemic period. In columns (4) and (7), we create half-month dummies around the COVID-19 outbreak. $COVID_{t-15,t}$ is a dummy variable that equals one from 15 days before the COVID-19 outbreak to the start of outbreak on February 20, 2020, and zero otherwise. $COVID_{t+1,t+15}$ and $COVID_{t+16,t+31}$ are defined similarly, and $COVID_{t+32,End}$ equals one from 32 days after February 20, 2020 to the end of the first acute phase of the COVID-19 pandemic on April 30, 2020. We also include the interactions of fixed income exposure with COVID dummies. All the continuous control variables are standardized with a mean of zero and a standard deviation of one. We include ETF and *Style* × *Day* fixed effects as indicated. The *t*-statistics are in parentheses.

Dep. Var. = *Raw* |*Premium*|							
		X = *FI Weight*			X = *D(FI Weight>Median)*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
COVID	21.009	17.715			18.118		
	(7.92)	(7.52)			(7.68)		
X×COVID		11.973	4.879		4.558	2.237	
		(3.88)	(1.80)		(3.56)	(2.03)	
$X \times COVID_{t-15,t}$				-1.457			-0.734
				(-0.79)			(-1.00)
$X \times COVID_{t+1,t+15}$				9.203			3.846
				(2.76)			(3.08)
$X \times COVID_{t+16,t+31}$				9.771			4.413
				(1.56)			(2.28)
$X \times COVID_{t+32,End}$				0.994			0.629
				(0.39)			(0.76)
Log (Size)			-3.425	-3.435		-3.317	-3.360
			(-1.63)	(-1.63)		(-1.57)	(-1.87)
STD			-0.807	-0.788		-0.774	-0.769
			(-0.63)	(-0.62)		(-0.61)	(-0.95)
BidAsk			3.470	3.530		3.549	3.533
			(6.05)	(6.12)		(6.21)	(6.49)
Turnover			1.000	0.979		0.985	0.969
			(2.68)	(2.63)		(2.65)	(2.72)
ETF FE	Y	Y	Y	Y	Y	Y	Y
Style×Day FE			Y	Y		Y	Y
Observations	152,170	152,170	150,889	150,889	152,170	150,889	150,889
R-squared	43.8%	43.9%	55.2%	55.2%	43.8%	55.2%	55.2%

Table 9. Subperiod analysis conditional on aggregate funding constraints

This table reports subperiod results stratified by levels of aggregate funding constraints. The regression specification follows Table 2. We divide the sample into halves based on *VIX*, credit spread (*CS*), the intermediary capital ratio (*HKM*) of He, Kelly, and Manela (2017), and the stock market return in the prior month (*Mktret*). High (Low) indicates periods with tightened (loosened) aggregate funding constraints. Up (Down) means market states with monthly return above (below) sample median. All independent variables are standardized with a mean of zero and a standard deviation of one. See Appendix B for variable definitions. The *t*-statistics are in parentheses. The sample period is from January 1, 2012 to December 31, 2020.

	Dep. Var. = |*Premium*|							
	VIX		*Credit Spread*		*HKM*		*Mktret*	
	Low	High	Low	High	Low	High	Down	Up
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LMM |*Premium*|	1.517	1.618	1.583	1.540	1.540	1.595	1.642	1.522
	(15.51)	(16.48)	(15.44)	(15.38)	(14.89)	(15.94)	(16.70)	(15.34)
Log (Size)	-2.722	-4.193	-3.278	-3.996	-3.511	-3.771	-3.646	-3.396
	(-6.42)	(-11.29)	(-7.86)	(-9.30)	(-8.36)	(-8.96)	(-9.76)	(-9.03)
STD	2.748	1.732	2.196	2.138	1.638	2.492	2.335	1.682
	(11.00)	(10.20)	(11.25)	(10.28)	(7.72)	(12.46)	(9.79)	(6.89)
BidAsk	7.294	5.792	6.259	6.725	6.343	6.409	6.492	6.580
	(22.39)	(21.40)	(21.20)	(19.71)	(18.59)	(21.87)	(24.13)	(21.48)
Turnover	0.358	0.288	0.381	0.254	0.265	0.706	0.493	0.239
	(1.75)	(2.12)	(2.28)	(1.45)	(1.79)	(3.89)	(3.06)	(1.55)
Style×*Day* FE	Y	Y	Y	Y	Y	Y	Y	Y
ETF FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	1,497,200	1,419,367	1,657,171	1,259,394	1,215,171	1,701,396	1,463,195	1,453,372
R-squared	8.9%	9.4%	7.9%	10.7%	11.3%	8.1%	8.5%	8.4%

