The Unexpected Activeness of Passive Investors: A Worldwide Analysis of ETFs

Si Cheng
Chinese University of Hong Kong

Massimo Massa
INSEAD

Hong Zhang
PBCSF, Tsinghua University

The global ETF industry provides more complicated investment vehicles than low-cost index trackers. Instead, we find that the real investments of ETFs may deviate from their benchmarks to leverage informational advantages (which leads to a surprising stock-selection ability) and to help affiliated OEFs through cross-trading. These effects are more prevalent in ETFs domiciled in Europe. Moreover, ETF flows seem to respond to additional risk. These results have important normative implications for consumer protection and financial stability. (JEL G20)

Received March 18, 2017; Editorial decision October 14, 2018 by Editor Raman Uppal. Authors have furnished an Internet Appendix, which is available on the Oxford University Press Web site next to the link to the final published paper online.

> The speed and breadth of financial innovation in the ETF market has been remarkable . . ., and has brought new elements of complexity and opacity into this standardized market.
>
> —Financial Stability Board[1]

We thank two anonymous referees, Warren Bailey, Pier

Introduction

Over the last decade, the market has witnessed the rise of exchange-traded funds (ETFs). According to the Financial Stability Board (FSB) (2011), the global ETF industry experienced an astonishing 40% annual growth rate over the 10-year period from 2001 to 2010, compared with the 5% annual growth rate in global open-end mutual funds (OEFs) and equity markets over the same period. The press has extolled the benefits of ETFs as cheap alternatives to traditional OEFs and even to index funds because they offer index-tracking investment opportunities to investors with low cost (i.e., no load fees and extremely limited management fees). In short, ETFs have been heralded as the harbingers of a new era of low-cost/low-risk investment opportunities that are available to the general public.

However, this brief narrative does not tell the entire story. Indeed, to charge low fees, many ETF sponsors may seek alternative investment techniques of active management, such as synthetic replication with affiliated banks, active divergence from the benchmark, and security lending (e.g., Ramaswamy 2011). These techniques may create an implicit "investment link" between the ETFs and their affiliated financial conglomerate. To illustrate this point, consider a Nikkei index ETF that receives \$100 of investment. Instead of investing this money as required by the index, the ETF can invest the entire \$100 in a different type of risky equity portfolio and at the same time enter a total return swap with its affiliated bank, whereby it swaps the total return on the invested portfolio with the return on the index. In this way, although the ETF is able to track the benchmark at a lower cost, the actual portfolio allocation will deviate from the benchmark. Because ETF investors only require the index return, ETF investors may not fully enjoy the upside of the actual portfolio investment, although they may be exposed to additional risk if the swap counterparty defaults on the promised delivery of the index return.

These features have raised the concerns of regulators on the incentives of ETFs. For example, the Financial Services Authority has identified the potential for conflicts of interest as one of the major concerns and suggested that it is "extremely important" for ETF providers to properly highlight the difference between a straightforward ETF and "more complex investment strategies" that may involve derivatives (Flood 2012). Practitioners have voiced similar concerns. BlackRock stated that "it believes that potential conflicts of interest arise when a synthetic ETF provider enters into a derivative agreement with its investment banking parent because the costs it pays for the swap could be uncompetitive and beneficial to the bank" (Davies 2012).

In this paper, we investigate the incentives for ETFs to engage in active management based on the universe of worldwide equity ETFs and OEFs during the 2001–2009 period.[2] In particular, we want to explore whether and how incentives related to information, subsidization, and security lending may give rise to ETF activeness. Because the former two types of incentives involve off-benchmark strategies, we expect non-full-replicating ETFs to dominate the potential effect (as full replicating ETFs are prohibited from adopting such off-benchmark strategies). Both full replicating and non-full-replicating ETFs can, however, participate in the security lending market. To provide a complete picture on the average effect of the entire ETF industry, we first include the entire sample of ETFs in our main analyses, and then examine how different types of ETFs are exposed to different kinds of incentives in subsample analyses (e.g., our subsample tests confirm that the effects of information or subsidization incentives concentrate on non-full-replicating ETFs).

We start by documenting a surprising selection ability of (non-full-replicating) ETFs, even though these funds in public views are largely passive index trackers. More explicitly, we find strong evidence that ETFs deviate from their benchmarks in stocks that have a lending relationship with their affiliated banks (“bank loan channel”). Moreover, such deviations are highly informative: for stocks receiving corporate loan services from banks, each 1% increase in benchmark-adjusted ownership of ETFs affiliated with these banks (hereafter, bank loan-related abnormal ownership of ETFs) is related to a 12 bps higher Daniel et al. (1997) (DGTW)-adjusted return per year. By contrast, for the same stock, the abnormal ownership of ETFs unaffiliated with the banks providing loan services conveys no such information. The difference between the informativeness of ownership changes of affiliated ETFs and that of unaffiliated ETFs strongly suggests that ETFs can overweight (underweight) stocks that are somehow confirmed to be good (bad) via their affiliate banks’ corporate loan services.

It is interesting to compare the informativeness of ETF trading to that of mutual funds. Affiliations with banks are known to accrue information to active OEFs in the mutual fund industry (e.g., Irvine, Lipson, and Puckett 2007; Massa and Rehman 2008; Massa and Zhang 2012), which is in spirit close to our finding in the ETF industry. However, this seemingly similar observation could imply drastically different incentives and normative implications in the two industries. On the one hand, it is perhaps not surprising for active mutual funds to collect superior information in order to create better performance for their investors. On the other hand, the same intuition does not directly apply to the ETF industry, as ETFs do not have the fiduciary

[2] In 2009, the so-called “funded swap model” was introduced in Europe. In this model, the counterparty posts collateral assets in a segregated account with a third-party custodian. The account can be held either in the name of the fund (in the case of a transfer of title) or in the name of the counterparty and pledged in favor of the fund (in the case of a pledge arrangement). The first case might dilute the validity of holding information for our tests. Thus, we restrict our testing sample to 2009. Interested readers should refer to Morningstar (2012) for institutional details.

duty to generate index-adjusted performance for investors. Rather, any performance that is not delivered to investors will be ultimately paid back to affiliated banks, for instance through the aforementioned swap design. Moreover, unlike OEFs, most ETFs are affiliated with bank conglomerates (as opposed to specialized asset management companies). These institutional differences imply that ETFs could be exposed more to the incentive to help affiliated financial conglomerates, which we can term as "proconglomerate incentives," than to the incentive of delivering superior performance to investors, which we label "proinvestor incentives." It is crucial to ask whether the surprising selection ability of ETFs reflects the former incentives because, if so, regulators and practitioners may have grounds to worry about the role played by ETFs in the financial market.

We further analyze three questions in order to better understand the incentives associated with ETFs. First, how pervasive are proconglomerate incentives; that is, is ETF activeness limited only to the collection of information through the bank loan channel, or could there be other mechanisms leading ETFs to deviate from their public image of index tracker as well? Second, do investors benefit or suffer from these mechanisms? Note that even though ETFs deliver index return in general, they can still share additional benefits with the investors through reductions in fees. In other words, proconglomerate and proinvestor incentives are not necessarily mutually exclusive. Finally, how do investors respond to ETF incentives? Proconglomerate incentive does not necessarily mean conflicts of interest as long as investors do not get hurt. But do investors nonetheless worry about such incentives? These questions are crucial to understand the potential opportunities and risks associated with the ETF industry.

Regarding the first question, we find that proconglomerate incentives are quite widely observed through several distinctive mechanisms. Similar to what we have observed in the bank loan channel, ETF investment in the stocks of the affiliated bank is related to better performance: an increase in abnormal ownership of affiliated ETFs predicts higher performance of the bank stock. ETF investment in stocks with high ownership of affiliated OEFs also appears informative. In addition to these informative trading mechanisms, ETFs also seem to engage in cross-trades with affiliated OEFs. Different from informative trading, cross-trades are typically related to negative future performance of ETFs. By contrast, affiliated OEFs seem to benefit from these cross-trades: a 1-standard-deviation increase in ETF/OEF cross-trades is related to an annualized 5.19% higher return and 12.17% higher inflow for the affiliated OEFs. If anything, such cross-trades could be associated with the proconglomerate incentive of subsidizing affiliated poorly performing OEFs.

To explore the second question—that is, the potential benefits and drawbacks for the investors—we link the level of ETF fees that investors need to pay and various types of risk that investors may face, such as tracking error and trading illiquidity, to ETF's proconglomerate activeness. Because security lending may generate income in addition to the aforementioned information or subsidization mechanisms, we also include security lending in this part of analysis.

In fee analysis, we find that ETF investors enjoy direct benefits from the security lending channel and may face some cost due to the subsidization channel. In fact, a 1-standard-deviation increase in lending fees translates into 0.6 bps lower fees that the ETF charges its investors. By contrast, every 1% negative return ofnel3filatid oO-2.6(E)F-4.5(F)-18391(i)-3.56F clean measure of the demand of sophisticated institutional investors. This allows us to gauge the impact of subsidization by examining whether sophisticated investors respond to various types of ETF activeness/incentives by changing their demand.

In particular, we construct an indirect measure to gauge the impact of potential subsidization as the difference between the holding-based return that ETFs can generate and the gross-of-fee reported return of the ETFs that investors can receive. A positive difference implies a potential transfer of benefits from the ETFs to the affiliated financial conglomerate; hence, we refer to this difference as Swapped transfer. In the lack of explicit information on the actual transactions across affiliated parties, this variable provides a rough approximation of what could be transferred, directly by—for example, a formal swap contract—or indirectly by, for example, cross-trades—from the ETFs to the affiliated financial conglomerate. More importantly, it captures the potential concern that investors and regulators may have on subsidizations; that is, ETFs may transfer benefits to their affiliated financial conglomerates rather than to their investors.

We find that positive Swapped transfer is typically associated with higher ETF outflows. A 1-standard-deviation increase in Swapped transfer is associated with an annual outflow of 3.56%. In other words, investors seem to have concerns when affiliated financial conglomerates benefit from the off-benchmark activities of ETFs. Moreover, ETF investors also withdraw capital when the affiliated bank's rating or return on assets (ROA) drops. A 1-standard-deviation deterioration in bank rating translates into 9.18% lower flows per year. More importantly, the outflow sensitivity with respect to Swapped transfer increases with worsening ratings/ROAs of the affiliated bank. These patterns suggest that investors regard potential subsidizations between ETFs and affiliated banks as detrimental especially when the latter become risky.

Our findings suggest that the global ETF industry is much more complicated than a simple offering of index trackers might indicate. These findings are the first—to the best of our knowledge—to provide evidence in support of the event that affiliated banks become distressed because a crisis that should be limited to the banking sector might spread to the equity market as a whole (e.g., FSB 2011; IMF 2011; Ramaswamy 2011).

In doing so, we also contribute to the burgeoning literature on ETFs. Although the entire ETF industry only has less than three decades of history, its growth speed has dwarfed other types of asset management firms and, subsequently, attracted tremendous academic attention (e.g., Boehmer and Boehmer 2003; Blocher and Whaley 2016; Bhattacharya et al. 2017; Israeli, Lee, and Sridharan 2017; Ben-David, Franzoni, and Moussawi 2018; Da and Shive 2018). Most existing studies focus on traditional passive ETFs as a trading instrument, from market traders' perspective. Our paper investigates the incentives of a broad set of international equity ETFs instead.

Our analysis also contributes to the literature on delegated asset management, particularly on passive benchmarking. There have been only a few attempts to address the issue of the relationship between ETFs and affiliated banks, despite their normative implications for both consumer protection and financial stability. Investors have been perceived as not fully aware or capable of understanding their exposure to distress risk. Our results on flows of ETFs should help alleviate such concerns.

Our findings also relate to the economics of mutual fund families. Research on the constraints and benefits that family affiliation imposes on funds has identified how family strategies condition fund performance, risk taking, and investment (Mamaysky and Spiegel 2001; Massa 2003; Nanda, Wang, and Zheng 2004; Gaspar, Massa, and Matos 2006). We broaden the focus to ETFs and their relationships with affiliated OEFs and banks.

2. The Industry, Data, and Main Variables

In this section, we first briefly describe the industry and then we define our data variables.

2.1 The ETF industry

Exchange Traded Funds, or ETFs, are index-tracking investment vehicles that allow investors to replicate an index cheaply. They represent a fixed combination of assets held as a function of their representation in the index they track, such as the S&P 500. Unlike Index Funds, investors can either invest the money in the fund/redeem its shares (for large orders) or buy/sell certificates representing ownership of ETFs. We will focus on ETFs that track equity indices and exclude leveraged or inverse ETFs.

Table 1 provides a snapshot of the ETF industry. For each year, we tabulate the number and total net assets (TNAs) (in billions of USD) of the ETFs in the first two columns of panel A. As of 2009, for instance, the ETF sample contains 921 ETFs with TNA of USD 760 billion. By contrast,

Table 1
Summary statistics

A. Snapshots of the ETF industry

	All ETFs		ETF replication methods						Sponsors affiliated with bank conglomerates				With valid benchmark	
			Full replication		Sampling		Synthetic		ETFs		OEFs			
Year	Number	TNA (in billions)	%number	%TNA	%number	%TNA	%number	%TNA	%number	%TNA	%number	%TNA	%number	%TNA
2001	109	61.12	18.35	67.30	77.06	31.08	4.59	1.62	98.17	98.03	36.10	23.50	52.29	93.06
2002	147	123.96	23.13	72.27	67.35	25.74	9.52	2.00	91.84	80.95	36.23	23.19	55.10	91.18
2003	166	185.08	22.29	62.42	65.66	34.76	12.05	2.83	87.95	79.87	40.05	24.50	55.42	86.18
2004	205	265.18	29.27	50.57	60.00	46.30	10.73	3.13	80.49	85.29	42.43	25.90	55.12	82.79
2005	315	356.93	38.10	45.73	52.38	50.90	9.52	3.38	77.78	87.14	44.80	39.98	56.19	78.72
2006	493	490.56	31.64	39.69	55.58	56.09	12.78	4.22	78.50	87.71	44.94	30.57	48.48	74.83
2007	687	670.64	29.99	38.97	52.98	55.87	17.03	5.16	76.13	86.53	44.00	30.93	45.71	72.96
2008	886	538.02	30.70	41.96	47.63	51.37	21.67	6.67	79.35	86.89	43.45	29.05	45.71	75.00
2009	921	759.91	30.62	35.51	48.43	55.87	20.96	8.62	79.59	83.93	43.01	29.19	45.39	67.52

B. Quantile distribution of ETF, OEF, and mutual fund family characteristics

			Quantile distribution				
	Mean	SD	10%	25%	Median	75%	90%
B1. ETF return (monthly, in %)							
Holding-based return	0.450	5.512	−7.580	−2.525	1.013	3.886	6.960
DGTW adjusted	−0.182	0.856	−1.180	−0.585	−0.135	0.232	0.639
ActiveShr performance	0.045	0.449	−0.343	−0.078	−0.002	0.090	0.453
Gross-of-fee NAV-based return	0.405	5.380	−6.922	−2.468	1.163	3.766	6.763
Swapped transfer	0.045	0.453	−0.331	−0.128	0.056	0.222	0.458
Fund return	0.374	5.380	−6.957	−2.497	1.132	3.740	6.730
Benchmark adjusted	0.047	0.651	−0.441	−0.109	−0.002	0.108	0.516
CAPM adjusted	0.127	0.946	−0.966	−0.471	0.032	0.741	1.424
FFC adjusted	−0.013	0.667	−0.794	−0.420	−0.006	0.367	0.816

(continued)

Table 1
Continued

B. Quantile distribution of ETF, OEF, and mutual fund family characteristics

			Quantile distribution				
	Mean	SD	10%	25%	Median	75%	90%
B2. ETF characteristics							
ActiveShr	0.277	0.307	0.021	0.050	0.130	0.382	0.824
Tracking error (in %)	0.515	1.001	0.021	0.039	0.105	0.618	1.391
log(fund illiquidity)	−0.786	3.944	−5.966	−3.322	−0.566	2.047	3.990
log(stock size in fund)	10.145	1.539	7.395	9.151	10.760	11.256	11.557
log(fund TNA)	19.598	2.065	16.908	18.118	19.475	21.233	22.282
log(fund age)	3.751	0.776	2.639	3.258	3.892	4.382	4.625
Expense ratio (annual, in %)	0.370	0.130	0.246	0.273	0.318	0.505	0.581
Fund flow (monthly, in %)	2.631	8.168	−4.182	−0.312	0.008	4.597	12.611
ETF premium (in %)	0.035	0.187	−0.080	−0.031	0.004	0.056	0.229
B3. OEF return (monthly, in %)							
Holding-based DGTW-adjusted return	−0.084	0.856	−1.020	−0.500	−0.089	0.287	0.808
OEF return	0.264	2.416	−3.801	−0.696	1.013	1.914	2.469
Benchmark adjusted	−0.014	0.903	−0.678	−0.259	0.000	0.246	0.758
CAPM adjusted	0.180	1.258	−1.117	−0.487	0.053	0.864	1.760
FFC adjusted	−0.027	0.850	−0.975	−0.467	−0.026	0.423	0.955
B4. OEF characteristics							
log(stock size in fund)	10.396	1.089	8.776	10.193	10.676	11.040	11.409
log(fund TNA)	18.862	1.736	16.490	17.636	19.117	20.186	20.495
log(fund age)	4.431	0.835	3.296	3.951	4.522	4.956	5.366
Expense ratio (annual, in %)	1.940	0.670	1.260	1.740	1.899	2.279	2.490
Fund flow (monthly, in %)	1.521	5.271	−3.116	−1.415	−0.032	3.058	9.720
B5. Cross-trades measures (quarterly, in %)							
ETF/OEF cross-trades	11.621	10.591	0.000	1.590	9.448	18.852	28.706
ETF/ETF cross-trades	9.251	11.766	0.577	1.780	4.711	11.784	24.252
B6. Family characteristics							
log(family TNA)	22.129	2.120	20.643	21.630	22.130	23.190	23.974
log(family age)	4.772	0.452	4.253	4.607	4.782	4.910	5.418
Family expense ratio	1.852	0.695	0.786	1.663	1.843	2.176	2.621
Family return	0.225	2.303	−3.728	−1.037	0.949	1.811	2.822
Family flow	5.497	5.322	−0.995	0.868	4.785	10.473	11.597

