



## Management Science

Publication details, including instructions for authors and subscription information:  
<http://pubsonline.informs.org>

### Exchange-Traded Funds and the Wash Sale Loophole

Michael Dambra, Andrew Glover, Charles M. C. Lee, Phillip J. Quinn

To cite this article:

Michael Dambra, Andrew Glover, Charles M. C. Lee, Phillip J. Quinn (2026) Exchange-Traded Funds and the Wash Sale Loophole. Management Science

Published online in Articles in Advance 03 Apr 2026

. <https://doi.org/10.1287/mnsc.2024.06944>

This work is licensed under a Creative Commons Attribution 4.0 International License. You are free to copy, distribute, transmit and adapt this work, but you must attribute this work as “*Management Science*. Copyright © 2026 The Author(s). <https://doi.org/10.1287/mnsc.2024.06944>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by/4.0/>.”

Copyright © 2026 The Author(s)

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

# Exchange-Traded Funds and the Wash Sale Loophole

Michael Dambra,<sup>a</sup> Andrew Glover,<sup>b</sup> Charles M. C. Lee,<sup>b,\*</sup> Phillip J. Quinn<sup>b</sup>

<sup>a</sup>School of Management, University at Buffalo, Buffalo, New York 14260; <sup>b</sup>Foster School of Business, University of Washington, Seattle, Washington 98195

\*Corresponding author

Contact: [mjdambra@buffalo.edu](mailto:mjdambra@buffalo.edu),  <https://orcid.org/0000-0001-5198-6197> (MD); [glov@uw.edu](mailto:glov@uw.edu),  <https://orcid.org/0009-0001-9951-0179> (AG); [clee27@uw.edu](mailto:clee27@uw.edu),  <https://orcid.org/0000-0002-9978-6044> (CMCL); [philq@uw.edu](mailto:philq@uw.edu),  <https://orcid.org/0000-0001-7922-0708> (PJQ)

Received: July 20, 2024  
 Revised: February 6, 2025; June 17, 2025  
 Accepted: August 11, 2025  
 Published Online in *Articles in Advance*:  
 April 3, 2026

<https://doi.org/10.1287/mnsc.2024.06944>

Copyright: © 2026 The Author(s)

**Abstract.** Tax wash sale rules prohibit the recognition of capital losses when substantially identical securities are sold and immediately repurchased within short windows. This study examines whether institutional investors use exchange-traded funds (ETFs) to circumvent wash sale rules. Consistent with tax-motivated demand for ETFs, incumbent ETFs both create more shares and experience more trading volume upon the introduction of nearly identical ETFs, particularly when recent returns are negative. We show that tax-sensitive institutions' investment in highly correlated ETFs has proliferated in recent years, exceeding a quarter of their assets under management. Furthermore, tax-sensitive institutions holding more ETFs are significantly more likely to engage in swapping nearly identical ETFs. This swapping behavior has become widespread, with tax-sensitive institutional investors swapping \$417 billion of nearly identical ETFs since 2001. We estimate that tax-sensitive institutions realized more than \$84 billion dollars in losses in highly correlated ETFs associated with the swapping activity since 2001.

**History:** Accepted by Shiva Rajgopal, accounting.



**Open Access Statement:** This work is licensed under a Creative Commons Attribution 4.0 International License. You are free to copy, distribute, transmit and adapt this work, but you must attribute this work as "*Management Science*. Copyright © 2026 The Author(s). <https://doi.org/10.1287/mnsc.2024.06944>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by/4.0/>."

**Funding:** The authors thank their institutions for financial support. M. Dambra thanks the Kenneth W. Colwell endowment for research support, C. M. C. Lee thanks the Kermit O. Hanson endowment for research support, and P. J. Quinn thanks PricewaterhouseCoopers for research support.

**Supplemental Material:** The online appendix and data files are available at <https://doi.org/10.1287/mnsc.2024.06944>.

