

Does Search Engine Visibility Help ETFs Attract Flows?

Olga A. Obizhaeva

Central University

Swedish House of Finance

16th research seminar, November 2025

Motivation

- Competition among exchange-traded funds (ETFs) led to a noticeable reduction in their fees (Investment Company Institute (2024)).
- When price and quality competitions are limited, other forms of competition may intensify (Scitovsky(1950)).
- Online space is important:
 - Consumers search for products online, for example, using Google.
 - Companies market and sell their products using online channels.

My main question: How online activities affect capital formation process in the asset management industry?

Main steps

1. Propose a novel methodology to calculate a set of descriptive metrics of search engine visibility using the web analytics data.
2. Apply the methodology to the U.S. ETFs.
3. Design empirical tests to deal with various econometrics issues (endogeneity, unobservability of consumers' characteristics, etc)
4. Study whether better visibility on the search engine helps ETFs to attract more capital?

Three datasets merged

I download and scrape data for the period of September 2019 through February 2021 from three sources:

1. ETFDB database.
 - ETFs' webpages URLs.
2. Bloomberg.
 - ETFs market data: performance, NAV, shares outstanding, contract.
3. SEMRush Web Analytics (main dataset).
 - Web analytics, 20B+ organic keywords, 1B+ ads, backlinks.

SEMRush web analytics data

SEMRush collects historical data on Google's Search Engine Reporting Pages (SERP).

- Provides database of keywords that exceeds 20 billion keywords and covers 142 geographical regions.
- Scrapes top 100 URLs on the Google SERP for each keyword in its database on a monthly basis.
- Records information on paid and organic search results,
 $\Omega_{i,t} = \{\Omega_{i,t}^{paid}, \Omega_{i,t}^{org}\}$, for which the webpage i appears on the Google SERP in month t .

Web analytics variables

For every keyword $\omega \in \Omega_{i,t}$ of the webpage i in month t :

- $\text{Rank}_{i,t}(\omega)$ is the rank of the webpage i on Google's SERP for the keyword ω ;
- $\text{PPC}_t(\omega)$, is the Google ads auction-determined price-per-click (PPC) for the keyword ω ;
- $\text{SVol}_t(\omega)$ is the estimated search volume for the keyword ω ;
- $\#\text{Clicks}_{i,t}(\omega)$ is the estimated number of clicks on the webpage i from the keyword search ω .

Illustrative example

ETF i has webpage URL “ETF URL”.

 1. Ad URL 1 2. Ad URL 2 3. URL 3 4. URL 4 5. URL 5 = ETF URL	 1. URL 1 2. URL 2 = ETF URL 3. URL 3 4. URL 4 5. URL 5	 1. Ad URL 1 = ETF URL 2. URL 2 3. URL 3 4. URL 4 5. URL 5
---	---	--

Organic keywords

$$\Omega_{i,t}^{org} = \{\text{Keyword 1, Keyword 2}\}$$

$$\text{Rank}_{i,t}(\text{Keyword 1}) = 5$$

$$\text{Rank}_{i,t}(\text{Keyword 2}) = 2$$

Paid keywords

$$\Omega_{i,t}^{paid} = \{\text{Keyword 3}\}$$

$$\text{Rank}_{i,t}(\text{Keyword 3}) = 1$$



Subsamples of keywords

- No ad auctions ($P_{PC} = 0$)
- Buy intentions ($\Omega^1 = \{\text{buy, invest}\}$)
- Sell intentions ($\Omega^2 = \{\text{sell}\}$)
- Fund types ($\Omega^3 = \{\text{index funds, mutual funds, ...}\}$)
- Asset classes and popular indices ($\Omega^4 = \{\text{stocks, bonds, ...}\}$)
- Names of ETF issuers ($\Omega^5 = \{\text{Blackrock, ishares, ...}\}$)

Search engine marketing

I construct the proxy for whether ETF i uses the search engine marketing (SEM) in month t :

$$\text{SEM}_{i,t} = \begin{cases} 1, & \text{if ETF engages in SEM} \\ 0, & \text{otherwise.} \end{cases}$$