C. Quantile distribution of bank, stock, and country characteristics

	Mean	SD	Quantile distribution				
			10%	25%	Median	75%	90%
C1. Bank characteristics							
DGTW-adjusted bank return (in %)	−0.222	4.130	−4.917	−2.225	−0.240	1.862	4.459
Bank rating	3.823	1.330	2.000	2.250	4.000	5.000	6.000
Bank ROA (annual, in %)	1.719	6.464	−1.008	−0.175	0.398	0.554	4.329
C2. Stock characteristics							
Stock return (monthly, in %)	1.161	5.418	−4.975	−1.367	0.950	3.741	7.530
DGTW-adjusted stock return (monthly, in %)	−0.024	4.126	−4.507	−2.125	−0.117	2.043	4.690
Corporate loan dummy	0.019	0.136	0.000	0.000	0.000	0.000	0.000
Affiliated bank stock dummy	0.001	0.038	0.000	0.000	0.000	0.000	0.000
Security lending fee	0.607	1.166	0.110	0.139	0.191	0.421	1.553
log(stock size)	5.041	2.147	2.451	3.642	4.980	6.432	7.835
Turnover	0.096	0.137	0.003	0.012	0.050	0.102	0.245
log(stock illiquidity)	3.768	3.038	−1.732	4.957	5.223	5.488	5.652
log(net income)	1.534	3.308	−3.017	−0.222	2.345	4.121	4.543
log(sales)	5.706	2.055	3.127	4.333	5.834	7.028	8.017
log(total assets)	6.340	2.239	3.549	4.679	6.186	8.120	8.599
C3. Country characteristics (annual, in %)							
Stock market turnover	1.164	0.615	0.480	0.757	1.081	1.417	1.933
Stock market/GDP	1.368	1.053	0.445	0.670	1.087	1.484	2.829
Private bond market/GDP	1.357	0.398	0.887	1.082	1.367	1.653	1.880
Active ETF/OEF	1.704	3.054	0.000	0.000	0.000	2.030	6.871
ETF TNA/GDP	1.202	2.004	0.008	0.052	0.483	1.164	3.701
OEF TNA/GDP	39.583	117.810	2.176	4.894	9.325	26.443	66.784

This table presents the summary statistics for the data used in the paper during the 2001–2009 period. Panel A reports the number and total net assets (TNAs) of ETFs, the percentage number and percentage of TNAs of three ETF replication methods, and the percentage number and percentage of TNAs of ETFs and OEFs affiliated with bank conglomerates on a year-by-year basis. Panel B reports the mean, median, standard deviation, and the quantile distribution of ETF, OEF, and mutual fund family characteristics. Panel C reports similar statistics for bank, stock, and country characteristics. Table A1 provides detailed definitions of each variable.

as of 2001, the number of ETFs totaled 109 with a TNA of 61 billion, which confirms the astonishing rate of growth in the industry. Among the 921 ETFs existing in 2009, 480 are from the United States, and 357 are from Europe, compared with 85 and 16, respectively, in the year 2001. Thus, the importance of ETFs has increased even more outside the United States.

In the United States, ETFs tend to physically replicate the underlying index, which seems to be driven by regulatory rules. For example, the Investment Act of 1940 requires ETFs to hold 80% of their assets in securities matching the fund's name. By contrast, more than 50% of the ETFs in Europe use synthetic structures. UCITS-compliant ETFs that are synthetically replicated tend to be registered in Luxemburg to reduce haircuts on the collateral assets posted.[4]

In the next few columns of Table 1, panel A, we report the three replication methods as reported by Morningstar: full replication, optimized sampling, and synthetic replication. Table 1 shows that only 30% of the ETFs in the world use "full replication." In our view, only full replication can prevent the ETF from deviating from its benchmark. The holdings for other types of ETFs might deviate from their benchmarks that are affected by various information and subsidization motivations.

The next few columns of panel A report the fraction of ETFs that are affiliated with bank conglomerates and the analogous statistics of OEFs reported on an annual basis. We define "bank conglomerates" as the finan-

ETFs can be roughly classified into those run by pure asset managers (e.g., Vanguard) and those affiliated with "bank conglomerates" (e.g., Barclays).

Returning to panel A of Table 1, the year-by-year statistics illustrate that the involvement of banks in the ETF industry is impressive: in any year, more than 70% of ETFs and more than 80% of the TNA of the industry are affiliated with banks. By contrast, less than 30% of OEF TNAs are typically affiliated with banks. It is notable that this affiliation pattern primarily prevails in Europe, whereas in the United States some of the largest ETF providers, such as Vanguard, are not part of bank conglomerates. This difference suggests that proconglomerate incentives could be more influential in Europe.

The last two columns of panel A report the fraction of ETFs for which we are able to construct the benchmarks.[7] Our sample typically covers between 45% and 55% of ETFs in terms of numbers and from 67% to 90% of the TNA of the industry. The final sample contains 420 ETFs, among which 107 are domiciled in the United States and 261 in Europe. Altogether, 16,365 stocks are held by ETFs, of which 8,809 are listed in the United States and 3,431 are listed in Europe.[8]

2.2 Data sources

Our data are drawn from different sources. The ETF and OEF holdings data are from the Factset/Lionshares database.[9] The Factset/Lionshares holdings data on international funds are sparse before 2001, so our sample is restricted to the 2001–2009 period. We match the database to the Morningstar mutual fund database, which reports monthly total returns for global mutual funds. We use Morningstar classifications to identify ETFs ("Exchange-Traded Funds Universe" in Morningstar), index funds ("Index Funds" from "Open End Funds Universe"), and OEFs (the rest of the "Open End Funds Universe"). From Morningstar, we obtain additional variables such as fund net asset value (NAV), fund TNA, fund age, management expenses, market price, volatility of fund returns, and the benchmark tracked by ETFs

[7] For each ETF, we proxy for the benchmark portfolio it should hold by using the average holdings of the open-end index funds that follow the same index. If index OEFs are not tracking the benchmark, we use the average holdings of ETFs using full replication to proxy for the index holding. As noted in Cremers et al. (2016), use of the actual weights of explicitly indexed funds tracking the benchmark has the advantage that some of the weights in the official benchmark include stocks that in practice may not be fully investable by mutual funds due to illiquidity or other constraints.

[8] Table 1 includes benchmarks that are only followed by one ETF, which occurs, for instance, with approximately 244 indices in the year 2009. Our main regressions further exclude those one-ETF indices not followed by index OEF funds. Our main results are robust if we exclude all indices not followed by index OEFs or if we use the average of all index OEF and full replicating ETF holdings to proxy for index holdings.

[9] A detailed description of the data set can be found in Ferreira and Matos (2008). We find that approximately 40% of investment vehicles in the Factset/Lionshares database report quarterly portfolio holdings and approximately 50% report semiannual holdings, the remaining 10% report either monthly or yearly holdings. We address the issue of different reporting frequencies by institution from different countries by using the latest available holdings updates at the quarter end.

and index OEFs ("Primary Prospectus Benchmark"). We focus on funds that have "Equity" as the Morningstar "Broad Category Group."

Monthly stock return data and annual stock characteristics, such as market capitalization, net income, sales and total assets, are obtained from Datastream/Worldscope for international stocks, with all the variables quoted in USD. Data on banks come from BvD BankScope. This data set contains annual financial data of banks, including total assets, ROA, equity/liabilities ratio, loan loss reserve/gross loans ratio, net interest margin, cost/income ratio, and net loans/total assets ratio. The characteristics of the loan contracts and the identities of the borrowers and lenders are taken from Thomson Reuters LPC DealScan. The monthly S&P long-term issuer credit ratings come from Compustat. Because bank variables are observed only on an annual basis, we adopt annual frequency in our main tests. Using quarterly frequency based on available quarterly variables leads to similar conclusions.

2.3 Variables

We will define the main variables in the subsequent sections as we use them. Here, we just summarize the main control variables we will use in this paper as well as the measures of performance.

We consider measures of both ETF and stock performance. The measures of ETF performance are the gross-of-fee NAV-based return, the benchmark-adjusted fund return, DGTW-adjusted Holding-based return, and the ActiveShr performance. Holding-based return of an ETF is defined as the investment value-weighted average of the returns of the stocks in its portfolio. It represents the return the ETF would have earned based on the stocks in its portfolio. The ActiveShr performance is computed as the difference between the holding-based return of an ETF and that of its benchmark. It captures the abnormal return that an ETF can generate by deviating from the holding portfolio of its benchmark.

Similarly, OEF performance is proxied by benchmark-adjusted fund returns or DGTW-adjusted holding-based returns. As an additional robustness check, we also construct performance as alpha net of the risk factors posited by the international capital asset pricing model (CAPM) and the international Fama-French-Carhart (FFC) model. The latter model extends the standard factor-based risk corrections used in the domestic literature to account for the international dimension. It includes four international factors as the value-weighted average of the four domestic factors (market, size, book-to-market, and momentum).[10] The construction of these international

[10] For a given country, we download all the (active and defunct) stocks from Thomson Datastream and complement them with necessary accounting data from the Worldscope database. Then, for each country, we construct market (RMF), size (SMB), value (HML), and momentum (MOM) factors, closely following the original methodology of Fama and French (1993) and Carhart (1997). The four international factors are the value-weighted average of the four domestic factors in all countries.

factors is in the spirit of Griffin (2002). We extend these international factors to include the momentum factor because of its importance in the mutual fund literature. Table A1 provides further details regarding the construction of the factors.[11]

To define stock performance, we use the Daniel et al. (1997, DGTW) methodology. That is, we first create stock styles by double-sorting all the stocks into 25 independent book-to-market and size portfolios within each country. We then adjust the return of a given stock by its style average to compute its DGTW-adjusted return. Finally, we obtain portfolio-level DGTW-adjusted return as the investment value-weighted average of stock-level DGTW-adjusted returns for all the stocks in the portfolio.

We control for lagged fund, stock, and bank characteristics. Fund characteristics include the following: log(stock size in fund), defined as the logarithm of the investment value-weighted average market value of stocks invested in by the fund; log(fund TNA), defined as the logarithm of fund TNA; log(fund age), defined as the logarithm of the number of operational months since inception; Expense ratio, defined as the annual expense ratio; Fund return, defined as the annual return of the fund; and Fund flow, defined as the annual fractional flow received by the fund. Stock characteristics include the following: log(stock size), defined as the logarithm of the market value of the stock; Turnover, defined as the annual turnover ratio of the stock; log(stock illiquidity), defined as the logarithm of the Amihud (2002) stock illiquidity; log(net income), defined as the logarithm of its net income; log(sales), defined as the logarithm of its sales; and log(total assets), defined as the logarithm of its total assets. Table A1 provides a detailed definition for each variable.

Panel B of Table 1 reports the descriptive statistics of the variables, including the mean, median, standard deviation, and quantile distribution of monthly ETF and OEF returns, and major characteristics (in annual frequency) of the funds and fund families. Panel C reports similar statistics related to monthly stock returns, and other annual bank, stock, and country characteristics. It is notable that the ETF Holding-based return and gross-of-fee NAV-based return have different distributions, which provides evidence of the existence of synthetic operations in the ETF industry.

The DGTW-adjusted return for the ETF holdings has a wide distribution. At the 75% quantile level, for example, the DGTW holding-based abnormal return is 23 bps per month. The economic magnitude involved is quite large, which suggests that ETFs invest in very good stocks. It is also notable that the characteristics of affiliated members of ETFs, such as banks and OEFs, also

[11] We use these three to be conservative. Indeed, the benchmark-adjusted return allows us to control for the benchmark and is closer in spirit to the performance that investors observe. The international Fama-French-Carhart four-factor model employs the broadest set of factors and has been used to estimate mutual fund performance (e.g., Carhart 1997; Bollen and Busse 2005; Avramov and Wermers 2006; Mamaysky, Spiegel, and Zhang 2007, 2008).

exhibit wide distributions. In the next section, we conduct more formal tests to explore how ETFs' deviations from their benchmarks may help transfer value to affiliated parties.

3. The Surprising Selection Ability of ETFs

In this section, we document a surprising selection ability of ETFs based on their affiliations with banks and discuss its implications. To capture the potential information benefits accruing from the affiliation with the bank conglomerate, we use the LPC DealScan data and define a dummy variable, $\mathrm{CorporateLoanDummy}_{i,f,t}$, that equals 1 if, with respect to ETF f, its affiliated bank provides bank loan services to firm i in year t and 0 otherwise. This dummy variable proxies for the information that ETFs may obtain from their affiliated banks based on such banks' processing of corporate loans.

Then we use this dummy to define bank loan-related abnormal (i.e., benchmark-adjusted) stock ownership for all the ETFs as follows: $\mathrm{ETFadjOwn(Loaned\ Corp)}_{i,t} = \sum_f (h_{i,f,t} - \widehat{h}_{i,f,t}) \times \mathrm{CorporateLoanDummy}_{i,f,t}$, where $\mathrm{ETFadjOwn(Loaned\ Corp)}_{i,t}$ refers to bank loan-related abnormal ETF ownership for stock i in year t, $h_{i,f,t}$ and $\widehat{h}_{i,f,t}$ refer to the real and benchmark-implied ownership of ETF f in stock i, respectively. If the bank loan channel is indeed motivated by information, a positive change in abnormal ownership should predict higher stock returns out-of-sample. We therefore estimate the annual panel regression:

$$\mathrm{Perf}_{i,t} = \alpha + \beta \times \Delta\mathrm{ETFadjOwn(Loaned\ Corp)}_{i,t-1} + \gamma M_{i,t-1} + e_{i,t}, \quad (1)$$

where $\mathrm{Perf}_{i,t}$ is the average monthly DGTW-adjusted return of a stock in year t, and $\Delta\mathrm{ETFadjOwn(Loaned\ Corp)}_{i,t-1}$ refers to changes in abnormal ETF ownership of stock i in year $t-1$ related to bank loan information. The vector M stacks all the other stock and fund control variables as defined previously. We use year and stock fixed effects and cluster the errors at the stock level.

Table 2 reports the results. Models 1 through 4 report the full sample results. The results document that bank loan-related abnormal ownership of ETFs can generate significant performance out of sample: in Model 2, for instance, each 1% increase in bank loan-related abnormal ownership of ETFs translates into a 12 bps higher DGTW-adjusted return per year.[12]

As a "Placebo" test, we also construct the abnormal ETF ownership unrelated to affiliated bank loan services as follows: $\mathrm{ETFadjOwn(Unloaned\ Corp)}_{i,t} = \sum_f \left(h_{i,f,t} - \widehat{h}_{i,f,t}\right) \times (1 - \mathrm{CorporateLoan}$

[12] The dependent variable is reported as a percentage of monthly abnormal return. Thus, the impact of a 1% increase in ΔETFadjOwn(Loaned Corp) can be estimated for Model 2, for instance, as $1.02\% \times 12 \times 1\% = 12.2$ bps, and 1.02% is the regression coefficient on ΔETFadjOwn(Loaned Corp). Unreported tests using raw stock return lead to very similar results, and each 1% increase in bank loan-related abnormal ownership of ETFs can be transferred to a 11 bps higher return per year.