Table 10. Co-movement in |Premium| for ETFs tracking the same or different benchmark (style)

This table presents the ETF premium comovement effect among ETFs following the same or different benchmark or style. In column (1), we construct *Non-LMM* $|Premium|^{same\ benchmark}$ as the equally weighted average $|Premium|$ of all ETFs managed by a different LMM but tracks the same benchmark index as the focal ETF. In column (2), we decompose *LMM* $|Premium|$ into *LMM* $|Premium|^{w/o\ same\ style}$ and *LMM* $|Premium|^{with\ same\ style}$, by separately computing the equally-weighted average $|Premium|$ for ETFs with and without the same style category as the focal ETF. All independent variables are standardized with a mean of zero and a standard deviation of one. See Appendix B for variable definitions. The *t*-statistics are in parentheses. The sample period is from January 1, 2012 to December 31, 2020.

Dep. Var. = $\|Premium\|$		
	(1)	(2)
LMM $\|Premium\|$	0.748	
	(4.73)	
Non-LMM $\|Premium\|^{same\ benchmark}$	-1.135	
	(-0.76)	
LMM $\|Premium\|^{w/o\ same\ style}$		1.301
		(14.35)
LMM $\|Premium\|^{with\ same\ style}$		0.638
		(3.07)
Log (Size)	-3.310	-3.593
	(-3.39)	(-10.14)
STD	1.666	2.073
	(3.55)	(10.36)
BidAsk	7.444	6.531
	(12.22)	(24.42)
Turnover	-0.267	0.366
	(-0.46)	(2.45)
Style×*Day* FE	Y	Y
ETF FE	Y	Y
Observations	334,836	2,871,587
R-squared	18.2%	7.9%

Table 11. Alternative specifications of fixed effects

This table presents results of the baseline regression model in Table 2 with additional fixed effects. In column (1), we further add a region × day fixed effect to the baseline specification, where region refers to the geographical focus of the ETF. Column (2) adds an exchange × day fixed effect, where exchange denotes the exchange on which the ETF is listed. Columns (3) and (4) add issuer × day and distributor × day fixed effects respectively. All independent variables are standardized with a mean of zero and a standard deviation of one. See Appendix B for variable definitions. The *t*-statistics are in parentheses. The sample period is from January 1, 2012 to December 31, 2020.

	Dep. Var. = |*Premium*|			
	Region×Day FE	Exchange×Day FE	Issuer×Day FE	Distributor×Day FE
	(1)	(2)	(3)	(4)
LMM |Premium|	1.422	1.667	1.339	1.542
	(17.51)	(17.23)	(13.84)	(16.14)
Log (Size)	-3.379	-3.174	-3.335	-3.473
	(-9.99)	(-8.96)	(-9.23)	(-9.79)
STD	2.079	2.126	2.069	2.014
	(10.73)	(10.56)	(10.51)	(10.24)
BidAsk	6.613	6.576	6.503	6.543
	(25.15)	(24.04)	(24.95)	(24.59)
Turnover	0.243	0.338	0.378	0.324
	(1.66)	(2.27)	(2.56)	(2.11)
ETF FE	Y	Y	Y	Y
Style×Day FE	Y	Y	Y	Y
Observations	2,944,230	2,762,153	2,903,631	2,885,231
R-squared	11.6%	6.7%	12.3%	8.1%

Table 12. Comovement in |Premium| for ETFs covering different regions

This table reports the estimates from regressions of ETF daily |*Premium*| on the equally-weighted average |*Premium*| of all ETFs sharing the same LMM. The sample includes U.S.-listed ETFs covering different geographic regions. We control for the contemporaneous ETF size, turnover, return volatility, and the bid-ask spread. All independent variables are standardized with a mean of zero and a standard deviation of one. ETF and *Style* × *Day* fixed effects are included in the regressions. Standard errors are double clustered at ETF and Day level. The *t*-statistics are in parentheses. See Appendix B for variable definitions. The sample is from January 1, 2012 to December 31, 2020.