Search engine marketing cost

I use the price-per-click (PPC) and the estimated number of clicks on the webpage across all paid keywords $\Omega_{i,t}^{paid}$ of the ETF i in month t to calculate the fund's expenses on search engine marketing:

$$\text{SEMCost}_{i,t} = \begin{cases} \sum_{\omega \in \Omega_{i,t}^{paid}} \text{PPC}_t(\omega) \cdot \#\text{Clicks}_{i,t}(\omega), & \text{if ETF engages in SEM} \\ 0, & \text{otherwise.} \end{cases}$$

Effects of Search Engine Visibility

The expected benefits for a fund from having its webpage on the search engine results page (SERP) depend on the two components:

$$\left[\text{Number of Clicks} \right] \times \left[\text{Value of Click} \right].$$

Since a fund appears at results pages for many queries, expected benefits have to be aggregated across all keywords.

Search engine visibility: Average SERP rank

I construct a proxy for the number of clicks as the average SERP rank by aggregating the rank of the ETF i 's webpage in month t across all keywords with weights proportional to keyword's overall search volume:

$$\overline{\text{Rank}}_{i,t} = \frac{\sum_{\omega \in \Omega_{i,t}} \text{SVol}_t(\omega) \cdot \text{Rank}_{i,t}(\omega)}{\sum_{\omega \in \Omega_{i,t}} \text{SVol}_t(\omega)}.$$

The webpages with lower average rank have better online visibility and higher chances to get clicks.

Search engine visibility: Average page

Google (usually) displays 10 URLs on each page of SERP. To capture potential non-linearity in relationships, I construct the dummies:

$$\text{Page}_{i,t}^1 = \mathbb{1}[1 \leq \overline{\text{Rank}}_{i,t} \leq 10];$$

$$\text{Page}_{i,t}^2 = \mathbb{1}[10 < \overline{\text{Rank}}_{i,t} \leq 20];$$

$$\text{Page}_{i,t}^3 = \mathbb{1}[20 < \overline{\text{Rank}}_{i,t} \leq 30];$$

$$\text{Page}_{i,t}^4 = \mathbb{1}[30 < \overline{\text{Rank}}_{i,t} \leq 40].$$

$$\text{Page}_{i,t}^5 = \mathbb{1}[40 < \overline{\text{Rank}}_{i,t} \leq 50].$$

Search engine visibility: Value of click

I construct a proxy for the value of a click as the average value of a click by aggregating $PPC_{i,t}(\omega)$ of the ETF i 's webpage in month t over all keywords with weights proportional to keyword's overall search volume:

$$\overline{PPC}_{i,t} = \frac{\sum_{\omega \in \Omega_{i,t}} SVol_t(\omega) \cdot PPC_{i,t}(\omega)}{\sum_{\omega \in \Omega_{i,t}} SVol_t(\omega)}.$$

A higher average PPC signals that advertisers especially value traffic from pages where a fund is shown among search engine results.

Bloomberg data

Sample of 1,784 US ETFs from September 2019 to May 2021.

For ETF i in month t :

- Benchmark index;
- Asset class;
- Net asset value, $P_{i,t}$;
- Shares outstanding, $\text{ShrOut}_{i,t}$;
- Expense ratio, Fee_i .

$$\% \text{Flow}_{i,t} = \frac{(\text{ShrOut}_{i,t} - \text{ShrOut}_{i,t-1}) \cdot P_{i,t}}{\text{ShrOut}_{i,t-1} \cdot P_{i,t-1}}.$$