Table 2
ETF stock selection based on bank lending (stock level)

Out-of-sample DGTW-adjusted stock return (in %) regressed on ΔAbnormal ETF ownership of lending-related stocks

	Full sample				Synthetic	Sampling	Full replication	U.S.	European
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
ΔETFadjOwn(loaned corp)	0.982**	1.020**	0.974**	1.014**	0.359*	1.071**	0.228	−1.318	0.627**
	(2.12)	(2.19)	(2.11)	(2.19)	(1.72)	(2.33)	(0.77)	(−0.10)	(2.46)
ΔETFadjOwn(unloaned corp)			−0.076	−0.043	−0.109	−0.104	0.090	−0.022	0.092
			(−0.93)	(−0.53)	(−1.62)	(−1.12)	(1.11)	(−0.27)	(1.21)
log(stock size)	−3.090***	−3.085***	−3.089***	−3.085***	−0.008	0.138***	−0.065***	0.000	−0.072***
	(−38.91)	(−38.85)	(−38.87)	(−38.81)	(−0.31)	(7.47)	(−3.66)	(0.01)	(−6.21)
Turnover	−1.390***	−1.314***	−1.387***	−1.313***	−0.039	−0.884***	0.252***	−0.080	0.600***
	(−5.06)	(−4.77)	(−5.05)	(−4.77)	(−0.29)	(−8.38)	(2.85)	(−0.73)	(9.11)
log(net income)	0.078***	0.079***	0.078***	0.079***	0.381	1.132***	1.557***	0.556***	−0.991***
	(8.06)	(8.22)	(8.06)	(8.22)	(0.31)	(4.77)	(5.90)	(2.83)	(−3.36)
log(sales)	0.181**	0.181**	0.181**	0.181**	−0.172***	−0.070***	−0.055**	−0.016	0.098***
	(2.08)	(2.08)	(2.08)	(2.08)	(−3.09)	(−3.19)	(−2.57)	(−0.75)	(6.71)
log(total assets)	−0.118	−0.143	−0.119	−0.143	0.008***	−0.027***	−0.001	0.010*	0.004
	(−1.34)	(−1.61)	(−1.34)	(−1.61)	(3.63)	(−6.59)	(−0.21)	(1.94)	(1.32)
log(fund TNA)		−0.067		−0.068	−2.871***	−2.826***	−2.867***	−2.833***	−2.393***
		(−1.62)		(−1.62)	(−41.55)	(−40.59)	(−41.49)	(−40.48)	(−40.63)
log(fund age)		0.219*		0.217*	−1.695***	−1.609***	−1.735***	−1.749***	−1.792***
		(1.73)		(1.71)	(−6.58)	(−6.28)	(−6.75)	(−6.82)	(−7.22)
Expense ratio		0.421*		0.418	0.073***	0.071***	0.072***	0.069***	0.090***
		(1.65)		(1.64)	(7.98)	(7.77)	(7.86)	(7.54)	(10.14)

(continued)

Table 2
Continued

Out-of-sample DGTW-adjusted stock return (in %) regressed on ΔAbnormal ETF ownership of lending-related stocks

	Full sample				Synthetic	Sampling	Full replication	U.S.	European
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Fund return		−0.141***		−0.141***	0.116	0.104	0.107	0.109	0.141*
		(−3.61)		(−3.61)	(1.36)	(1.22)	(1.24)	(1.28)	(1.68)
Fund flow		−0.012**		−0.012**	−0.207**	−0.202**	−0.199**	−0.201**	−0.233***
		(−1.98)		(−1.98)	(−2.52)	(−2.46)	(−2.41)	(−2.44)	(−2.90)
Intercept	20.748***	20.773***	20.747***	20.789***	19.146***	19.508***	19.269***	19.389***	17.119***
	(34.67)	(20.06)	(34.67)	(20.08)	(37.70)	(38.08)	(38.22)	(38.54)	(42.61)
R-squared	.168	.170	.168	.170	.162	.163	.161	.160	.138
Obs	46,863	46,863	46,863	46,863	46,863	46,863	46,863	46,863	46,863

This table presents the results of the following annual panel regressions with year and stock fixed effects and their corresponding t-statistics with standard errors clustered at the stock level:

$$\mathrm{Perf}_{i,t} = \alpha + \beta_1 \Delta \mathrm{ETFadjOwn}(\mathrm{Loaned\ Corp})_{i,t-1} + \beta_2 \Delta \mathrm{ETFadjOwn}(\mathrm{Unloaned\ Corp})_{i,t-1} + \gamma \mathrm{M}_{i,t-1} + \mathrm{e}_{i,t},$$

where $Perf_{i,t}$ refers to the average monthly DGTW-adjusted return of a stock in year t, $\Delta ETFadjOwn(Loaned\ Corp)_{i,t-1}$ refers to the change in bank-loan-related abnormal ETF ownership of stock i in year $t-1$, and $\Delta ETFadjOwn(Unloaned\ Corp)_{i,t-1}$ refers to the change in abnormal ETF ownership unrelated to bank loans. Vector M stacks all other stock and fund control variables, including log(stock size), Turnover, log(net income), log(sales), log(total assets), log(fund TNA), log(fund age), Expense ratio, Fund return, and Fund flow. Table A1 provides detailed definitions of each variable. Models 5 to 9 further apply Model 4 to subsamples of ETFs, including synthetic replication ETFs, optimized sampling ETFs, full replication ETFs, U.S. ETFs, and European ETFs. *p .1; **p .05; ***p .01.

$\text{Dummy}_{i,f,t}$), in which all variables are defined the same as in ETFadjOwn (Loaned Corp)$_{i,t}$. The results reported in Models 3 and 4 show that the abnormal ETF ownership changes unrelated to affiliated bank loan services do not predict stock return. These results are consistent with our working hypothesis that the link with affiliated banks allows ETFs to select superior stocks.

Next, in Models 5 to 9, we break down the analysis into different subsamples. In Models 5 to 7, we consider the synthetic, sampling, and full replication ETFs, whereas in Models 8 and 9, we consider U.S. and European ETFs. We see that lending-related abnormal ownership changes for both synthetic and sampling ETFs as well as European ETFs forecast stock performance. In contrast, full replication and U.S. ETFs do not seem to be affected. Additional (unreported) robustness checks indicate that including bank characteristics aggregated at the stock level leads to similar results.

Jointly, these results indicate that ETFs deviate from their benchmarks—and their public image of index tracker—in stocks that have a lending relationship with affiliated banks and that such deviations result in higher performance because ETFs can overweight (underweight) stocks that are somehow confirmed to be good (bad) via the affiliate banks' corporate loan services. In brief, ETFs trading is highly informative in this bank loan channel, exhibiting some selection ability.

Although active OEFs are known to accrue information from affiliated banks (e.g., Irvine, Lipson, and Puckett 2007; Massa and Rehman 2008; Massa and Zhang 2012), their incentives are very different. Active OEFs want to collect superior information to deliver better performance to their investors. By contrast, ETFs do not have the fiduciary duty to generate index-adjusted performance for investors. Why, in this case, do ETFs have incentives to collect information?

It is important to understand the incentives behind the surprising selection ability of ETFs, not the least because the latter may indicate an incentive of helping affiliated financial conglomerates (i.e., "proconglomerate incentives") rather than that of delivering superior performance to investors (i.e., "proinvestor incentives"). Indeed, precisely because ETFs do not have any duty to deliver extra benefits to their investors, and additional benefits generated from this surprising selection ability could be ultimately delivered back to affiliated conglomerate through, for instance, the swap design between ETFs and their affiliated conglomerates. In this case, conglomerates have incentives to accrue information to their affiliated ETFs, allowing the former to indirectly benefit from the information advantage that they obtain from the financial services they offer. In this regard, compared to OEFs, ETFs are more suitable instruments to implement proconglomerate incentives because of the combination of their index-tracking fiduciary duty (i.e., extra benefits do not need to be passed on to investors) and the potential swap designs between ETFs and affiliated conglomerates.

4. How Pervasive Are Proconglomerate Incentives?

We now address the three questions raised in the introduction regarding the scope and influence of ETF activeness. We start with the first question of how pervasive is ETF activeness that could be related to proconglomerate incentives. We extend the bank loan channel to stocks that ETFs may accrue information from their affiliations: stocks of the affiliated bank and stocks (co)invested by the affiliated OEFs. We also explore the channel of cross-trading between affiliated ETFs and OEFs.

4.1 Selection ability on affiliated bank stocks

ETFs may directly accrue information (from their affiliated conglomerate) regarding the stocks of their affiliated bank. Does their trading predict abnormal performance of the bank stock as the bank loan channel? To address this issue, we estimate the following panel specification:

$$\mathrm{Perf}_{b,q} = \alpha + \beta \times \Delta\mathrm{ETFadjOwn}(\mathrm{Affiliated\ Bank})_{b,q-1} + \gamma \mathrm{M}_{b,q-1} + \mathrm{e}_{b,q}, \tag{2}$$

where $\mathrm{Perf}_{b,q}$ is the average monthly DGTW-adjusted return of a bank stock in quarter q, and $\Delta\mathrm{ETFadjOwn}(\mathrm{Affiliated\ Bank})_{b,q-1}$ refers to changes in abnormal ETF ownership in bank b by the affiliated ETFs (by netting out their benchmark-implied ownership). The vector M stacks all other stock and fund control variables as defined previously. We add quarter and bank fixed effects, and cluster the errors at the bank level.

Table 3 reports the results. Models 1 through 4 report the full sample results. Especially, Models 1 and 2 document that an increase in abnormal ownership of affiliated ETFs predicts higher performance of the bank stock. Models 3 and 4 present a "Placebo" test, where we construct $\Delta\mathrm{ETFadjOwn}(\mathrm{Unaffiliated\ Bank})_{b,q-1}$ as the change in percentage abnormal stock ownership held by the ETFs for unaffiliated banks. Models 5 to 9 further apply Model 4 to subsamples of ETFs, including synthetic replication ETFs, optimized sampling ETFs, full replication ETFs, U.S. ETFs, and European ETFs. We see that affiliated ETF abnormal ownership changes positively forecast bank performance in all sub-groups, except for full replication and U.S. ETFs. In contrast, ETFs unaffiliated with the bank, their abnormal ownership changes do not predict bank return in both full sample and subsamples. This provides evidence of selection ability related to superior information.[13]

[13] This evidence is also consistent with the possibility of bank support; that is, the investment of the ETF is used to prop up the value of the shares of the affiliated bank. Although this is possible, still the effect for the fund is to invest in assets whose price goes up in value. In the presence of pure price support, it is not clear this will translate in a consistent positive relationship between investment and future return.

Table 3
ETF stock selection of affiliated bank stocks (bank level)

Out-of-sample DGTW-adjusted bank stock return (in %) regressed on Δabnormal ETF ownership of affiliated bank stocks

	Full sample				Synthetic	Sampling	Full replication	U.S.	European
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
ΔETFadjOwn(affiliated bank)	0.062***	0.060***	0.065***	0.063***	0.066***	0.064***	−0.203	−5.398*	0.065***
	(2.78)	(2.86)	(2.85)	(2.92)	(3.16)	(2.99)	(−1.14)	(−1.84)	(3.20)
ΔETFadjOwn(unaffiliated bank)			−0.029	−0.028	−0.030	−0.034	−0.032	−0.143	−0.027
			(−1.31)	(−1.23)	(−1.10)	(−1.18)	(−1.12)	(−1.09)	(−1.21)
log(stock size)	0.648*	0.421	0.659*	0.435	0.428	0.433	0.404	0.407	0.432
	(1.73)	(1.19)	(1.75)	(1.22)	(1.21)	(1.22)	(1.12)	(1.08)	(1.22)
Turnover	2.254	1.905	2.237	1.891	1.893	1.904	2.042	2.195	1.875
	(0.70)	(0.62)	(0.70)	(0.62)	(0.62)	(0.62)	(0.68)	(0.73)	(0.61)
log(net income)	0.071	0.085	0.072	0.085	0.085	0.085	0.092	0.089	0.085
	(1.19)	(1.55)	(1.22)	(1.58)	(1.56)	(1.57)	(1.64)	(1.58)	(1.58)
log(sales)	0.174	0.203	0.214	0.242	0.240	0.240	0.234	0.243	0.241
	(0.22)	(0.25)	(0.27)	(0.29)	(0.29)	(0.29)	(0.28)	(0.29)	(0.29)
log(total assets)	−0.401	−0.463	−0.452	−0.514	−0.499	−0.509	−0.507	−0.489	−0.509
	(−0.47)	(−0.52)	(−0.52)	(−0.58)	(−0.56)	(−0.57)	(−0.57)	(−0.54)	(−0.57)
log(fund TNA)		0.053		0.063	0.066	0.071	0.062	0.077	0.060
		(0.20)		(0.24)	(0.25)	(0.27)	(0.22)	(0.28)	(0.22)
log(fund age)		1.843**		1.778**	1.785**	1.775**	1.748**	1.830**	1.787**
		(2.28)		(2.20)	(2.20)	(2.19)	(2.12)	(2.22)	(2.21)
Expense ratio		−1.044		−0.941	−0.927	−0.905	−1.028	−1.226	−0.947
		(−0.77)		(−0.70)	(−0.68)	(−0.67)	(−0.75)	(−0.89)	(−0.71)

(continued)

Table 3
Continued

Out-of-sample DGTW-adjusted bank stock return (in %) regressed on Δabnormal ETF ownership of affiliated bank stocks

	Full sample				Synthetic	Sampling	Full replication	U.S.	European
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Fund return		0.186		0.188	0.190	0.190	0.194	0.197	0.188
		(0.60)		(0.61)	(0.62)	(0.62)	(0.63)	(0.64)	(0.62)
Fund flow		0.030		0.028	0.029	0.028	0.027	0.027	0.029
		(1.51)		(1.44)	(1.45)	(1.42)	(1.37)	(1.32)	(1.47)
Intercept	−3.224	−9.743	−3.095	−9.643	−9.851	−9.872	−9.253	−10.235	−9.621
	(−0.58)	(−1.10)	(−0.55)	(−1.08)	(−1.11)	(−1.11)	(−1.04)	(−1.12)	(−1.08)
R-squared	.145	.149	.147	.152	.151	.152	.149	.153	.152
Obs	785	785	785	785	785	785	785	785	785

This table presents the results of the following quarterly panel regressions with quarter and bank fixed effects and their corresponding t-statistics clustered at the bank level:

$$\mathrm{Perf}_{b,q} = \alpha + \beta_1 \Delta \mathrm{ETFadjOwn(Affiliated\ Bank)}_{b,q-1} + \beta_2 \Delta \mathrm{ETFadjOwn(Unaffiliated\ Bank)}_{b,q-1} + \gamma \mathrm{M}_{b,q-1} + \mathrm{e}_{b,q},$$

where $Perf_{b,q}$ refers to the average monthly DGTW-adjusted return of bank b in quarter q, $\Delta ETFadjOwn(Affiliated\ Bank)_{b,q-1}$ refers to the change in percentage abnormal bank ownership held by the affiliated ETFs (by netting out their benchmark-implied ownership), $\Delta ETFadjOwn(Unaffiliated\ Bank)_{b,q-1}$ refers to the change in percentage abnormal bank ownership held by the ETFs not affiliated with the bank. Vector M stacks all other stock and fund control variables, including log(stock size), Turnover, log(net income), log(sales), log(total assets), log(fund TNA), log(fund age), Expense ratio, Fund return, and Fund flow. Table A1 provides detailed definitions of each variable. Models 5 to 9 further apply Model 4 to subsamples of ETFs, including synthetic replication ETFs, optimized sampling ETFs, full replication ETFs, U.S. ETFs, and European ETFs. *p < .1; **p < .05; ***p < .01.

4.2 Co-ownership and cross-trading between ETFs and affiliated OEFs

Another potential source of information for the ETFs may flow from the affiliated OEFs. Therefore, we investigate the possibility of a relation between the stock selection of the ETFs and the position of the affiliated OEFs. More specifically, we estimate:

$$\mathrm{Perf}_{i,t} = \alpha + \beta \times \Delta\mathrm{ETFadjOwn}(\mathrm{High\ ETF/OEF\ Co\text{-}Ownership})_{i,t-1} + \gamma \mathrm{M}_{i,t-1} + \mathrm{e}_{i,t}, \tag{3}$$

where $\Delta\mathrm{ETFadjOwn}(\mathrm{High\ ETF/OEF\ Co\text{-}Ownership})_{i,t-1}$ refers to changes in abnormal ETF ownership in stocks in which the affiliated OEFs have high degree of coinvestment (i.e., co-ownership). High and low co-ownership are constructed according to the median break point in the cross section of each period. For stocks in which ETFs and OEFs have high co-ownership, we expect the ETF trading to be informative because information could be accrued from affiliated OEFs or conglomerate for these relative "important" stocks. The use of this information is beneficial to affiliated conglomerate in a way that is similar to the bank loan channel.

Another important way of benefiting affiliated conglomerate is cross-trading ("cross-trades"), through which ETFs and affiliated OEFs could swap trading benefits/losses (see Gaspar, Massa, and Matos 2006 for the cross-trading mechanism within the OEF industry). To explore whether this channel also exist, we first compute the cross-trades between ETF i and affiliated OEF j (ETF/OEF Cross-Trades) in a given quarter q as follows: $\mathrm{CrossTra}_{ij,q} = [(\sum_{s\in S_1\cap S_2} N_{s,i,q}P_{s,q} + N_{s,j,q}P_{s,q}) \times I\{\Delta N_{s,i,q} \times \Delta N_{s,j,q} < 0\}] / (\sum_{s\in S_1} N_{s,i,q}P_{s,q} + \sum_{s\in S_2} N_{s,j,q}P_{s,q})$, where S_1 and S_2 represent the set of companies held by fund i and j, $P_{s,q}$ is the price of company s at quarter q, $N_{s,i,q}$ and $N_{s,j,q}$ are the number of shares of company s held by fund i and j, respectively, and $I\{\cdot\}$ is an indicator function that equals 1 if $N_{s,i,q}$ and $N_{s,j,q}$ change in opposite directions and 0 otherwise, following Gaspar, Massa, and Matos (2006). We then average the ETF/OEF Cross-Trades at ETF-stock level across all ETF-OEF pairs within the same fund family, and those above (below) the median break point in the cross section of each period are labeled as High (Low) ETF/OEF cross-trades. Finally, we construct a variable $\Delta\mathrm{ETFadjOwn}(\mathrm{High\ ETF/OEF\ CrossTrades})$, referring to changes in abnormal ETF ownership in stocks in which the affiliated OEFs have high degree of cross-trading. We then link stock performance to this cross-trading related ownership variable replacing the co-ownership variable with this variable in Equation (3).