	Dep.Var. = |*Premium*|						
	Emerging Markets	Developed Markets	Asia-Pacific	Europe	Global Ex-U.S.	Global	North America
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
LMM |Premium|	3.291	2.761	3.558	2.690	2.410	1.273	0.375
	(6.19)	(6.48)	(8.21)	(6.45)	(5.48)	(8.49)	(6.90)
Log (Size)	-2.538	-5.805	-1.604	-3.955	-3.679	-2.567	-3.165
	(-1.33)	(-4.88)	(-1.23)	(-2.46)	(-2.83)	(-3.37)	(-7.36)
STD	1.719	0.548	5.206	4.377	1.965	1.742	0.512
	(1.63)	(0.39)	(9.10)	(4.79)	(1.88)	(4.43)	(2.50)
BidAsk	4.920	8.629	5.613	5.868	6.387	6.315	7.855
	(7.64)	(11.39)	(6.78)	(8.34)	(8.05)	(13.94)	(15.19)
Turnover	0.736	0.531	1.589	1.963	-0.089	0.234	-0.066
	(0.98)	(0.67)	(3.69)	(4.19)	(-0.14)	(0.43)	(-0.47)
Style×*Day* FE	Y	Y	Y	Y	Y	Y	Y
ETF FE	Y	Y	Y	Y	Y	Y	Y
Observations	139,804	121,393	210,767	140,852	153,003	518,523	1,548,140
R-squared	22.4%	28.9%	19.0%	23.7%	19.6%	15.2%	13.7%

Table 13. Comovement in |Premium| for ETFs tracking different asset classes

This table presents results of the baseline regression model of Table 2 for ETFs tracking different assets (columns (1) to (6)) and for leveraged/inverse ETF (column (7)). We regress ETF |*Premium*| on the equally-weighted average |*Premium*| of all ETFs sharing the same LMM (excluding the focal ETF itself). Controls include ETF size, turnover, bid-ask spreads, and return volatility. All independent variables are standardized with a mean of zero and a standard deviation of one. See Appendix B for variable definitions. We include ETF and *Style* × *Day* fixed effects as indicated. Standard errors are double clustered at the ETF and Day level. The *t*-statistics are in parentheses. The sample period is from January 1, 2012 to December 31, 2020.

	Dep. Var. = |*Premium*|						
	Commodities	Currencies	Equities	Fixed income	Real estates	Multi-asset	Leveraged
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
LMM |*Premium*|	0.826	0.687	1.781	0.832	0.941	2.157	1.163
	(2.20)	(1.80)	(15.41)	(5.56)	(3.53)	(3.54)	(2.16)
Log (Size)	-0.846	-0.470	-3.834	-2.905	-0.558	-4.958	3.331
	(-0.42)	(-0.30)	(-10.41)	(-2.68)	(-0.41)	(-3.04)	(1.06)
STD	5.649	4.353	1.957	0.97	3.042	2.215	2.117
	(4.21)	(2.49)	(9.24)	(1.03)	(3.77)	(2.64)	(2.59)
BidAsk	5.964	4.255	6.181	8.528	6.220	5.870	7.309
	(5.28)	(4.91)	(19.90)	(13.75)	(6.43)	(4.49)	(13.47)
Turnover	0.394	0.316	0.660	-0.23	-0.859	-0.482	-0.934
	(0.45)	(1.23)	(4.07)	(-0.74)	(-0.63)	(-0.79)	(-1.89)
ETF FE	Y	Y	Y	Y	Y	Y	Y
Style×*Day* FE	Y	Y	Y	Y	Y	Y	Y
Observations	51,759	9,597	2,121,627	531,418	129,374	71,459	302,021
R-squared	38.8%	31.2%	5.1%	17.5%	9.3%	9.0%	34.4%

Internet Appendix to "Financial Intermediaries and Contagion in Market Efficiency: The Case of ETFs"

Table IA.1. List of lead makers

This table lists the information about the lead market makers (LMMs) in our sample. We first calculate the number and size (in billions USD) of ETFs, as well as the equally weighted average (raw) |*Premium*| of ETFs managed by each LMM at the daily level. We then report the time series average statistics of these variables from January 2012 to December 2020.