Summary Statistics

Main Fund Flows Regression

I estimate the relationship between ETF's fund flows and measures of online visibility using the following panel regressions with controls and fixed effects:

$$\begin{aligned}
 \text{Flow}_{i,t+1} = & \alpha_0 + \alpha_{rank} \cdot \overline{\text{Rank}}_{i,t} + \alpha_{PPC} \cdot \ln(\overline{\text{PPC}}_{it}) + \alpha_{link} \cdot \ln(\# \text{Links}_{it}) + \\
 & + \sum_{n=1}^4 \alpha_R^n \cdot R_{it}^n + \alpha_{flow} \cdot \text{Flow}_{i,t} + \alpha_{size} \cdot \ln(\text{Size}_{i,t}) + \\
 & + \alpha_{age} \cdot \ln(\text{Age}_{i,t}) + \alpha_{fee} \cdot \text{Fee}_i + \alpha_{\%ba} \cdot \% \text{BASpread}_{it} + \\
 & + \lambda_{asset} + \lambda_{issuer} + \lambda_t + \epsilon_{i,t+1}.
 \end{aligned}$$

Problem: Ranking endogeneity

In regressions of fund flows on rankings, SERP rankings are endogenous:

- Google employs sophisticated algorithms and use information about users' choices for determining positions.
- Products with higher likelihood of being clicked are likely to be positioned at the top of SERP.

The previous literature suggested a number of solutions: Narayanan and Kalyanam (2015), Ghose et.al.(2009), Ghose et.al.(2014), Rutz et.al.(2012), Blake et.al(2015), and Ursu(2018).

My approach

- Exclude sponsored listings and focus on organic search results.
 - Users search for various purposes, so the correlation should be weaker.
 - Janssen et al. (2023) provide theoretical arguments for why it might be optimal for Google to randomly assign organic search positions.
- Control for variables, such as the number of backlinks, that are known to affect both ranks and flows. Control for time, issuer, asset class fixed effects.
- Perform analysis on subsamples of keywords to better control for customers' characteristics.
- Check results for ETFs tracking the same index.

Fund Flows and SERP Positions: Results

I obtain the following results:

$$\begin{aligned} \%Flow_{i,t+1} = & \underset{(0.012)}{0.051^{***}} \cdot Page_{it}^1 + \underset{(0.009)}{0.045^{***}} \cdot Page_{it}^2 + \underset{(0.008)}{0.028^{***}} \cdot Page_{it}^3 + \\ & + \underset{(0.008)}{0.021^{**}} \cdot Page_{it}^4 + \underset{(0.008)}{0.007} \cdot Page_{it}^5 + \underset{(0.002)}{0.013^{***}} \cdot \ln(\overline{PPC}) + \\ & + Controls_{it} + Fixed\ Effects + \epsilon_{i,t+1}. \end{aligned}$$

These results are robust to various specifications (only organic keywords, subsamples with ad auctions and w/o them, subsamples that exclude Ω_1 , Ω_2 , Ω_3 , Ω_4 , and Ω_5).

Results: Sample of ETFs with the Same Benchmark



	(1)	(2)	(3)	(4)	(5)
$\overline{\text{Rank}}$	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001** (0.000)	-0.000 (0.000)
$\ln(\overline{\text{PPC}})$	0.022*** (0.006)	0.014** (0.007)	—	0.016** (0.006)	0.008*** (0.003)
$\ln(\# \text{Links})$	0.003 (0.003)	0.007** (0.003)	0.008** (0.003)	0.007** (0.003)	0.007* (0.004)
Controls	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Issuer FE	Yes	Yes	Yes	Yes	Yes
Index FE	Yes	Yes	Yes	Yes	Yes
Sample	Ω	Ω^{org}	$\Omega^{org} PPC = 0$	$\Omega^{org} \setminus \Omega^3$	$\Omega^{org} \setminus \bigcup_{i=1}^5 \Omega^i$
#Obs.	1,798	1,798	1,798	1,798	1,798
Adj. R^2	0.130	0.213	0.212	0.213	0.213

Results are similar even for the smaller sample of funds tracking the same indices with index FE.