Table 4 reports the results. Models 1 through 4 report the results for the co-ownership-based variable and Models 5 to 8 report the results for the cross-trades-based variable. We find that changes in abnormal ETF ownership in

Table 4
ETF stock selection based on affiliated OEFs (stock level)

Out-of-sample DGTW-adjusted stock return (in %) regressed on Δabnormal ETF ownership of OEF-related stocks

OEF-related corp =	ETF/OEF co-ownership				ETF/OEF cross-trades			
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
ΔETFadjOwn(high OEF-related corp)	0.360***	0.362***	0.369***	0.373***	−0.088	−0.098	−0.068	−0.077
	(4.38)	(4.38)	(4.41)	(4.41)	(−0.54)	(−0.60)	(−0.42)	(−0.47)
ΔETFadjOwn(low OEF-related corp)			0.077	0.088			0.244***	0.255***
			(1.06)	(1.21)			(3.93)	(4.07)
log(stock size)	−3.127***	−3.130***	−3.128***	−3.131***	−3.125***	−3.127***	−3.127***	−3.130***
	(−37.93)	(−37.92)	(−37.92)	(−37.91)	(−37.92)	(−37.90)	(−37.92)	(−37.91)
Turnover	−1.812***	−1.806***	−1.811***	−1.805***	−1.817***	−1.811***	−1.816***	−1.809***
	(−6.20)	(−6.18)	(−6.20)	(−6.18)	(−6.22)	(−6.20)	(−6.22)	(−6.20)
log(net income)	0.076***	0.076***	0.076***	0.076***	0.076***	0.077***	0.076***	0.077***
	(7.71)	(7.76)	(7.72)	(7.76)	(7.72)	(7.77)	(7.72)	(7.77)
log(sales)	0.083	0.081	0.083	0.081	0.083	0.081	0.083	0.081
	(0.86)	(0.84)	(0.86)	(0.84)	(0.87)	(0.85)	(0.86)	(0.84)
log(total assets)	−0.208**	−0.214**	−0.208**	−0.215**	−0.209**	−0.215**	−0.209**	−0.215**
	(−2.22)	(−2.28)	(−2.22)	(−2.28)	(−2.22)	(−2.28)	(−2.23)	(−2.29)
log(fund TNA)		−0.040		−0.039		−0.044		−0.039
		(−1.07)		(−1.03)		(−1.16)		(−1.04)
log(fund age)		0.288***		0.289***		0.293***		0.295***
		(2.95)		(2.97)		(3.00)		(3.03)
Expense ratio		0.480		0.486		0.505		0.513
		(1.32)		(1.34)		(1.39)		(1.41)
Fund return		−0.030		−0.030		−0.025		−0.028
		(−0.95)		(−0.94)		(−0.81)		(−0.88)

Fund flow		0.006		0.006		0.006		0.006
		(1.08)		(1.09)		(1.09)		(1.10)
Intercept	22.656***	21.717***	22.660***	21.693***	22.645***	21.697***	22.661***	21.647***
	(40.65)	(23.27)	(40.64)	(23.24)	(40.65)	(23.25)	(40.65)	(23.20)
R-squared	.175	.176	.175	.176	.175	.175	.175	.176
Obs	40,608	40,608	40,608	40,608	40,608	40,608	40,608	40,608

Models 1 to 4 present the results of the following annual panel regressions with year and stock fixed effects and their corresponding t-statistics with standard errors clustered at the stock level:

$$\mathrm{Perf}_{\mathrm{i,t}} = \alpha + \beta_1 \Delta \mathrm{ETFadjOwn}(\mathrm{High\ \ ETF/OEF\ \ Co\text{-}Ownership})_{\mathrm{i,t-1}} + \gamma \mathrm{M}_{\mathrm{i,t-1}} + \mathrm{e}_{\mathrm{i,t}},$$

where $Perf_{i,t}$ refers to the average monthly DGTW-adjusted return of a stock in year t, $\Delta ETFadjOwn(High\ \ ETF/OEF\ \ Co\text{-}Ownership)_{i,t-1}$ refers to the change in abnormal ETF ownership in stocks in which the affiliated OEFs have high degree of coinvestment (i.e., co-ownership). High (low) ETF/OEF co-ownership refers to the stock with above (below) median common ownership between ETFs and affiliated OEFs in the cross section of each period. Vector M stacks all other stock and fund control variables, including log(stock size), Turnover, log(net income), log(sales), log(total assets), log(fund TNA), log(fund age), Expense ratio, Fund return, and Fund flow. Models 5 to 8 present similar statistics when $\Delta ETFadjOwn(High\ \ ETF/OEF\ \ Co\text{-}Ownership)$ is replaced with $\Delta ETFadjOwn(High\ \ ETF/OEF\ \ CrossTrades)$, referring to changes in abnormal ETF ownership in stocks in which the affiliated OEFs have high degree of cross-trading, and cross-trading is defined following Gaspar, Massa, and Matos (2006). Table A1 provides detailed definitions of each variable. *p .1; **p .05; ***p .01.

stocks also heavily held by affiliated OEFs increase the subsequent performance. For instance, each 1% increase in ETF/OEF co-ownership-based abnormal ownership of ETFs is related to a 4.5 bps higher DGTW-adjusted return per year (Model 4).[14] By contrast, cross-trade between ETFs and affiliated OEFs is negatively related to performance. In particular, each 1% increase in abnormal ownership of ETFs in stocks with low ETF/OEF cross-trades activities is related to a 3 bps higher DGTW-adjusted return per year (Model 8).

One potential explanation on the difference between ETF/OEF co-ownership and ETF/OEF cross-trades is that the former could result from the information spillover between affiliated OEFs and ETFs, while the latter could be largely driven by the need to subsidize affiliated OEFs (e.g., subsidizing ETFs may end up holding more underperforming stocks after cross-trades with subsidized OEFs). To further verify the information channel, we investigate their impact jointly with the aforementioned bank loan channel, and show that ETFs and OEFs receive common information related to bank loans from the affiliated bank. To save space, we relegate these additional results to the Internet Appendix (Table IN1). Our remaining task is to examine whether cross-trades can be associated with the incentives to subsidize OEFs in the same group. If the intuition is correct, cross-trades should be associated with higher subsequent OEF performance. We take on this task in the next subsection.

4.3 Proconglomerate incentives behind ETF/OEF cross-trades

We now further explore the implications of the cross-trades channel. It is well known that investors withdraw capital from poorly performing OEFs, which is a cost to OEFs. By help promoting OEFs' performance, cross-trades from ETFs could effectively help OEFs to avoid outflows/attract new flows.[15] We test this hypothesis in a two-stage framework. In the first stage, we consider the possibility that both ETFs and OEFs deviate from their benchmarks to allow for convenient cross-trading. In this case, their portfolios exhibit off-benchmark active shares in the spirit of Cremers and Petajisto (2009). For cross-trading to be possible on a particular stock, both parties need to deviate on the same stock. Hence, we need to go beyond the original active share

[14] The dependent variable is reported as a percentage of monthly abnormal return. Thus, the impact of a 1% increase in ΔETFadjOwn(High ETF/OEF Co-Ownership) can be estimated for Model 4, for instance, as $0.373\% \times 12 \times 1\% = 4.5$ bps, and 0.373% is the regression coefficient on ΔETFadjOwn(High ETF/OEF Co-Ownership).

[15] Note that it makes sense for performance to be transferred from a "closed-end" instrument to an "open-end" fund, as the former will not suffer from outflows due to the "closed-end" property, whereas the latter will gain inflows due to the "open-end" property. ETFs are effectively "closed-end" funds to retail investors as explained in Footnote 3, making this subsidization feasible. The opposite direction of trading is unlikely to happen, as it will lead to outflows for the latter without obvious inflows of the former. Based on the same argument, this mechanism is unlikely to benefit two "closed-end" instruments, an implication we will use to conduct a placebo test. The mechanism could benefit two "open-end" funds (see, e.g., Gaspar, Massa, and Matos 2006).

measure and investigate how common active ownership between ETFs and OEFs (which we label

Table 5
ETF and OEF cross-trades

A. Two-stage OEF flow (in %) and performance (in %) regression (OEF level)

	First-stage	Second-stage						
	ETF/OEF cross-trades	Fund flow	BMK-adj volatility	Return	BMK-adj	DGTW	CAPM	FFC

B. Two-stage ETF flow (in %) and performance (in %) regression (ETF level)

	First-stage	Second-stage						
	ETF/ETF cross-trades Model 1	Fund flow Model 2	BMK-adj volatility Model 3	Return Model 4	BMK-adj Model 5	DGTW Model 6	CAPM Model 7	FFC Model 8
ETF/ETF Co-ActiveShr	4.189*							
	(1.67)							
ETF/ETF cross-trades		0.350	0.028	0.055	−0.031	−0.084	0.050	0.058
		(1.25)	(0.47)	(0.31)	(−0.37)	(−1.21)	(0.71)	(1.05)
log(stock size in fund)	1.827***	−0.507	0.004	−0.272	0.046	0.099	−0.165	−0.071
	(8.97)	(−0.95)	(0.03)	(−0.84)	(0.29)	(0.73)	(−1.24)	(−0.66)
log(fund TNA)	−1.119***	0.108	−0.075	0.078	−0.050	−0.077	0.079	0.061
	(−4.57)	(0.33)	(−1.14)	(0.41)	(−0.56)	(−0.99)	(1.02)	(1.05)
log(fund age)	2.354*	−2.589***	−0.424**	−1.157**	0.079	0.217	−0.243	−0.276*
	(1.90)	(−2.90)	(−2.54)	(−2.13)	(0.39)	(1.06)	(−1.22)	(−1.93)
Expense ratio	3.595	−1.965	−0.041	−1.280	0.368	0.277	0.518	−0.208
	(1.33)	(−1.46)	(−0.14)	(−1.58)	(1.10)	(0.83)	(1.57)	(−0.83)
Fund return	−0.254	0.099	0.037	−0.561***	0.057**	0.048*	0.080***	0.094***
	(−1.28)	(0.80)	(1.34)	(−9.20)	(2.04)	(1.71)	(2.93)	(4.36)
Fund flow	−0.026	−0.052	−0.009	0.021	−0.005	−0.013	−0.013	−0.008
	(−0.15)	(−0.59)	(−0.75)	(0.74)	(−0.44)	(−0.76)	(−1.38)	(−0.79)
Intercept	1.576	12.003***	2.581***	6.032***	0.506	0.357	0.395	−0.100
	(0.25)	(3.91)	(3.66)	(3.48)	(0.84)	(0.45)	(0.59)	(−0.20)
Obs	561	561	561	561	561	561	561	561

Panel A presents the results of the following two-stage panel regressions at the OEF level and their corresponding t-statistics clustered by fund after controlling for the year and fund fixed effects:

$$\text{First stage: CrossTrades}_{f,t} = \alpha + \beta \text{ETF/OEF Co-ActiveShr}_{f,t} + \gamma M_{f,t} + e_{f,t},$$

$$\text{Second stage: OEF\_Char}_{f,t+1} = \alpha + \beta \widehat{\text{CrossTrades}}_{f,t} + \gamma M_{f,t} + e_{f,t+1},$$

where $CrossTrades_{f,t}$ refers to the average quarterly cross-trades of fund f with other affiliated ETF(s) in year t. $ETF/OEF\ \ Co\text{-}ActiveShr_{f,t}$ refers to the benchmark-adjusted common active shares between an OEF f and its affiliated ETF(s), defined as $\sum_{i\in\{f\cap ETF\}}\left|w_{i,f,t}-\widehat{w}_{i,f,t}\right|/2$ for all stock i that is held by both OEF f and its affiliated ETF(s), where $w_{i,f,t}$ is the investment weight of stock i in fund f in year t, $\widehat{w}_{i,f,t}$ is the benchmark investment weight. $OEF\_Char_{f,t+1}$ refers to the OEF characteristics including average monthly flow, benchmark-adjusted return volatility (by netting out the benchmark average of return volatility, defined as the standard deviation of monthly return), monthly return, and risk-adjusted OEF return. $\widehat{CrossTrades}_{f,t}$ is the projected value of $CrossTrades_{f,t}$ from the first stage, and vector M stacks all other control variables, including log(stock size in fund), log(fund TNA), log(fund age), Expense ratio, OEF return, and Fund flow. OEF returns are adjusted by subtracting the benchmark return, the DGTW portfolio return, the CAPM, and the international Fama-French-Carhart (FFC) model. Panel B reports similar statistics of the following two-stage regressions at the ETF level:

$$\text{First stage: CrossTrades}_{f,t}=\alpha+\beta \text{ETF/ETF Co-ActiveShr}_{f,t}+\gamma \text{M}_{f,t}+\text{e}_{f,t},$$

$$\text{Second stage: ETF\_Char}_{f,t+1}=\alpha+\beta\widehat{\text{CrossTrades}}_{f,t}+\gamma \text{M}_{f,t}+\text{e}_{f,t+1}$$

where ETF/ETF Co-ActiveShr$_{f,t}$ refers to the benchmark-adjusted common active shares between an ETF f and other affiliated ETF(s), ETF_Char$_{f,t+1}$ refers to the ETF characteristics including average monthly flow, benchmark-adjusted return volatility, monthly return, and risk-adjusted ETF return, as defined above. *p $< .1$; **p $< .05$; ***p $< .01$.

As noted in Footnote 15, this type of subsidization can only occur from ETFs to OEFs. It makes less sense, from the conglomerate's perspective, to let ETFs to conduct similar cross-subsidization. Panel B provides a placebo test based on this intuition. There, we tabulate the formation (first stage) and influence (second stage) of cross-trades between affiliated ETFs. Unlike the case of ETF/OEF cross-trades, ETF/ETF cross-trades are much weaker and have no effect on either flow or performance. Therefore, cross-trades help OEFs, but not ETFs. Note that cross-trades may not directly affect ETF returns (recall that index returns can be delivered by the affiliated banks). Nonetheless, the costs of cross-trades may be indirectly charged via increased fees, a topic we will explore in our later sections.

Next, as a robustness check, we split the sample by type of ETF and report the results in Table 6. Panel A reports subsample results for ETF/OEF cross-trades with synthetic ETFs (Models 1 to 4), optimized sampling ETFs (Models 5 to 8), and full replication ETFs (Models 9 to 12), and panel B shows similar subsample results for U.S. ETFs (Models 1 to 4) and European ETFs (Models 5 to 8). We find that the cross-trades channel is significant for optimized sampling ETFs, but not for synthetic and full replication ETFs. U.S. ETFs are not subject to this problem; instead, the problem is concentrated in European ETFs.

Overall, we find that proconglomerate incentives are quite widely observed through several distinctive mechanisms. Similar to what we have observed in the bank loan channel, ETFs investment in the stocks of the affiliated bank are related to better performance. ETF investments in stocks with high co-ownership of affiliated OEFs also appear informative. Finally, ETFs also seem to engage in cross-trades with affiliated OEFs.

Although these mechanisms could arise to benefit conglomerates, we are yet to test their impact on ETF investors. In principle, proconglomerate and proinvestor incentives may not be mutually exclusive. Therefore, our next section takes on the task of investigating the impact of these mechanisms on ETF investors.