List of Lead Market Makers				
LMM	#ETF	Size (billion USD)	Raw |Premium|	|Premium|
Goldman Sachs	280	634.5	20.56	-0.97
KCG	364	489.5	30.92	1.41
Virtu Financial	203	377.6	17.25	-0.43
Jane Street	209	316.3	33.99	1.75
Susquehanna	215	302.6	29.06	0.16
IMC Chicago	105	204.2	16.71	0.42
Cantor Fitzgerald	109	122.2	26.31	0.35
Latour Trading	25	97.1	5.08	-0.05
Pundion	23	74.2	22.24	2.40
Credit Suisse	38	63.0	21.48	-1.13
RBC Capital Markets	32	37.7	19.17	0.15
Citadel	22	32.6	14.80	-0.78
Deutsche Bank	19	17.3	19.11	-0.32
Flow Traders	10	12.9	26.15	0.73
Societe Generale	9	6.4	23.78	-2.81
Wolverine Trading	5	4.1	19.01	-0.21
CLP	3	1.4	48.38	0.77
C&C Trading	4	1.0	31.91	-4.72

Table IA.2. Comovement in |Premium| when ETFs change the lead market makers (excluding the acquisition of KCG by Virtu)

This table replicates Table 4 on the comovement effect for ETF |*Premium*| served by their previous and new LMMs, by excluding the 476 events driven by the acquisition of KCG by Virtu. The sample includes ETF-Day observations within the [-120, 120] trading days around the events when an ETF changes its LMM. We include the same set of controls as in Table 2. LMM_{Old} $|Premium|$ (LMM_{New} $|Premium|$) is the equally weighted average |*Premium*| of ETF *i*'s old (new) LMM, excluding ETF *i*. Old (new) LMM is the LMM before (after) the ETF changes its LMM. *Post* is a dummy variable that equals one for the period after an ETF changes its LMM. We also control for aggregate funding constraints, including *VIX*, *CS*, and *HKM*, the returns on the five Fama-French factors, and the ten Fama-French industries as indicated. All independent variables are standardized with a mean of zero and a standard deviation of one. The standard errors are clustered at the event levels. The *t*-statistics are in parentheses. See Appendix B for variable definitions.

	Dep. Var. = |*Premium*|				
	(1)	(2)	(3)	(4)	(5)
LMM_{Old} $\|Premium\|$	1.241	1.230	1.200	1.179	1.189
	(3.61)	(4.13)	(3.87)	(3.64)	(3.62)
$Post$×LMM_{Old} $\|Premium\|$	-0.968	-0.997	-0.989	-1.049	-1.059
	(-2.88)	(-3.27)	(-3.33)	(-3.44)	(-3.52)
LMM_{New} $\|Premium\|$	0.305	0.247	0.233	0.235	0.230
	(1.54)	(1.23)	(1.18)	(1.19)	(1.18)
$Post$×LMM_{New} $\|Premium\|$	1.428	1.566	1.573	1.496	1.483
	(2.82)	(2.93)	(2.93)	(3.10)	(3.09)
Post	0.221	0.484	0.195	0.251	0.274
	(0.46)	(0.67)	(0.25)	(0.34)	(0.38)
Log (Size)		1.194	1.153	1.147	1.150
		(3.54)	(3.56)	(3.53)	(3.53)
STD		0.543	0.545	0.556	0.559
		(1.81)	(1.88)	(1.83)	(1.83)
BidAsk		3.012	3.015	3.020	3.021
		(5.25)	(5.33)	(5.37)	(5.37)
Turnover		-0.184	-0.192	-0.194	-0.196
		(-0.78)	(-0.86)	(-0.86)	(-0.87)
Controls for Aggregate Funding Constraints			Y	Y	Y
FF 5 factors				Y	Y
FF 10 Industries					Y
Observations	135,420	135,386	135,386	135,386	135,386
R-squared	0.70%	2.00%	2.10%	2.10%	2.10%

Table IA.3. Subsample analysis to address concerns about missing LMMs

This table reports the subsample estimates of the specification in Table 2. Columns (1)-(4) report the estimates for observations up to the year 2018. Columns (5)-(8) report the estimates for ETFs with an age longer than 24 months. All the independent variables are standardized with a mean of zero and a standard deviation of one. We include ETF and *Style* × *Day* fixed effects as indicated. Standard errors are double clustered at the ETF and Day level. The *t*-statistics are in parentheses. See Appendix B for variable definitions. The sample period is from January 1, 2012 to December 31, 2020.