SEM Results



	(1)	(2)	(3)
SEM	0.020* (0.010)	—	0.016 (0.023)
ln(SEM Expense)	—	0.002 (0.002)	—
SEM * Top5%	—	—	0.005 (0.024)
Controls	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Issuer FE	Yes	Yes	Yes
Asset class FE	Yes	Yes	Yes
Sample	Ω	Ω	Ω
#Obs.	21,219	21,219	21,219
Adj. R^2	0.122	0.122	0.122

ETF's fund flows: Variance decomposition

I do variance decomposition to analyze how important are measures of online activities for explaining fund flows:

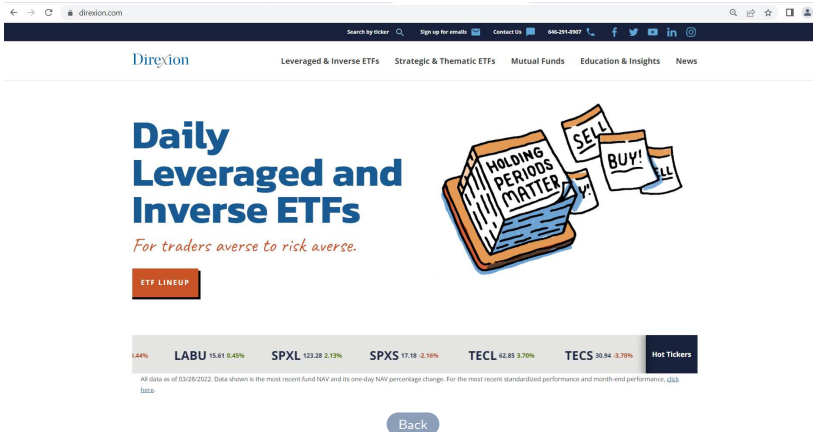
Fixed Effects	Age	Past returns	Size	Google	BA spread	SEM	Fees	Past flows
45.6%	23.0%	16.3%	11.5%	1.9%	0.8%	0.5%	0.2%	0.2%

In addition to conventional factors, the fund's visibility on Google explains 1.9% of fund flows; this is more than funds' fees, past flows, and liquidity. SEM explains 0.5%.

Conclusion

1. Financial companies do use online marketing tools to gain advantage in attracting funds.
2. Search engine visibility is important for an ETF's capital formation:
 - ETFs that engage in search engine marketing(SEM) attract more flows in the next quarter.
 - ETFs with web pages that, on average, position higher on Google get higher future fund flows.
 - The quality of traffic proxied by the average pay-per-click is important.
3. The online visibility and SEM of an ETF explains an additional 2.5% of the variation in ETF fund flows (besides age, past performance, and size of the fund).

Paid search: Direxion homepage



The screenshot shows the Direxion homepage with a dark blue header containing navigation links and social media icons. The main content area features a large blue headline, a sub-headline in red, an orange button, and a table of ETFs. An illustration of a book and trading cards is also present.

Daily Leveraged and Inverse ETFs

For traders averse to risk averse.

ETF LINEUP

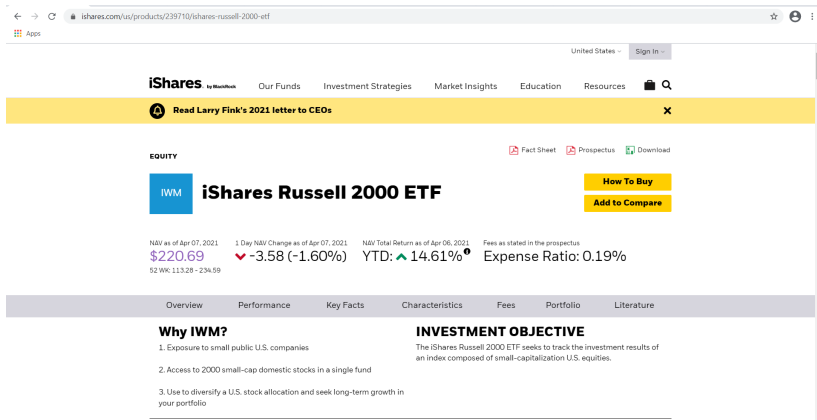
ETF	NAV	1-Day % Change
LABU	15.61	0.45%
SPXL	123.28	2.13%
SPXS	17.18	-2.16%
TECL	62.85	3.70%
TECS	30.94	-3.70%

Hot Tickers

All data as of 03/28/2022. Data shown is the most recent fund NAV and its one-day NAV percentage change. For the most recent standardized performance and month-end performance, [click here](#).