5. Benefits and Costs of ETF Incentives from Investors' Perspectives

To explore the second question of whether investors benefit or suffer from ETF activeness, we link major benefits/costs of ETF investments (from investors' perspective) to aforementioned mechanisms of proconglomerate incentives. In addition to aforementioned mechanisms (i.e., ETF's investment in stocks with affiliated bank loan services, ETF's investment in stocks of affiliated bank, and the need to help affiliated OEFs proxied by the performance of affiliated OEFs), we also include the channel of security lending, because many ETFs lend out their stocks in its portfolio in order to generate additional income. The major benefits/costs of ETF investments are

Table 6

Robustness checks on ETF and OEF cross-trades

A. Two-stage OEF flow (in %) and performance (in %) regression (replication method)

	Synthetic replication ETF				Optimized sampling ETF				Full replication ETF			
	First-stage	Second-stage			First-stage	Second-stage			First-stage	Second-stage		
	ETF/OEF cross-trades	Fund flow	Return	DGTW	ETF/OEF cross-trades	Fund flow	Return	DGTW	ETF/OEF cross-trades	Fund flow	Return	DGTW
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
ETF/OEF Co-ActiveShr	18.492***				19.563***				25.400***			
	(11.52)				(17.32)				(12.52)			
ETF/OEF cross-trades		0.038	0.085**	0.019*		0.050**	0.068***	0.008		0.007	0.021	0.008
		(1.24)	(2.05)	(1.77)		(2.25)	(2.78)	(1.10)		(0.19)	(0.75)	(1.43)
log(stock size in fund)	0.536	−0.466**	−0.249	−0.100	1.241***	−0.685***	−0.289***	−0.083**	0.386	0.043	−0.468*	−0.149*
	(1.24)	(−2.56)	(−1.07)	(−1.23)	(3.14)	(−4.68)	(−2.91)	(−2.23)	(0.67)	(0.15)	(−1.73)	(−1.94)
log(fund TNA)	−0.429	−0.518***	0.420***	0.024	0.615**	−0.683***	0.290***	0.013	0.052	−0.303	−0.070	−0.074*
	(−1.22)	(−3.79)	(3.32)	(0.58)	(2.15)	(−5.91)	(4.33)	(0.59)	(0.13)	(−1.37)	(−0.44)	(−1.88)
log(fund age)	0.401	0.638**	0.149	−0.029	0.271	0.031	−0.046	0.029	1.933***	0.280	−0.040	−0.025
	(0.61)	(2.55)	(0.56)	(−0.41)	(0.50)	(0.15)	(−0.27)	(0.52)	(2.90)	(0.87)	(−0.15)	(−0.48)
Expense ratio	0.299	0.472	1.277**	0.039	4.327***	−0.945***	−0.160	0.047	1.268	0.323	1.231***	−0.023
	(0.36)	(1.25)	(2.24)	(0.46)	(6.62)	(−4.35)	(−0.96)	(1.18)	(1.05)	(0.58)	(3.74)	(−0.44)
OEF return	0.099	−0.306***	−0.805***	−0.027*	−0.289*	−0.095*	−0.657***	0.055***	0.040	−0.177*	−0.537***	−0.010
	(0.55)	(−4.74)	(−11.06)	(−1.72)	(−1.81)	(−1.71)	(−17.03)	(4.46)	(0.18)	(−1.88)	(−9.16)	(−0.73)
Fund flow	−0.157**	0.168***	0.085*	0.026***	−0.162***	0.178***	0.061***	0.000	0.130	0.134**	−0.051	0.003
	(−2.54)	(5.54)	(1.83)	(2.71)	(−2.71)	(5.43)	(3.18)	(0.02)	(1.25)	(2.15)	(−1.12)	(0.46)
Intercept	6.168	9.444***	−11.496***	0.041	−29.490***	21.142***	−3.905**	0.070	−13.518	3.256	3.487	2.883***
	(0.78)	(2.80)	(−2.98)	(0.03)	(−3.91)	(7.46)	(−2.20)	(0.13)	(−1.18)	(0.54)	(0.84)	(2.62)
Obs	634	634	634	634	1,233	1,233	1,233	1,233	210	210	210	210

B. Two-stage OEF flow (in %) and performance (in %) regression (domicile country)

	U.S. ETF				European ETF			
	First-stage	Second-stage			First-stage	Second-stage		
	ETF/OEF cross-trades Model 1	Fund flow Model 2	Return Model 3	DGTW Model 4	ETF/OEF cross-trades Model 5	Fund flow Model 6	Return Model 7	DGTW Model 8
ETF/OEF Co-ActiveShr	11.442***				11.853***			
	(8.29)				(8.50)			
ETF/OEF cross-trades		0.005	0.017	−0.015		0.196***	0.135***	−0.007
		(0.04)	(0.47)	(−1.28)		(2.61)	(4.09)	(−0.97)
log(stock size in fund)	0.767	−0.779**	−0.104	−0.036	2.072***	−0.421	−0.794***	−0.163***
	(1.50)	(−2.37)	(−1.09)	(−0.80)	(4.52)	(−1.40)	(−5.07)	(−3.83)
log(fund TNA)	−0.699**	−1.245***	0.145*	−0.023	0.946**	−1.267***	−0.156	−0.026
	(−2.29)	(−4.04)	(1.94)	(−0.88)	(2.43)	(−4.49)	(−1.49)	(−1.00)
log(fund age)	−1.508*	0.511	0.175	0.082	−0.118	0.571	0.249	0.037
	(−1.74)	(0.66)	(0.79)	(0.85)	(−0.14)	(1.63)	(1.35)	(0.86)
Expense ratio	−0.106	0.093	0.106	0.117**	1.060	−0.327	1.829***	0.121***
	(−0.11)	(0.15)	(0.62)	(2.25)	(1.02)	(−0.46)	(7.81)	(2.58)
OEF return	−0.365*	−0.475***	−0.512***	0.144***	−0.301*	0.150	−0.556***	−0.016
	(−1.95)	(−2.83)	(−11.21)	(6.79)	(−1.66)	(1.48)	(−11.51)	(−1.42)
Fund flow	−0.009	0.278***	0.027*	−0.004	−0.181***	0.236***	0.055***	0.011***
	(−0.11)	(3.28)	(1.87)	(−0.51)	(−3.31)	(4.03)	(3.46)	(2.80)
Intercept	15.930*	32.032***	−3.037	0.381	−31.060***	23.596***	3.362	1.633**
	(1.72)	(4.10)	(−1.49)	(0.60)	(−3.27)	(3.77)	(1.25)	(2.50)
Obs	557	557	557	557	1,251	1,251	1,251	1,251

This table presents the results of the following two-stage panel regressions at the OEF level and their corresponding t-statistics clustered by fund after controlling for the year and fund fixed effects:

$$\text{First stage}: \mathrm{CrossTrades}_{f,t} = \alpha + \beta \mathrm{ETF/OEF\ Co\text{-}ActiveShr}_{f,t} + \gamma \mathrm{M}_{f,t} + e_{f,t},$$

$$\text{Second stage}: \mathrm{OEF\_Char}_{f,t+1} = \alpha + \beta \widehat{\mathrm{CrossTrades}}_{f,t} + \gamma \mathrm{M}_{f,t} + e_{f,t+1},$$

where $CrossTrades_{f,t}$ refers to the average quarterly cross-trades of fund f with other affiliated ETF(s) in year t. $ETF/OEF\ Co\text{-}ActiveShr_{f,t}$ refers to the benchmark-adjusted common active shares between an OEF f and its affiliated ETF(s). $OEF\_Char_{f,t+1}$ refers to the OEF characteristics including average monthly flow, monthly return and risk-adjusted OEF return (by subtracting the DGTW portfolio return), $\widehat{CrossTrades}_{f,t}$ is the projected value of $CrossTrades_{f,t}$ from the first stage, and vector M stacks all other control variables, including log(stock size in fund), log(fund TNA), log(fund age), Expense ratio, OEF return, and Fund flow. Panel A reports subsample results for cross-trades with synthetic replication ETFs (Models 1 to 4), optimized sampling ETFs (Models 5 to 8), and full replication ETFs (Models 9 to 12). Panel B reports similar subsample results for U.S. ETFs (Models 1 to 4) and European ETFs (Models 5 to 8).

*p < .1; **p < .05; ***p < .01.

measured by ETF fees and various types of risk that investors may face, including the degree of ETFs' deviation from their benchmarks in terms of holdings (i.e., active share of Cremers and Petajisto 2009, labeled "ActiveShr"), the degree of ETFs' deviation from their benchmarks in terms of returns (i.e., Tracking error), and the liquidity of ETF trading (i.e., illiquidity of Amihud 2002, labeled as log(fund illiquidity)). In addition, we also examine the performance generated by deviation in holdings (i.e., ActiveShr performance) as a proxy for the gross benefit that various mechanisms can generate. Jointly, these variables can help us understand how benefits and costs/risks associated with ETF activeness are shared between investors and affiliated financial conglomerates.

5.1 Mechanisms of proconglomerate incentives versus fees

ETFs are only required to deliver gross-of-fee returns as high as index returns, so reductions in fees provide the only way for ETFs to benefit their investors. We first investigate the link between ETF fees and the four mechanisms of information, subsidization, and security lending. To capture the information benefits accruing from the affiliation with the bank conglomerate, we use $\text{CorporateLoanDummy}_{i,f,t}$, defined as before, to proxy for the information that ETFs may obtain from their affiliated banks based on such banks' processing of corporate loans. To capture the informational advantages for the ETFs to invest in affiliated banks, we define a dummy variable $\text{AffiliatedBankStockDummy}_{i,f,t}$ that equals 1 if ETF f holds the stock of its affiliated bank i in year t and 0 otherwise. To proxy for the need to engage in cross-trades with the affiliated OEFs, we define a variable $\text{AffiliatedOEFPerformance}_{f,t}$ that equals the lagged TNA-weighted average benchmark-adjusted return of all other OEFs affiliated with the ETF, where the benchmark-adjusted OEF return is computed as the OEF returns minus the average return of OEFs tracking the same benchmark. Given that the need to help the affiliated OEFs concentrates in periods when they underperform, this variable can be used to detect the incentives for ETFs to deviate from their benchmarks to subsidize their affiliated OEFs when the latter have experienced poor performance. To capture the potential security lending income generated by ETFs, we define a variable $\text{SecurityLendingFee}_{i,t}$ that equals the loan-value-weighted average short-selling lending fee of stock i in year t. Most of our tests are conducted at the stock level, so here we also provide similar tests by aggregating ETF fees and proconglomerate mechanisms at the stock level. We then estimate the annual panel regression as follows:

$$\text{Fee}_{i,t} = \alpha + \beta \times \text{Channel}_{i,t-1} + \gamma \text{M}_{i,t-1} + \text{e}_{i,t}, \tag{6}$$

where the dependent variable measures fees (annual expense ratio) charged by ETF and $Channel_{i,t-1}$ is the vector that contains proxies for proconglomerate mechanisms (i.e., Corporate loan dummy, Affiliated bank stock dummy, Affiliated OEF performance, and Security lending fee). When applicable, the stock-level measures involving fund characteristics are computed as the investment value-weighted average of fund characteristics for all funds that invest in the stock. The vector M stacks all other stock and fund control variables defined before. We estimate a panel specification with year and stock fixed effects and clustering at the stock level.

Table 7 reports the results. In Models 1 to 4, we separately report the four channels, whereas we consider a specification with all the channels in Model 5. Model 6 reports similar regression parameters in joint models when we replace stock Turnover with log(stock illiquidity) as an alternative control for the Amihud illiquidity of stocks. We find that the benefit of the information channel does not accrue to ETF investors; that is, the Corporate loan dummy is uncorrelated with fees. Additionally, we do not find a direct link between fees and excess investment in the stock of affiliated banks. Instead, we find a strong negative relationship between fees and Affiliated OEF performance: every 1% negative return of affiliated OEFs is associated with 14 bps of additional ETF fees (Model 3). That the negative performance of affiliated OEFs implies a higher chance for ETFs to subsidize OEFs means the negative correlation could be interpreted as the cost of subsidization. In addition, it is interesting to notice that investors get some solace from the stock lending activity. Indeed, security lending fee is negatively related to ETF fees. In particular, a 1-standard-deviation increase in Security lending fee translates into 0.6 bps lower fees that the ETF charges its investors (Model 4). Overall, these results suggest that ETF investors enjoy direct benefits from the security lending channel and may face some cost due to the subsidization channel. Meanwhile, information channel does not seem to harm investors.

5.2 Mechanisms of proconglomerate incentives versus tracking risk

After we understand the fee implication of proconglomerate mechanisms from investors' perspective, we now move on to test how ETF activeness affects its tracking risk. The ETF tracking risk is measured by its deviation in holdings, that is, active share (ActiveShr) constructed following Cremers and Petajisto (2009), and deviation in returns, that is, Tracking error. For Tracking error, we obtain the fund's total return (net of fees) in U.S. dollars from Morningstar. We add back the fees, and we refer to the resultant gross-of-fee return as the NAV-based return.[17] We then define Tracking error as the standard deviation of the difference between the monthly ETF gross-of-fee

[17] When a portfolio has multiple share classes, we compute its total return as the lagged TNA-weighted return of all the share classes of the portfolio. Similarly, we construct the gross-of-fee benchmark return by using the index funds that track the same benchmark.

Table 7
Impact of proconglomerate incentives on fees (stock level)

Out-of-sample fees (in %) regressed on stock and ETF characteristics

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Corporate loan dummy	0.001				0.000	−0.002
	(0.09)				(0.08)	(−0.25)
Affiliated bank stock dummy		0.026			0.022	0.019
		(0.73)			(0.72)	(0.62)
Affiliated OEF performance			−0.142***		−0.142***	−0.141***
			(−21.73)		(−21.75)	(−21.66)
Security lending fee				−0.005***	−0.005***	−0.005***
				(−5.66)	(−5.79)	(−5.75)
log(stock size)	0.005***	0.005***	0.005***	0.005***	0.004***	0.003***
	(4.23)	(4.23)	(3.69)	(3.87)	(3.31)	(2.75)
Stock return	−0.000***	−0.000***	−0.000***	−0.000***	−0.000***	−0.000***
	(−4.69)	(−4.69)	(−3.90)	(−4.59)	(−3.80)	(−3.51)
Turnover	−0.001	−0.001	0.001	−0.000	0.002	
	(−0.13)	(−0.13)	(0.28)	(−0.01)	(0.41)	
log(stock illiquidity)						−0.010***
						(−3.99)
log(net income)	−0.000	−0.000	−0.000	−0.000	−0.000	−0.000*
	(−1.38)	(−1.38)	(−1.26)	(−1.61)	(−1.49)	(−1.76)
log(sales)	−0.006***	−0.006***	−0.005***	−0.006***	−0.005***	−0.005***
	(−3.37)	(−3.38)	(−3.08)	(−3.36)	(−3.08)	(−3.16)
log(total assets)	0.008***	0.008***	0.008***	0.008***	0.008***	0.008***
	(4.17)	(4.18)	(4.20)	(4.15)	(4.19)	(4.26)
log(fund TNA)	−0.007***	−0.007***	−0.008***	−0.007***	−0.008***	−0.009***
	(−8.31)	(−8.32)	(−10.04)	(−8.48)	(−10.24)	(−10.94)
log(fund age)	−0.009***	−0.009***	−0.012***	−0.008***	−0.011***	−0.010***
	(−5.91)	(−5.92)	(−8.04)	(−5.53)	(−7.57)	(−6.87)
Fund flow	0.000***	0.000***	0.000***	0.000***	0.000***	0.000**
	(3.55)	(3.55)	(2.94)	(3.49)	(2.87)	(2.54)
Intercept	0.467***	0.467***	0.517***	0.472***	0.523***	0.577***
	(25.01)	(24.98)	(27.88)	(25.26)	(28.21)	(25.96)
R-squared	.354	.354	.368	.355	.369	.370
Obs	46,527	46,527	46,527	46,527	46,527	46,527

This table presents the results of the following annual panel regressions with year and stock fixed effects and their corresponding t-statistics with standard errors clustered at the stock level,

$$\mathrm{Fee}_{i,t} = \alpha + \beta \mathrm{Channel}_{i,t-1} + \gamma \mathrm{M}_{i,t-1} + \mathrm{e}_{i,t},$$

where $Fee_{i,t}$ refers to the investment value-weighted average of the ETF-level annualized percentage expense ratio across all funds holding stock i in year t, $Channel_{i,t-1}$ refers to four channels of impact: Corporate loan dummy (a dummy variable taking a value of one if it is a lending-related stock), Affiliated bank stock dummy (a dummy variable taking a value of one if the ETF invests in its affiliated bank), Affiliated OEF performance (the benchmark-adjusted return of other affiliated OEFs), and Security lending fee (the average short selling lending fee). The stock-level Corporate loan dummy and Affiliated bank stock dummy (Affiliated OEF performance) are computed as the investment value-weighted average of the ETF-stock-level (ETF-level) proxies across all funds holding a stock. Vector M stacks all other stock and fund control variables, including log(stock size), Stock return, Turnover, log(net income), log(sales), log(total assets), log(fund TNA), log(fund age), and Fund flow. Table A1 provides detailed definitions for each variable.
$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

NAV-based return and its gross-of-fee benchmark return during a particular year. Tracking error is a standard measure used by the market to assess the ability of the fund to replicate the benchmark.

Investors use ETFs to track index returns, so more deviations imply a risk of tracking the underlining index from investors' perspective. Note that ETF investors will view active share as a risk—similar to tracking error except in the holding space—because, unlike the OEF industry, ETF performance will not be automatically distributed to investors unless fees are reduced. More active share in the holding portfolio, in this regard, only imposes a potential risk for an ETF not to deliver index return in bad economic scenarios (i.e., a counterparty risk if the swap counterparty of the ETF fail to deliver index return in bad economic states).

We first investigate the link between ETFs' tracking risk and the four mechanisms of information, subsidization, and security lending. The results are reported in Table 8, panel A, for active share (ActiveShr), and panel B for Tracking error. The layout of the columns is the same as that of Table 7. The results show a strong correlation between the variables that proxy for the channels and the proxies for deviation. In particular, across the different specifications, we find a strong positive relationship between Corporate loan dummy and ActiveShr and a similar pattern between Corporate loan dummy and Tracking error. ETFs that own stock in firms that receive corporate loan services from the bank affiliated with the applicable ETF present a higher ActiveShr (Tracking error) of 11.8% (9.4 bps), which is consistent with the idea that ETFs tend to diverge more in stocks on which they presumably have more information. Also, the stocks of affiliated banks display an ActiveShr (Tracking error) that is higher by 18.9% (30.6 bps).