	Dep. Var. = |*Premium*|							
	Year<=2018				Age>24 months			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
non-LMM |Premium| (a)	-0.039				0.040			
	(-0.41)				(0.46)			
LMM |Premium| (b)	2.116	2.006	1.513	1.577	1.968	1.912	1.370	1.417
	(19.18)	(19.41)	(15.82)	(16.72)	(18.04)	(17.28)	(14.35)	(14.73)
Log (Size)		-3.265	0.376	-3.240		-1.715	1.522	-2.690
		(-9.45)	(2.87)	(-7.96)		(-5.48)	(11.93)	(-7.37)
STD		1.024	0.733	2.901		-0.335	0.988	1.972
		(7.01)	(4.65)	(12.23)		(-2.35)	(6.88)	(9.37)
BidAsk		6.991	2.829	7.134		6.153	2.857	6.445
		(23.69)	(16.11)	(23.93)		(21.13)	(15.65)	(21.32)
Turnover		0.297	0.302	0.321		0.334	0.260	0.311
		(1.77)	(3.56)	(1.89)		(2.48)	(3.59)	(2.38)
Style×Day FE	N	N	Y	Y	N	N	Y	Y
ETF FE	N	Y	N	Y	N	Y	N	Y
Observations	2,184,402	2,184,402	2,160,533	2,160,533	2,367,745	2,367,732	2,338,205	2,338,192
R-squared	0.8%	5.3%	5.6%	9.1%	0.7%	4.5%	7.1%	10.1%
(b)-(a)	2.155				1.928			
F-stat.	(19.13)				(16.19)			

Table IA.4. Comovement in positive vs. negative ETF premium

This table reports estimates from regressing the ETF daily positive (negative) premium on the equally weighted average positive (negative) premium of all ETFs sharing the same LMM (LMM Premium). *Positive premium* is calculated as the ETF premium when the sign of the premium is positive, and zero otherwise. Similarly, *negative premium* is calculated as the absolute value of ETF premium when the sign of premium is negative, and zero otherwise. *LMM positive (negative) premium* is calculated as the equally weighted average *positive (negative) premium* of all ETFs sharing the same lead market maker, excluding the focal ETF itself. The controls are the same as in Table 2. We include ETF and *Style* × *Day* fixed effects as indicated. Standard errors are double clustered at ETF and Day level. The *t*-statistics are in parentheses. See Appendix B for variable definitions. The sample is from January 1, 2012 to December 31, 2020.

Dep.Var. = *Positive Premium*					Dep.Var. = *Negative Premium*				
	(1)	(2)	(3)	(4)		(5)	(6)	(7)	(8)
non-LMM Positive Premium (a)	-0.957				*non-LMM Negative Premium* (a)	-0.456			
	(-2.84)					(-1.16)			
LMM Positive Premium (b)	6.291	3.149	3.559	1.099	*LMM Negative Premium* (b)	6.398	3.978	2.630	0.823
	(16.82)	(17.45)	(10.47)	(8.03)		(15.36)	(18.58)	(7.82)	(5.21)
Log (Size)		-4.444	-1.061	-3.245	*Log (Size)*		-5.360	-1.486	-3.322
		(-13.60)	(-3.17)	(-9.27)			(-15.73)	(-4.38)	(-9.45)
STD		3.433	3.934	2.217	*STD*		2.157	3.905	2.210
		(12.43)	(8.23)	(9.96)			(10.39)	(8.12)	(9.94)
BidAsk		7.120	10.497	6.857	*BidAsk*		6.694	10.343	6.822
		(26.84)	(28.78)	(25.56)			(24.99)	(28.00)	(25.42)
Turnover		0.364	1.294	0.291	*Turnover*		0.355	1.240	0.309
		(2.29)	(3.56)	(1.93)			(2.20)	(3.41)	(2.03)
Style×*Day* FE	N	N	Y	Y	*Style*×*Day* FE	N	N	Y	Y
ETF FE	N	Y	N	Y	ETF FE	N	Y	N	Y
Observations	2,946,280	2,946,280	2,916,567	2,916,567	Observations	2,946,280	2,946,280	2,916,567	2,916,567
R-squared	3.3%	41.1%	28.9%	47.2%	R-squared	3.6%	4.2%	28.6%	47.2%
(b)-(a)	7.248				(b)-(a)	6.854			
F-stat.	(10.83)				*F*-stat.	(8.86)			