[Back](#)

Organic search: iShare IWM webpage



The screenshot shows the iShares website for the iShares Russell 2000 ETF (IWM). The page includes a navigation bar with links to 'Our Funds', 'Investment Strategies', 'Market Insights', 'Education', 'Resources', and a search icon. A yellow banner at the top promotes 'Read Larry Fink's 2021 letter to CEOs'. Below this, the 'EQUITY' section features the IWM logo and the title 'iShares Russell 2000 ETF'. To the right of the title are buttons for 'How To Buy' and 'Add to Compare'. The main content area displays key performance indicators: NAV as of Apr 07, 2021 at \$220.69, 1 Day NAV Change as of Apr 07, 2021 at -3.58 (-1.60%), YTD Total Return as of Apr 06, 2021 at 14.61%, and an Expense Ratio of 0.19%. A horizontal menu below the main content area includes links to 'Overview', 'Performance', 'Key Facts', 'Characteristics', 'Fees', 'Portfolio', and 'Literature'. The 'Why IWM?' section lists three points: 1. Exposure to small public U.S. companies, 2. Access to 2000 small-cap domestic stocks in a single fund, and 3. Use to diversify a U.S. stock allocation and seek long-term growth in your portfolio. The 'INVESTMENT OBJECTIVE' section states that the iShares Russell 2000 ETF seeks to track the investment results of an index composed of small-capitalization U.S. equities.

← → ↻ ishares.com/us/products/239710/ishares-russell-2000-etf ☆ ⓘ ⋮

Apps

United States - Sign In -

iShares by BlackRock Our Funds Investment Strategies Market Insights Education Resources 🔍

Read Larry Fink's 2021 letter to CEOs ✕

EQUITY [Fact Sheet](#) [Prospectus](#) [Download](#)

IWM **iShares Russell 2000 ETF** [How To Buy](#) [Add to Compare](#)

NAV as of Apr 07, 2021 **\$220.69** 1 Day NAV Change as of Apr 07, 2021 ▼ -3.58 (-1.60%) YTD Total Return as of Apr 06, 2021 ▲ 14.61%^o Fees as stated in the prospectus Expense Ratio: 0.19%

52 WK: 113.28 - 234.59

Overview Performance Key Facts Characteristics Fees Portfolio Literature

Why IWM?

1. Exposure to small public U.S. companies
2. Access to 2000 small-cap domestic stocks in a single fund
3. Use to diversify a U.S. stock allocation and seek long-term growth in your portfolio

INVESTMENT OBJECTIVE

The iShares Russell 2000 ETF seeks to track the investment results of an index composed of small-capitalization U.S. equities.

Back

Summary stats: All keywords

	#Keywords (millions)	PPC=0 (%)	PPC(\$) avg	PPC > 0 med
Set Ω	1.674	64%	5.17	3.82
Set with auctions	0.609	0%	5.17	3.82
Set w/o auctions	1.065	100%	0.00	0.00

[Back](#)

Bloomberg: Summary stats

	AuM (\$ (billions)	Volume (\$ (billions)	Return	Fee (bps)	%BA spread (bps)	Age (years)
All	3.12	0.80	0.002	44	25	8.1
Page 1	11.62	1.49	-0.020	21	9	10.6
Page 2	5.94	2.05	0.019	40	18	11.6
Page 3	3.31	1.30	0.004	43	23	9.2
Page 4	1.97	0.45	0.001	43	27	7.4
Page 5	1.82	0.33	-0.006	46	29	6.5
Pages Below	0.86	0.16	0.005	58	32	6.4

[Back](#)