Moreover, we find a negative relationship between deviation and the performance of the affiliated OEFs, that is, Affiliated OEF performance. A benchmark-adjusted performance of the affiliated OEFs that is worse by 1-standard-deviation raises ActiveShr (Tracking error) by 4.06% (31.5 bps). This negative sign implies that when affiliated OEFs underperform their benchmarks—and are therefore more exposed to investor withdrawals—ETFs tend to deviate more from their indices. Our previous results suggest that the assistance is transferred to OEFs through cross-trades.

Finally, security lending fees are negatively correlated with both ActiveShr and Tracking error. This effect is also economically relevant: a 1-standard-deviation higher level of Security lending fee reduces ActiveShr (Tracking error) by 0.35% (2 bps), which suggests that the benefits accruing from security lending allow the ETF to diverge less. Thus, the fact that the ETF can generate performance by simply holding the benchmark and lending the shares reduces the need to diverge from the benchmark.

5.3 Mechanisms of proconglomerate incentives versus performance and liquidity

Next, we examine the performance and liquidity side of ETF activeness. When positive performance is generated, the ETFs can pass on this

Table 8

Impact of proconglomerate incentives on tracking risk, performance, and liquidity (stock level)

A. Out-of-sample active share regressed on stock and ETF characteristics

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Corporate loan dummy	0.118***				0.117***	0.111***
	(7.65)				(7.61)	(7.37)
Affiliated bank stock dummy		0.189***			0.180***	0.171***
		(4.13)			(4.05)	(3.76)
Affiliated OEF performance			−0.045***		−0.045***	−0.045***
			(−4.73)		(−4.81)	(−4.71)
Security lending fee				−0.003***	−0.003***	−0.003***
				(−6.38)	(−5.83)	(−5.68)
Stock and fund controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	.088	.072	.072	.071	.091	.099
Obs	46,526	46,526	46,526	46,526	46,526	46,526

B. Out-of-sample tracking error (in %) regressed on stock and ETF characteristics

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Corporate loan dummy	0.094***				0.092***	0.080***
	(4.34)				(4.26)	(3.68)
Affiliated bank stock dummy		0.306***			0.290***	0.275***
		(6.82)			(5.95)	(5.94)
Affiliated OEF performance			−0.349***		−0.350***	−0.352***
			(−11.85)		(−11.88)	(−11.97)
Security lending fee				−0.016***	−0.016***	−0.017***
				(−7.91)	(−7.79)	(−8.00)
Stock and fund controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	.378	.378	.382	.378	.384	.383
Obs	46,526	46,526	46,526	46,526	46,526	46,526

C. Out-of-sample active share performance (in %) regressed on stock and ETF characteristics

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Corporate loan dummy	0.057***				0.057***	0.066***
	(4.48)				(4.47)	(5.03)
Affiliated bank stock dummy		−0.020			−0.032	−0.018
		(−0.46)			(−0.61)	(−0.36)
Affiliated OEF performance			−0.335***		−0.335***	−0.336***
			(−12.22)		(−12.23)	(−12.34)
Security lending fee				−0.002	−0.002	−0.003*
				(−1.38)	(−1.51)	(−1.67)
Stock and fund controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	.244	.243	.250	.243	.251	.252
Obs	46,434	46,434	46,434	46,434	46,434	46,434

D. Out-of-sample Amihud illiquidity regressed on stock and ETF characteristics

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Corporate loan dummy	−0.072				−0.074	−0.084
	(−0.93)				(−0.95)	(−1.09)
Affiliated bank stock dummy		0.049			0.056	0.037
		(0.17)			(0.20)	(0.13)
Affiliated OEF performance			0.163		0.163	0.170
			(1.56)		(1.56)	(1.61)
Security lending fee				−0.010	−0.010	−0.009
				(−0.58)	(−0.58)	(−0.52)

(continued)

Table 8
Continued

D. Out-of-sample Amihud illiquidity regressed on stock and ETF characteristics

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Stock and fund controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	.365	.365	.365	.365	.365	.365
Obs	46,392	46,392	46,392	46,392	46,392	46,392

Panel A presents the results of the following annual panel regressions with year and stock fixed effects and their corresponding t-statistics with standard errors clustered at the stock level,

$$\mathrm{ActiveShr}_{\mathrm{i,t}} = \alpha + \beta \mathrm{Channel}_{\mathrm{i,t-1}} + \gamma \mathrm{M}_{\mathrm{i,t-1}} + \mathrm{e}_{\mathrm{i,t}},$$

where $ActiveShr_{i,t}$ refers to the average quarterly active share of stock i in year t, $Channel_{i,t-1}$ refers to the four channels of impact described in Table 7. Vector M stacks all other stock and fund control variables, including log(stock size), Stock return, Turnover, log(net income), log(sales), log(total assets), log(fund TNA), log(fund age), Expense ratio, and Fund flow. The stock-level ActiveShr is computed as the investment value-weighted average of the ETF-stock-level active share across all funds holding a stock. Panels B to D report similar regression parameters when the dependent variable is Tracking error (the investment value-weighted average of the ETF-level tracking error), ActiveShr performance (the investment value-weighted average of the ETF-level active share performance), and log(fund illiquidity) (the investment value-weighted average of the logarithm of ETF-level Amihud illiquidity). Only the main variables are tabulated for brevity. Table A1 provides detailed definitions for each variable.
*p < .1; **p < .05; ***p < .01.

additional benefit to the investors as reduced fees. Therefore, we must consider ETF performance jointly with fees to derive an overall picture.

We begin with performance (ActiveShr performance) and report the results in panel C of Table 8. The layout of the columns is the same as that of Table 7. We find that the information channel is positively related to performance. This result is expected as higher quality information derived from bank loans can be used to generate performance that outperforms the benchmark. In contrast, we do not find a link between performance and either Affiliated bank stock dummy or Security lending fee. Moreover, affiliated OEF performance is negatively related to ETF performance, implying a “substitute” effect between ETF and OEF return. If we consider these results jointly with the fee results in Table 7, we can gauge both the overall benefit of ETF active share (ActiveShr) and the part that is passed on to the ETF investors. We find that informational advantage derived from lending relationship seems to benefit the affiliated financial conglomerates but is not transferred to investors, while ETFs use the benefits accruing from security lending to reduce the fees charged to their investors.

We also examine how ETF activeness affects its liquidity (log(fund illiquidity)), and report the results in panel D of Table 8. The layout of the columns is the same as that of Table 7. None of the four mechanisms of information, subsidization, and security lending is the related to ETF liquidity. Hence, ETF activeness does not affect trading liquidity regardless of the underlying motivations.

Overall, our results suggest that ETF investors expose to both costs and benefits from ETF activeness associated with proconglomerate incentives. Among the three broad channels of ETF activeness we examined, the security lending channel benefits ETF investors via reduced fees and lower degrees of tracking errors, whereas the information and subsidization channels may expose investors to higher tracking errors without direct benefit in terms of fees. The subsidization channel may even impose some additional cost in terms of fees. These observations suggest that ETF off-benchmark activeness, except for security lending, may involve a transfer between the ETF and its sponsor (and affiliated OEFs). This provides a sort of "pool of capital" to the affiliated financial conglomerate, which, in principle, can be invested in anything. Synthetic operations, such as swaps, allow the conglomerate to deliver the committed return of the benchmark to the ETF investors and—in return—to receive whatever performance can be generated by the actual holdings of the ETF.

6. The Market's Reaction to ETF's Activeness

Our findings suggest that ETFs deviate from the benchmark to leverage their information advantage from the affiliated bank, and to help other members of their financial conglomerate. Although the information-motivated active share may boost performance, subsidization channel might lead to inferior performance during those very periods in which the affiliated bank or OEFs are most in need of subsidization. The existence of the swap with the affiliated bank is designed to protect ETF investors from such risks, but the potential distress of the bank at a time when the performance of the ETF portfolio is particularly poor may nonetheless expose ETF investors to credit risk. Therefore, the remaining question is whether this proconglomerate incentive is perceived by sophisticated ETF investors as detrimental.

To answer this question, we relate the ETF flows[18]—a proxy for the sophisticated ETF investor demand—to the potential impact of ETF proconglomerate incentives in the following regression with year fixed effects and clustering at the fund level:

$$\begin{aligned} \mathrm{Flow}_{f,t} = {} & \alpha + \beta_1 \mathrm{Welfare}_{f,t} + \beta_2 \mathrm{Rating}_{f,t} + \beta_3 \mathrm{Welfare}_{f,t} \times \mathrm{Rating}_{f,t} \\ & + \gamma \mathrm{M}_{f,t-1} + e_{f,t}, \end{aligned} \tag{7}$$

where $\mathrm{Flow}_{f,t}$ refers to the average monthly flows of ETF f in year t; $\mathrm{Welfare}_{f,t}$ refers to the potential impact of ETF off-benchmark activities which we will

[18] Although ETF investors typically exit by selling the ETF in the market as opposed to redeeming the shares (see Footnote 3), investors can nonetheless create inflows and outflows at the fund level. This feature allows us to use ETF flows to proxy for fund demand. We compute monthly ETF flows as $\mathrm{Flow}_{f,m} = \left[\mathrm{TNA}_{f,m} - \mathrm{TNA}_{f,m-1} \times \left(1 + \mathrm{R}_{f,m}\right)\right] / \mathrm{TNA}_{f,m-1}$, where $\mathrm{TNA}_{f,m}$ refers to the total net asset of fund f in month m, and $\mathrm{R}_{f,m}$ refers to fund total return in the same month. Annual ETF flows are computed as the average of monthly flows within a year. In additional robustness checks, we also compute the flows using annual frequency. The results do not change.

specify shortly. $\text{Rating}_{f,t}$ refers to the S&P long-term domestic issuer credit rating of its affiliated bank and proxies for distress risk (we also use bank ROA to replace bank rating in a few specifications). We follow Avramov et al. (2009) in creating our bank rating score, which transforms the S&P ratings into ascending numbers as follows: AAA = 1, AA+ = 2, AA = 3, AA– = 4, A+ = 5, A = 6, A– = 7, BBB+ = 8, BBB = 9, BBB– = 10, BB+ = 11, BB = 12, BB– = 13, B+ = 14, B = 15, B– = 16, CCC+ = 17, CCC = 18, CCC– = 19, CC = 20, C = 21, and D = 22. Vector M stacks the control variables.[19]

Two important effects emerge: any additional benefits/costs generated by the off-benchmark activities of ETFs can be transferred to investors or to affiliated conglomerates. We have already discussed that ETF fee provides a reasonable proxy for the beneficial effect that can be received by ETF investors. What is still missing is a measure on the beneficial impact of such activities on the affiliated conglomerates. Ideally, we will need explicit cash flow transaction data between ETFs and affi (e.g., cross-trades). To further verify this interpretation, we test the flow implication of Swapped transfer side by side with ETF Fees. We also directly test the flow implication of tracking risk (i.e., ActiveShr, Tracking error) in order to understand whether investors care more about tracking risk per se or its implication in terms of subsidization.

Table 9 presents the results. Models 1 to 5 illustrate that ETF flows are uncorrelated with ActiveShr or Tracking error but are negatively related to both Swapped transfer and Fees. It is reasonable that investors are not particularly worried about tracking risk per se because tracking risk may be related to superior information, as we have discussed above, which does not necessarily hurt investors. However, positive Swapped transfer and higher Fees signal a net detrimental effect to the investor, and investors respond to such net negative effects by withdrawing capital. An increase in the Swapped transfer (Fees) of 1-standard-deviation is associated with a lower annual flow of 3.56% (8.61%) in Model 3 (Model 4). These results suggest that investors consider these negative effects to be detrimental.

In addition, Model 6 reports a negative relationship between flows and bank rating (recall that a higher numerical value means a lower rating). A 1-standard-deviation deterioration in bank rating translates to 9.18% lower flows per year. The fact that ETF investors withdraw capital when affiliated banks have poor ratings suggests that investors view the affiliation with a bad bank as detrimental. This result is not surprising because both the incentive for subsidization and the risk (for the affiliated bank) to default on the promised index return are amplified in poor ratings. Meanwhile, if deteriorating bank ratings appear detrimental to the investors, then deterioration in bank performance should also appear detrimental to the investors. Model 7 verifies this equivalence by replacing bank rating with ROA. We observe that negative ROA is associated with outflows, which is a pattern that is consistent with what we observe with bank rating.[21]

More importantly, from the perspective of investors, the detrimental impact of Swapped transfer should be more significant when the affiliated banks are riskier. Models 8 to 11 test this intuition by interacting Swapped transfer with bank rating or ROA. Indeed, we observe that the outflow sensitivity with respect to Swapped transfer increases in the poor ratings/ROAs of affiliated banks. Thus, investors do not seem to appreciate the links between ETFs and affiliated banks, particularly when Swapped transfer signals potential conflicts of interest and when the banks become riskier. Models 12 to 14 confirm that the results remain unchanged after controlling for the four mechanisms of information, subsidization, and security lending as discussed earlier. As a robustness check, we also estimate a Fama and MacBeth (1973)

[21] To validate the interpretation of ROA, we created a dummy variable that equals 1 when bank ROA is below the median. Unreported results show that below-median bank ROA discourage monthly flows by 4.43% in the affiliated ETFs.

Table 9
ETF flows (ETF level)

ETF flow (in %) regressed on ETF and affiliated bank characteristics

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14
ActiveShr	0.161				−0.282					−0.225	−0.397	−0.196	−0.246	−0.549
	(0.34)				(−0.55)					(−0.38)	(−0.75)	(−0.39)	(−0.38)	(−1.05)
Tracking error		0.527			−0.909*					−0.305	−0.605	−0.663	−0.375	−0.893*
		(1.00)			(−1.86)					(−0.45)	(−1.29)	(−1.14)	(−0.55)	(−1.78)
Swapped transfer			−0.654**		−0.737***			0.728*	−1.155***	0.703*	−1.242***	−0.224	0.641	−1.245***
			(−2.44)		(−2.78)			(1.72)	(−2.69)	(1.67)	(−2.97)	(−1.08)	(1.53)	(−2.90)
Expense ratio				−5.517***	−5.675***					−4.027***	−4.300***	−7.835***	−5.525***	−6.185***
				(−3.93)	(−4.13)					(−3.14)	(−3.13)	(−5.12)	(−3.94)	(−4.31)
Bank rating						−0.575***		−0.821***		−0.718***			−0.559***	
						(−3.70)		(−5.89)		(−5.02)			(−3.63)	
Bank ROA							2.309***		0.858**		0.314			−0.040
							(2.62)		(2.48)		(0.85)			(−0.11)
Swapped transfer × Bank rating								−0.212***		−0.202***			−0.167**	
								(−2.89)		(−2.77)			(−2.27)	
Swapped transfer × Bank ROA									0.661**		0.670**			0.722**
									(2.16)		(2.23)			(2.36)
Corporate loan dummy												2.133**	1.420	1.795**
												(2.29)	(1.44)	(1.99)
Affiliated bank stock dummy												−1.908***	−1.727**	−0.704
												(−3.32)	(−2.20)	(−1.30)
Affiliated OEF performance												4.149***	3.028**	3.417**
												(2.67)	(2.03)	(2.22)
Security lending fee												1.413***	1.172**	1.100*
												(2.69)	(2.27)	(1.76)

(continued)

Table 9
Continued

ETF flow (in %) regressed on ETF and affiliated bank characteristics

	[illegible]	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14

specification with Newey and West (1987) adjustment. The (unreported) results are similar to the reported results.

Overall, these findings show that the market is aware of the potential implications of the link between ETFs and their affiliated financial conglomerates. It appears that investors' concerns, which are expressed in lower flows, are consistent with regulatory concerns (FSB 2011; IMF 2011; Ramaswamy 2011).

7. Additional Analyses and Robustness Checks

We finally provide additional analysis on the relationship between ETF off-benchmark incentives and market development. Intuitively, the relationship could be two-way. On the one hand, the off-benchmark incentives of ETFs can shape the development of the ETF industry by influencing fund family's decisions to launch new ETFs. On the other hand, financial market development may also affect the incentives for ETF families to share benefits with

Stock market/GDP, and Private bond market/GDP. When a fund family has multiple funds, we aggregate the TNA from all funds within the family as family TNA. All other family characteristics (such as expense ratio, return, and flow) are computed as the lagged TNA-weighted fund characteristics of all OEFs within the family. We include year and country fixed effects and cluster the standard errors at the family level.

We report the results in Table 10, with Models 1 to 6 for logistic specifications and Models 7 to 12 for probit specifications. We find that across all specifications, the launch of full replication ETF is negatively related to the cross-trades between existing OEFs within the family (OEF/OEF cross-trades). This implies that the subsidization needs to cross-trade with affiliated OEFs play an important role in determining the type of ETF inception. In contrast, the mechanisms of information and security lending do not affect the ETF launch decision. We can see that ETF off-benchmark incentives may have played an important role in shaping the development of the ETF industry.