Table IA.5. Controlling for LMM |Premium| in the prior month and past ETF returns

This table reports the estimates from regressing the ETF daily raw (or residual) |*Premium*| on the equally weighted average raw (or residual) |*Premium*| for all ETFs sharing the same LMM (excluding the focal ETF itself). The specification follows that of Table 2, except that we further control for *LMM raw |Premium| (LMM |Premium| (m-1))*, which is calculated as the average LMM raw (or residual) |Premium| in the prior month. We also control for the ETF's return in the prior month (Ret_{m-1}) and cumulative returns from past 12 to 2 months ($Ret_{m-2,m-12}$). All the independent variables are standardized with a mean of zero and a standard deviation of one. We include ETF and *Style* × *Day* fixed effects as indicated. Standard errors are double clustered at the ETF and Day level. The *t*-statistics are in parentheses. See Appendix B for variable definitions. The sample period is from January 1, 2012 to December 31, 2020.

	Dep. Var. = *Raw* |*Premium*|					Dep. Var. = |*Premium*|			
	(1)	(2)	(3)	(4)		(5)	(6)	(7)	(8)
non-LMM raw |Premium| (a)	1.522				*non-LMM |Premium| (a)*	-0.115			
	(4.99)					(-1.26)			
LMM raw |Premium| (b)	7.173	7.602	3.074	2.447	*LMM |Premium| (b)*	2.057	2.016	1.655	1.577
	(18.20)	(29.24)	(9.51)	(10.95)		(19.17)	(19.97)	(17.77)	(17.22)
LMM raw |Premium| (m-1)	1.025	-2.851	1.390	-1.603	*LMM |Premium| (m-1)*	0.115	-0.103	-0.699	0.054
	(2.77)	(-10.95)	(2.59)	(-4.72)		(1.04)	(-0.70)	(-3.46)	(0.16)
Ret_{m-1}	-0.637	0.207	-4.792	0.673	Ret_{m-1}	0.026	0.323	0.981	2.034
	(-3.66)	(2.36)	(-2.33)	(0.56)		(0.56)	(3.48)	(0.91)	(1.63)
$Ret_{m-2,m-12}$	-1.650	0.491	-5.238	-0.029	$Ret_{m-2,m-12}$	-0.202	0.222	-0.831	0.19
	(-2.71)	(3.72)	(-2.77)	(-0.06)		(-1.95)	(2.67)	(-1.90)	(0.39)
Log (Size)		-4.275	-0.917	-3.212	*Log (Size)*		-3.125	0.800	-3.525
		(-12.89)	(-2.70)	(-8.86)			(-9.41)	(7.67)	(-9.79)
STD		0.977	3.812	2.232	*STD*		-0.199	0.780	2.087
		(5.47)	(7.73)	(9.69)			(-1.35)	(6.15)	(10.13)
BidAsk		6.532	10.355	6.808	*BidAsk*		6.184	2.723	6.491
		(23.95)	(28.17)	(25.01)			(23.60)	(17.88)	(24.33)
Turnover		0.154	1.387	0.292	*Turnover*		0.267	0.363	0.344
		(0.97)	(3.73)	(1.87)			(1.79)	(4.97)	(2.25)
Style×*Day* FE	N	N	Y	Y	*Style*×*Day* FE	N	N	Y	Y
ETF FE	N	Y	N	Y	ETF FE	N	Y	N	Y
Observations	2,870,933	2,870,933	2,841,669	2,841,669	Observations	2,870,933	2,870,933	2,841,669	2,841,669
R-squared	8.2%	43.6%	29.3%	47.4%	R-squared	0.008	0.035	6.0%	8.1%
(b)-(a)	5.651				(b)-(a)	2.174			
F-stat.	(9.16)				*F*-stat.	(19.72)			