7.2 Impact of market development

The global financial industry has been growing rapidly over time, and the growth rate is much higher in ETF industry comparing with OEFs. More importantly, do ETF investors benefit from the overall development in financial market? As before, we focus on ETF fees to examine the welfare implication of investors. We estimate the following annual panel regression:

$$\begin{aligned}\mathrm{Fee}_{i,t} = \alpha + \beta_1 \mathrm{Channel}_{i,t-1} + \beta_2 \mathrm{Channel}_{i,t-1} \times (\mathrm{ActiveETF/OEF})_{c,t-1} \\ + \beta_3 (\mathrm{ActiveETF/OEF})_{c,t-1} + \gamma \mathrm{M}_{i,t-1} + \mathrm{e}_{i,t},\end{aligned} \tag{9}$$

where $(\mathrm{ActiveETF/OEF})_{c,t-1}$ refers to the TNA of active ETFs (including synthetic replication and optimized sampling ETFs) as a percentage of OEFs in country c (where stock i is traded) in year t. ActiveETF/OEF captures the relative importance of active ETFs comparing with the OEF industry. All other variables are defined as in Equation (6), except for we further control for financial market development, proxied by ETF TNA/GDP, defined as the end-of-year TNA of ETFs divided by nominal GDP; and OEF TNA/GDP, defined as the end-of-year TNA of OEFs divided by nominal GDP. We estimate a panel specification with year and stock fixed effects and clustering at the stock level.

Table 11 reports the results. We first confirm that ETF investors enjoy direct benefits from the security lending channel and may face some cost due to the subsidization channel, after controlling for the financial market development. Second, the growth of active ETFs (comparing with OEFs) encourages more income sharing between ETFs and their investors. ETFs charge lower fees when they benefit from information channel, that is, through their investment in loan-related stocks and affiliated bank. Finally, the growth of

Table 10
The decision to launch full replication ETFs

Out-of-sample launch of full replication ETFs regressed on family characteristics

	Logistic						Probit					
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
Bank affiliation dummy	−0.555					−0.354	−0.396					−0.275
	(−0.77)					(−0.48)	(−0.91)					(−0.64)
Corporate loan ActiveShr		0.057				0.007		0.040				0.016
		(0.97)				(0.07)		(1.07)				(0.31)
Affiliated bank stock ActiveShr			5.317*			5.657			3.193**			3.181*
			(1.84)			(1.41)			(2.27)			(1.87)
OEF/OEF cross-trades				−0.162***		−0.154*				−0.099***		−0.089**
				(−2.69)		(−1.86)				(−3.20)		(−2.41)
OEF security lending fee					0.581	−1.743					0.037	−1.093
					(0.26)	(−0.54)					(0.03)	(−0.79)
log(family TNA)	0.555**	0.548***	0.510***	0.636***	0.574***	0.548***	0.304***	0.301***	0.289***	0.368***	0.315***	0.311***
	(2.49)	(2.66)	(2.67)	(3.30)	(2.71)	(2.77)	(2.73)	(3.04)	(2.98)	(3.84)	(3.05)	(3.03)
log(family age)	1.485	1.482	1.525	0.981	1.370	1.598	0.825	0.815	0.797	0.439	0.785	0.852
	(1.47)	(1.42)	(1.35)	(0.86)	(1.46)	(1.49)	(1.56)	(1.49)	(1.34)	(0.71)	(1.53)	(1.45)
Family expense ratio	0.015	0.019	0.014	0.009	0.022	−0.008	0.010	0.012	0.010	0.007	0.013	−0.000
	(0.64)	(1.04)	(0.70)	(0.41)	(1.10)	(−0.15)	(0.91)	(1.28)	(1.12)	(0.75)	(1.27)	(−0.01)
Family return	−0.151	−0.164	−0.225	−0.199	−0.165	−0.236	−0.125	−0.131	−0.162	−0.138	−0.124	−0.146
	(−0.80)	(−0.86)	(−1.12)	(−0.74)	(−1.00)	(−1.01)	(−1.18)	(−1.22)	(−1.44)	(−1.10)	(−1.25)	(−1.31)
Family flow	−0.016	−0.014	−0.010	−0.009	−0.013	−0.007	−0.009	−0.007	−0.006	−0.005	−0.007	−0.005

(continued)

Table 10
Continued

Out-of-sample launch of full replication ETFs regressed on family characteristics

	Logistic						Probit					
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
	(−1.46)	(−1.19)	(−1.05)	(−0.88)	(−1.21)	(−0.83)	(−1.47)	(−1.08)	(−1.02)	(−0.78)	(−1.12)	(−0.93)
Stock market	−0.761	−0.803	−0.945	−1.392*	−0.818	−1.369*	−0.443	−0.495	−0.575	−0.812*	−0.482	−0.764
turnover	(−1.00)	(−1.07)	(−1.25)	(−1.82)	(−1.09)	(−1.67)	(−1.03)	(−1.17)	(−1.34)	(−1.82)	(−1.13)	(−1.59)
Stock market/	1.930	2.282	2.049	3.020	2.315	2.202	1.041	1.284	1.110	1.612	1.220	1.165
GDP	(0.83)	(1.05)	(0.96)	(1.17)	(0.96)	(0.83)	(0.87)	(1.12)	(0.97)	(1.14)	(1.04)	(0.82)
Private bond	−2.324	−2.046	−1.339	−3.897	−2.180	−2.784	−1.130	−0.932	−0.562	−2.045*	−1.007	−1.510
market/GDP	(−0.96)	(−0.83)	(−0.58)	(−1.52)	(−0.91)	(−1.17)	(−1.01)	(−0.80)	(−0.49)	(−1.80)	(−0.89)	(−1.37)
Intercept	−19.036**	−19.589**	−19.782**	−11.971	−19.522**	−14.328**	−10.802***	−11.134***	−11.172***	−6.490	−11.162***	−7.975**
	(−2.28)	(−2.38)	(−2.39)	(−1.48)	(−2.39)	(−1.98)	(−2.63)	(−2.72)	(−2.72)	(−1.46)	(−2.83)	(−2.00)
Obs	127	127	127	127	127	127	127	127	127	127	127	127

This table presents the results of the following annual logistic or probit regressions with year and country fixed effects and their corresponding t-statistics with standard errors clustered at the family level:

$$\mathrm{FullRep}_{\mathrm{F,t}} = \alpha + \beta \mathrm{Fam\_Char}_{\mathrm{F,t-1}} + \gamma \mathrm{M}_{\mathrm{F,t-1}} + \mathrm{e}_{\mathrm{F,t}},$$

where $FullRep_{F,t}$ refers to a dummy variable that equals 1 if the mutual fund family F launches full replication ETF(s) in year t and 0 otherwise, and $Fam\_Char_{F,t-1}$ refers to the family characteristics including Bank affiliation dummy (a dummy variable equal to 1 if the ETF is affiliated with a bank conglomerate), Corporate loan ActiveShr (the average active share on lending-related stocks held by affiliated OEFs), Affiliated bank stock ActiveShr (the average active share on affiliated bank held by affiliated OEFs), OEF/OEF cross-trades (the average cross-trades of between affiliated OEFs), and OEF security lending fee (the investment value-weighted average of short selling lending fee based on affiliated OEF holdings). Vector M stacks all other family and country control variables, including log(family TNA), log(family age), Family expense ratio, Family return, Family flow, Stock market turnover, Stock market/GDP, and Private bond market/GDP. Models 1 to 6 present the results of logistic regressions, and Models 7 to 12 present the results of probit regressions. Table A1 provides detailed definitions for each variable.
*p < .1; **p < .05; ***p < .01.

Table 11

Impact of proconglomerate incentives and market development on fees (stock level)

Table 11
Continued

Out-of-sample fees (in %) regressed on stock, ETF, and country characteristics

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
R-squared	.453	.454	.453	.455	.453	.455
Obs	34,111	34,111	34,111	34,111	34,111	34,111

This table presents the results of the following annual panel regressions with year and stock fixed effects and their corresponding t-statistics with standard errors clustered at the stock level,

$$\mathrm{Fee}_{i,t} = \alpha + \beta_1 \mathrm{Channel}_{i,t-1} + \beta_2 \mathrm{Channel}_{i,t-1} \times (\mathrm{ActiveETF/OEF})_{c,t-1} + \beta_3 (\mathrm{ActiveETF/OEF})_{c,t-1} + \gamma \mathrm{M}_{i,t-1} + \mathrm{e}_{i,t},$$

where $Fee_{i,t}$ refers to the investment value-weighted average of the ETF-level annualized percentage expense ratio across all funds holding stock i in year t, $Channel_{i,t-1}$ refers to the four channels of impact described in Table 7. $(ActiveETF/OEF)_{c,t-1}$ refers to the total net asset of active ETFs (including synthetic replication and optimized sampling ETFs) as a percentage of OEFs in country c (where stock i is traded). Vector M stacks all other country, stock, and fund control variables, including ETF TNA/GDP, OEF TNA/GDP, log(stock size), Stock return, Turnover, log(net income), log(sales), log(total assets), log(fund TNA), log(fund age0(rn,)-339.ap2(ta351meyo

active ETFs magnifies the cost of subsidization channel. In particular, when there is a higher chance for ETFs to subsidize OEFs (implied by poor affiliated OEF performance) and active ETFs are more prevailing, ETFs tend to charge higher fees.

Overall, financial market development not only enhances the competition and incentivizes active ETFs to share income with investors to attract capital but also magnifies the subsidization needs and the corresponding cost. Together with the previous tests, we can see that ETF off-benchmark incentives may intertwine with market development in influencing investors' welfare. These findings further highlight the importance for regulators and investors to better understand ETF incentives above and beyond the role of a passive index tracker.

Finally, the development of ETF industry could benefit a broader set of investors that are interested in index-linked investment. Although a comprehensive analysis of welfare implication for all market participants is beyond the scope of the paper, unreported results suggest that index fund investors can benefit from lower fees when ETFs introduce competitions into the index-tracking business. However, we find that only the growth in full replication ETFs is associated with fee reduction from index funds. Active ETFs, by contrast, achieve an opposite effect, if any.

7.3 Robustness checks

As a robustness check, we repeat the main analyses in the paper and cluster the standard errors by time. Table 12 reports the results. Panel A repeats the analyses on stock selection in Table 2 (Models 1 to 2), Table 3 (Models 3 to 4),

Table 12

Robustness checks on ETF stock selection and cross-trades (cluster by time)

A. Out-of-sample DGTW-adjusted stock return (in %) regressed on Δabnormal ETF ownership

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
ΔETFadjOwn(loaned corp)	0.974*	1.014*						
	(1.87)	(2.29)						
ΔETFadjOwn(unloaned corp)	−0.076	−0.043						
	(−0.62)	(−0.40)						
ΔETFadjOwn(affiliated bank)			0.065*	0.063*				
			(1.95)	(2.06)				
ΔETFadjOwn(unaffiliated bank)			−0.029	−0.028				
			(−1.29)	(−1.17)				
ΔETFadjOwn(high OEF-co-ownership corp)					0.369***	0.373***		
					(7.71)	(6.70)		
ΔETFadjOwn(low OEF-co-ownership corp)					0.077	0.088		
					(1.32)	(1.59)		
ΔETFadjOwn(high OEF-cross-trades corp)							−0.068	−0.077
							(−0.28)	(−0.30)
ΔETFadjOwn(low OEF-cross-trades corp)							0.244***	0.255***
							(5.45)	(4.75)
log(stock size)	−3.089***	−3.085***	0.659	0.435	−3.128***	−3.131***	−3.127***	−3.130***
	(−9.99)	(−10.24)	(0.60)	(0.38)	(−12.43)	(−12.95)	(−12.45)	(−12.96)
Turnover	−1.387**	−1.313*	2.237	1.891	−1.811***	−1.805***	−1.816***	−1.809***
	(−2.33)	(−2.14)	(0.78)	(0.63)	(−3.55)	(−3.59)	(−3.56)	(−3.60)
log(net income)	0.078**	0.079**	0.072	0.085	0.076***	0.076***	0.076***	0.077***
	(3.00)	(3.05)	(1.37)	(1.74)	(3.45)	(3.45)	(3.46)	(3.46)
log(sales)	0.181	0.181	0.214	0.242	0.083	0.081	0.083	0.081
	(1.49)	(1.45)	(0.24)	(0.25)	(0.62)	(0.60)	(0.62)	(0.61)
log(total assets)	−0.119	−0.143	−0.452	−0.514	−0.208	−0.215	−0.209	−0.215
	(−0.84)	(−0.94)	(−0.53)	(−0.53)	(−1.35)	(−1.37)	(−1.36)	(−1.38)
log(fund TNA)		−0.068		0.063		−0.039		−0.039
		(−0.86)		(0.15)		(−0.82)		(−0.81)
log(fund age)		0.217		1.778		0.289*		0.295*
		(1.23)		(1.47)		(2.11)		(2.14)
Expense ratio		0.418		−0.941		0.486		0.513
		(0.67)		(−0.45)		(0.43)		(0.45)
Fund return		−0.141		0.188		−0.030		−0.028
		(−0.97)		(0.70)		(−0.38)		(−0.35)

(continued)

Table 12
Continued

A. Out-of-sample DGTW-adjusted stock return (in %) regressed on Δabnormal ETF ownership

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Fund flow		−0.012		0.028		0.006		0.006
		(−0.72)		(0.83)		(0.43)		(0.44)
R-squared	.168	.170	.147	.152	.175	.176	.175	.176
Obs	46,863	46,863	785	785	40,608	40,608	40,608	40,608

B. Two-stage OEF flow (in %) and performance (in %) regression (OEF level)

	First-stage	Second-stage						
	ETF/OEF cross-trades Model 1	Flow Model 2	BMK-adjusted volatility Model 3	Return Model 4	BMK-adjusted Model 5	DGTW Model 6	CAPM Model 7	FFC Model 8
ETF/OEF Co-ActiveShr	11.380***							
	(7.71)							
ETF/OEF cross-trades		0.251***	0.036*	0.107***	0.008	−0.014**	0.010	0.000
		(3.11)	(1.75)	(4.12)	(0.96)	(−2.29)	(1.26)	(0.08)
log(stock size in fund)	1.452***	−1.142***	−0.325***	−0.501***	−0.081***	−0.073***	−0.142***	0.003
	(3.27)	(−4.30)	(−3.98)	(−5.22)	(−2.71)	(−2.58)	(−4.85)	(0.16)
log(fund TNA)	0.278	−1.619***	−0.129***	0.124**	0.010	0.009	0.047**	0.047***
	(0.80)	(−6.40)	(−2.82)	(2.00)	(0.59)	(0.54)	(2.47)	(4.30)
log(fund age)	−0.483	0.415	0.100	0.174	0.047	0.026	0.057	−0.053*
	(−0.55)	(1.01)	(1.23)	(1.33)	(1.02)	(0.73)	(1.21)	(−1.81)
Expense ratio	3.656***	−1.760***	−0.071	0.078	0.035	0.097***	0.255***	0.216***
	(4.48)	(−3.86)	(−0.82)	(0.57)	(1.00)	(2.76)	(6.40)	(7.14)
OEF return	−0.350***	0.079	−0.087***	−0.524***	−0.033**	0.022*	−0.015	0.059***
	(−3.07)	(0.84)	(−4.15)	(−16.09)	(−2.28)	(1.91)	(−0.96)	(6.28)
Fund flow	−0.144***	0.281***	0.022***	0.048***	0.007*	−0.001	0.008**	0.001
	(−3.85)	(5.77)	(2.65)	(5.80)	(1.89)	(−0.24)	(2.08)	(0.29)
Intercept	−16.781*	41.685***	5.015***	−0.097	0.253	0.274	−0.266	−1.162***
	(−1.73)	(7.26)	(4.25)	(−0.06)	(0.63)	(0.63)	(−0.62)	(−4.27)
Obs	1,959	1,959	1,653	1,959	1,653	1,959	1,959	1,959

In panel A, Models 1 to 2 repeat the analyses in Table 2, Models 3 to 4, and report the corresponding t-statistics clustered by year. Models 3 and 4 report statistics similar to those reported in Table 3, Models 3 to 4; Models 5 and 6 report statistics similar to those reported in Table 4, Models 3 to 4; and Models 7 and 8 report statistics similar to those reported in Table 4, Models 7 to 8, with t-statistics clustered by year. Panel B report statistics similar to those reported in Table 5, panel A, with t-statistics clustered by year. Table A1 provides detailed definitions of each variable.

*p < .1; **p < .05; ***p < .01.

and Table 4 (Models 5 to 8). Panel B repeats the analyses on cross-trades in Table 5. Our main findings are robust to the alternative clustering approach.

8. Conclusion

Appendix

Table A1
Variable definitions

Variables	Definitions
A. ETF performance measures (in %)	
Holding-based return	The investment-value weighted average of stock returns of a fund's most recently reported holding portfolio
Holding-based DGTW-adjusted return	The investment-value weighted average of stock-level DGTW adjusted returns, according to a fund's most recently reported holding information. More specifically, stock returns are adjusted by the style average, where stock styles are created by double-sorting remaining stocks into 25 independent book-to-market and size portfolios within each country, following Daniel et al. (1997)
ActiveShr performance	ETF holding-based return minus the holding-based return of index funds tracking the same benchmark
Gross-of-fee NAV-based return	Monthly fund total returns as reported by Morningstar plus one-twelfth of the annualized expense ratio. When a portfolio has multiple share classes, its total return is computed as the share class TNA-weighted return of all share classes, in which the total net asset (TNA) values are 1-month lagged
Swapped transfer	Holding-based return minus gross-of-fee NAV-based return
Fund return	Monthly affiliated ETF total returns as reported by Morningstar. When a portfolio has multiple share classes, its total return is computed as the share class TNA-weighted return of all share classes, where the TNA values are 1-month lagged
Benchmark-adjusted return	ETF return minus the return of index funds tracking the same benchmark
International Fama-French-Carhart-adjusted return	Realized fund returns minus the productions between a fund's four-factor betas multiplied by the realized four factor returns in a given month. The four international factors are the value weighted average of four domestic Fama-French-Carhart factors (market, size, book-to-market, and momentum). The betas of the fund are estimated as the exposures of the fund to the relevant risk factors in its entire sample period
B. ETF characteristics	
ActiveShr	Active share in a given year t is computed as follows: $\mathrm{ActiveShr}_{f,t} = \sum_{i \in f} \mathrm{DivStock}_{i,f,t} = \sum_{i \in f} \left| w_{i,f,t} - \widehat{w}_{i,f,t} \right| / 2$, where $\mathrm{DivStock}_{i,f,t}$ refers to the stock-level active share of stock i in fund f in year t, $w_{i,f,t}$ is the investment weight of stock i by fund f in year t, and $\widehat{w}_{i,f,t}$ is the benchmark investment weight. When quarterly or semi-annual holdings are available, stock-level active share is computed first at the quarterly or semi-annual level and then averaged within a year
Tracking error (in %)	Tracking error in a given year t is computed as the standard deviation of the difference between monthly ETF gross-of-fee NAV-based return and its gross-of-fee benchmark index return

log(fund illiquidity)	The logarithm of annual fund illiquidity. The fund illiquidity measure in a given month m is computed as follows: $\mathrm{ILLIQ}_{f,m} = \left(\sum_{d \in m} \lvert R_{f,d,m} \rvert / \mathrm{VOLD}_{f,d,m}\right)/D_{f,m} \times 10^6$, where $R_{f,d,m}$ refers to the percentage return of fund f in day d of month m, $\mathrm{VOLD}_{f,d,m}$ refers to the dollar trading volume at the same time, and $D_{f,m}$ is the number of trading days for fund f in month m, following Amihud (2002). The annual fund illiquidity is the average of the monthly fund illiquidity within a year
log(stock size in fund)	The logarithm of the value weighted average of market capitalization, in millions, of stocks in a fund's most recently reported holding portfolio
log(fund TNA)	The logarithm of total net asset as reported in Morningstar
log(fund age)	The logarithm of the number of operational months since inception
Expense ratio (in %)	The annualized expense ratio as reported in Morningstar
Fund flow (in %)	Fund flow in a given month m is computed as follows: $\mathrm{Flow}_{f,m} = \left[\mathrm{TNA}_{f,m} - \mathrm{TNA}_{f,m-1} \times \left(1 + R_{f,m}\right)\right]/\mathrm{TNA}_{f,m-1}$, where $\mathrm{TNA}_{f,m}$ refers to the total net asset of fund f in month m, and $R_{f,m}$ refers to fund total return in the same month. The annual ETF flow is the average of monthly flows within a year
ETF premium (in %)	ETF premium in a given month m is computed as follows: $\mathrm{Premium}_{f,m} = \left(\mathrm{Price}_{f,m} - \mathrm{NAV}_{f,m}\right)/\mathrm{NAV}_{f,m}$, where $\mathrm{Price}_{f,m}$ refers to the market price of fund f in month m, $\mathrm{NAV}_{f,m}$ refers to the net asset value in the same month. The annual premium is the average of monthly premium within a year
C. Affiliated OEF characteristics	
Affiliated OEF performance (benchmark-adjusted, in %)	Monthly total return of affiliated open-end funds (OEFs) as reported by Morningstar, minus the average return of the OEFs tracking the same benchmark. When a portfolio has multiple share classes, its total return is computed as the share class TNA-weighted return of all share classes, where the TNA values are 1-month lagged
D. Cross-trades measures	
ETF/OEF cross-trades (in %)	Cross-trades between ETF i and affiliated OEF j in a given quarter q is computed as follows: $\mathrm{CrossTra}_{ij,q} = \left[\left(\sum_{s \in S_1 \cap S_2} N_{s,i,q} P_{s,q} + N_{s,j,q} P_{s,q}\right) \times I\left\{\Delta N_{s,i,q} \times \Delta N_{s,j,q} < 0\right\}\right]/\left(\sum_{s \in S_1} N_{s,i,q} P_{s,q} + \sum_{s \in S_2} N_{s,j,q} P_{s,q}\right)$, where S_1 and S_2 represent the set of companies held by fund i and j, $P_{s,q}$ is the price of company s at quarter q, $N_{s,i,q}$ and $N_{s,j,q}$ are the number of shares of company s held by fund i and j, respectively, and $I\{\cdot\}$ is an indicator function that equals 1 if $N_{s,i,q}$ and $N_{s,j,q}$ change in opposite directions and 0 otherwise, following Gaspar, Massa, and Matos (2006). Annual cross-trades is the average of quarterly cross-trades within a year
ETF/ETF cross-trades (in %)	Cross-trades between ETF i and affiliated ETF j in a given quarter q is computed as follows: $\mathrm{CrossTra}_{ij,q} = \left[\left(\sum_{s \in S_1 \cap S_2} N_{s,i,q} P_{s,q} + N_{s,j,q} P_{s,q}\right) \times I\left\{\Delta N_{s,i,q} \times \Delta N_{s,j,q} < 0\right\}\right]/\left(\sum_{s \in S_1} N_{s,i,q} P_{s,q} + \sum_{s \in S_2} N_{s,j,q} P_{s,q}\right)$, where all variables are defined like in ETF/OEF cross-trades. Annual cross-trades is the average of quarterly cross-trades within a year

(continued)

Table A1

Continued

Variables	Definitions
E. Affiliated bank characteristics	
Bank rating	The monthly S&P long-term domestic issuer credit rating of the affiliated bank as reported in Compustat. We transform the S&P ratings into ascending numerical scores, where AAA = 1, AA+ = 2, AA = 3, AA– = 4, A+ = 5, A = 6, A– =

Affiliated bank stock ActiveShr	Active share of affiliated bank in a given year t is computed as follows: $\text{ActiveShr(Affiliated Bank)}_{f,t} = \sum_{i \in f} \left| w_{i,f,t} - \widehat{w}_{i,f,t} \right| / 2 \times \text{AffiliatedBankDummy}_{i,f,t}$, where $\text{AffiliatedBankDummy}_{i,}$

Table B1
List of ETF sponsors, 2009

Rank	Conglomerate name for ETF sponsors	Domicile	Bank dummy	TNA (in millions of $)	Market share (in %)
1	Barclays Plc	United Kingdom	1	354,751.15	46.68
2	State Street Corp.	United States	1	165,888.96	21.83
3	Vanguard Group, Inc.	United States	0	79,649.11	10.48
4	Société Générale SA	France	1	32,391.69	4.26
5	INVESCO Ltd.	United States	0	29,107.02	3.83
6	Nomura Holdings, Inc.	Japan	1	12,653.94	1.67
7	American International Group, Inc.	United States	1	11,363.42	1.50
8	MidCap SPDR Trust Services	United States	0	8,484.97	1.12
9	Credit Suisse Group	Switzerland	1	7,296.41	0.96
10	DekaBank Deutsche Girozentrale	Germany	1	5,679.00	0.75
11	Sumitomo Trust & Banking Co. Ltd.	Japan	1	5,577.27	0.73
12	Bank of New York Mellon Corp.	United States	1	5,065.86	0.67
13	Daiwa Securities Group Co. Ltd.	Japan	1	4,889.99	0.64
14	HSBC Holdings Plc	United Kingdom	1	4,695.56	0.62
15	CITIC Securities Co. Ltd.	China	1	4,283.76	0.56
16	Commerzbank AG	Germany	1	4,080.14	0.54
17	UBS AG	Switzerland	1	3,610.78	0.48
18	Guggenheim Capital LLC	United States	1	3,530.26	0.46
19	The Security Benefit Group of Cos.	United States	1	2,724.24	0.36
20	BNP Paribas SA	France	1	2,403.26	0.32
21	First Trust Advisors LP	United States	0	1,974.46	0.26
22	Polaris Securities Co. Ltd.	Taiwan	0	1,939.38	0.26
23	NASDAQ OMX Group, Inc.	United States	1	1,579.42	0.21
24	Svenska Handelsbanken AB	Sweden	1	1,437.14	0.19
25	Banco Bilbao Vizcaya Argentaria SA	Spain	1	1,157.12	0.15
26	BOCI-Prudential Asset Management Ltd.	Hong Kong	1	881.97	0.12
27	AXA SA	France	0	793.13	0.10
28	Rue de la Boetie SAS	France	1	713.58	0.09
29	Crédit Agricole SA	France	1	404.67	0.05
30	DnB NOR ASA	Norway	1	246.25	0.03
31	Fubon Financial Holding Co. Ltd.	Taiwan	1	160.98	0.02
32	RFS Holdings BV	Netherlands	1	150.96	0.02
33	Geode Capital Management LLC	United States	0	134.09	0.02
34	Alpha Bank SA	Greece	1	96.67	0.01
35	DBS Group Holdings Ltd.	Singapore	1	32.08	0.00
36	Bank of Ireland	Ireland	1	31.01	0.00
37	Esposito Partners LLC	United States	0	24.03	0.00
38	The Capital Group Cos., Inc.	United States	1	10.21	0.00
39	Global X Management Co. LLC	United States	0	7.18	0.00
40	Medvesek Pusnik DZU	Slovenia	1	6.77	0.00
41	TMB Bank Public Co. Ltd.	Thailand	1	6.34	0.00
42	ICICI Prudential Asset Management Co. Ltd.	India	1	0.20	0.00

Table C1
Top-three ETF sponsors over time

Year	Rank	Conglomerate name for ETF sponsors	Domicile	TNA (in millions of $)	Market share (in %)
2001	1	State Street Corp.	United States	33,894.65	55.46
2001	2	Barclays Plc	United Kingdom	17,593.23	28.79
2001	3	Nomura Holdings, Inc.	Japan	52,97.33	8.67
2002	1	State Street Corp.	United States	48,344.80	39.00
2002	2	Barclays Plc	United Kingdom	26,933.96	21.73
2002	3	INVESCO Ltd.	United States	17,034.31	13.74
2003	1	State Street Corp.	United States	58,615.59	31.67
2003	2	Barclays Plc	United Kingdom	53,765.70	29.05
2003	3	INVESCO Ltd.	United States	25,689.52	13.88
2004	1	Barclays Plc	United Kingdom	105,739.02	39.87
2004	2	State Street Corp.	United States	75,839.27	28.60
2004	3	INVESCO Ltd.	United States	22,610.88	8.53
2005	1	Barclays Plc	United Kingdom	175,199.32	49.09
2005	2	State Street Corp.	United States	82,445.74	23.10
2005	3	INVESCO Ltd.	United States	23,200.80	6.50
2006	1	Barclays Plc	United Kingdom	263,631.41	53.74
2006	2	State Street Corp.	United States	94,356.45	19.23
2006	3	INVESCO Ltd.	United States	26,077.16	5.32
2007	1	Barclays Plc	United Kingdom	340,059.47	50.71
2007	2	State Street Corp.	United States	149,426.53	22.28
2007	3	Vanguard Group, Inc.	United States	40,350.88	6.02
2008	1	Barclays Plc	United Kingdom	243,692.34	45.29
2008	2	State Street Corp.	United States	145,673.55	27.08
2008	3	Vanguard Group, Inc.	United States	40,609.81	7.55
2009	1	Barclays Plc	United Kingdom	354,751.15	46.68
2009	2	State Street Corp.	United States	165,888.96	21.83
2009	3	Vanguard Group, Inc.	United States	79,649.11	10.48

References

Amihud, Y. 2002. Illiquidity and stock returns: Cross-section and time-series effects. Journal of Financial Markets 5:31–56.

Avramov, D., T. Chordia, G. Jostova, and A. Philipov. 2009. Dispersion in analysts' earnings forecasts and credit rating. Journal of Financial Economics 91:83–101.

Avramov, D., and R. Wermers. 2006. Investing in mutual funds when returns are predictable. Journal of Financial Economics 81:339–77.

Bhattacharya, U., B. Loos, S. Meyer, and A. Hackethal. 2017. Abusing ETFs. Review of Finance 21:1217–50.

Ben-David, I., F. Franzoni, and R. Moussawi. 2018. Do ETFs increase volatility? Journal of Finance. Advance Access published September 22, 2018, 10.1111/jofi.12727.

Blocher, J., and R. E. Whaley. 2016. Two-sided markets in asset management: exchange-traded funds and securities lending. Working Paper.

Boehmer, B., and E. Boehmer. 2003. Trading your neighbor's ETFs: Competition or fragmentation? Journal of Banking and Finance 27:1667–703.

Bollen, N., and J. Busse. 2005. Short-term persistence in mutual fund performance. Review of Financial Studies 18:569–97.

Carhart, M. M. 1997. On persistence in mutual fund performance. Journal of Finance 52:57–82.

Cremers, K. M., M. Ferreira, P. Matos, and L. Starks. 2016. Indexing and active fund management: international evidence. Journal of Financial Economics 120:539–60.

Cremers, K. M., and A. Petajisto. 2009. How active is your fund manager? A new measure that predicts performance. Review of Financial Studies 22:3329–65.

Da, Z., and S. Shive. 2018. Exchange traded funds and asset return correlations. European Financial Management 24:136–68.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers. 1997. Measuring mutual fund performance with characteristic-based benchmarks. Journal of Finance 52:1035–58.

Davies, A. 2012. BlackRock calls for action on conflicts of interest in ETFs. Reuters, March 30. https://www.reuters.com/article/financial-etfs-regulation/blackrock-calls-for-action-on-conflicts-of-interest-in-etfs-idUSL6E8EU2YJ20120330

Fama, E. F., and K. R. French. 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33:3–56.

Fama, E. F., and J. D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. Journal of Political Economy 71:607–36.

Ferreira, M. A., and P. Matos. 2008. The colors of investors' money: The role of institutional investors around the world. Journal of Financial Economics 88:499–533.

Financial Stability Board (FSB). 2011. Potential financial stability issues arising from recent trends in exchange-traded funds (ETFs). Consultation Document, April 11. Basel, Switzerland: Financial Stability Board. http://www.fsb.org/wp-content/uploads/r_110412b.pdf

Flood, C. 2012. UK regulator declares ETF concerns. Financial Times, April 1.

Gaspar, J. M., M. Massa, and P. Matos. 2006. Favoritism in mutual fund families? Evidence on strategic cross-fund subsidization. Journal of Finance 61:73–104.

Griffin, J. 2002. Are the Fama and French factors global or country-specific? Review of Financial Studies 15:783–803.

International Monetary Fund (IMF). 2011. Global Financial Stability Report. April 13. Washington, DC: International Monetary Fund. https://www.imf.org/en/Publications/GFSR/Issues/2016/12/31/Durable-Financial-Stability-Getting-There-from-Here

Irvine, P., M. Lipson, and A. Puckett. 2007. Tipping. Review of Financial Studies 20:741–68.

Israeli, D., C. M.C. Lee, and S. A. Sridharan. 2017. Is there a dark side to exchange traded funds? An information perspective. Review of Accounting Studies 22:1048–83.

Kacperczyk, M., C. Sialm, and L. Zheng. 2008. Unobserved actions of mutual funds. Review of Financial Studies 21:2379–416.

Mamaysky, H., and M. Spiegel. 2001. A theory of mutual funds: Optimal fund objectives and industry organization. Working Paper.

Mamaysky, H., M. Spiegel, and H. Zhang. 2007. Improved forecasting of mutual fund alphas and betas. Review of Finance 11:359–400.

———. 2008. Estimating the dynamics of mutual fund alphas and betas. Review of Financial Studies 21:233–64.

Massa, M. 2003. How do family strategies affect fund performance? When performance-maximization is not the only game in town. Journal of Financial Economics 67:249–304.

Massa, M., and Z. Rehman. 2008. Information flows within financial conglomerates: Evidence from the banks-mutual fund relationship. Journal of Financial Economics 89:288–306.

Massa, M., and L. Zhang. 2012. Fire-sales and information advantage: When bank-affiliation helps. Working Paper.

Morningstar. 2012. Synthetic ETFs under the microscope: A global study. Morningstar report.

Nanda, V., Z. J. Wang, and L. Zheng. 2004. Family values and the star phenomenon: strategies of mutual-fund families. Review of Financial Studies 17:667–98.

Newey, W. K., and K. D. West. 1987. A simple positive-definite heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55:703–8.

Ramaswamy, S. 2011. Market structures and systemic risks of exchange-traded funds. Working Paper.

Slater, S. 2012. Barclays to sell $6.1 billion BlackRock stake. Reuters, May 21. https://www.reuters.com/article/us-barclays-blackrock/barclays-to-sell-6-billion-blackrock-stake-idUSBRE84K0FT20120521