**Keywords:** exchange-traded funds • wash sale rule • institutional investors • personal taxes • Internal Revenue Service

Tax loss harvesting is the crown jewel of tax management strategies, and it represents what we consider to be the only reliable way for investors to outperform the market ... (Malkiel 2023)

## 1. Introduction

For over 100 years, U.S. law has prohibited claiming a tax loss on a wash sale—where a security is sold for a realized loss (RL) within 30 days of purchasing a “substantially identical” security (26 CFR section 1091-1). By disallowing tax losses from the sale and repurchase of substantially identical securities within a short window, the wash sale rule prevents harvesting capital losses with trades that lack economic substance. Although theory argues that investors should immediately harvest capital losses for tax benefits (Constantinides 1983), the wash sale rule significantly limits investors' ability to do so (Grinblatt and Keloharju 2004, Jensen and Marekwicka 2011).

Exchange-traded funds (ETFs) provide an efficient way for investors to circumvent the trading frictions associated with the wash sale rule. Specifically, investors can sell a depreciated ETF security and realize a capital loss while simultaneously purchasing another nearly identical ETF security. This form of swap trading allows investors to maintain a substantially identical economic position while harvesting a capital loss that can be used to offset realized gains (RG) and other taxable income.<sup>1</sup> With an explosion in available ETFs over the past two decades, these securities have become ideal vehicles for circumventing the wash sale rule. More than 2,000 ETFs trade today on U.S. exchanges (Lundholm and Zheng 2025), and many track economically equivalent baskets of securities. These broad-based ETFs typically offer greater liquidity and lower fees than traditional diversified investment vehicles, such as mutual funds (Huang and Guedj 2009, Ben-David et al. 2023). More importantly for our purposes,

investors in ETFs can buy and sell their shares directly throughout the day; they do not need to purchase (or redeem) their shares from the mutual fund at end-of-day prices.

Our study examines the extent to which investors swap shares of highly correlated (HC) ETFs to circumvent the wash sale rule. Although the economic intent of the wash sale rule is straightforward, significant uncertainty remains as to the permissibility of tax deductions achieved through ETF swaps. The Internal Revenue Service (IRS) has not ruled on what constitutes a “substantially identical” security, leaving financial advisors to navigate a foggy legal landscape. Some advisors seem to take the regulatory silence as tacit permission to swap ETFs that hold identical securities or that are benchmarked to the same index (e.g., Lasser 2011). Others argue that if an investor’s economic position has not changed after swapping ETFs, the spirit of the wash sale rule has likely been violated (e.g., Fischer 2010). Against this backdrop of legal uncertainty, the extent to which investors engage in tax avoidance through ETF wash sales remains largely unknown.

Our analyses proceed in three stages, each employing a different database. In the first stage, we use ETF-level data to examine primary and secondary market demand for incumbent ETFs when a near-identical ETF alternative is introduced. Absent tax effects, economic theory predicts that the introduction of a near-identical alternative security will reduce investor demand for the incumbent ETF as the two funds are economic substitutes. However, if the second ETF provides a wash sale benefit to investors, the two ETFs may serve as complementary goods, leading to increased demand for both assets. We explore this possibility by examining the change in demand for incumbent ETF shares around the introduction of a new near-identical ETF.

Using index ETF data from 1993 through 2022, we identify ETFs with highly correlated returns (i.e.,  $\geq 99\%$ ) as potential loss-harvesting peers.<sup>2</sup> Our focus is on the change in demand for an incumbent ETF at the introduction of a near-identical peer ETF. We approximate primary market demand with ETF shares outstanding (e.g., Brown et al. 2021) and secondary market demand with ETF trading volume (e.g., Ben-David et al. 2018, Li 2024). We then conduct a difference-in-differences (DiD) analysis to determine whether the introduction of a highly correlated ETF increases or decreases demand for the incumbent ETF. To avoid potential heterogeneity in treatment effects that could confound a staggered DiD design, we align the introductions of loss harvest pairs in event time and run a stacked regression that compares the treated ETF with a clean control group of unpaired ETFs (Baker et al. 2022).

Our results show that demand for incumbent ETFs’ shares increase with the introduction of a new near-

identical ETF. First, consistent with an increase in primary market demand, we find a significant increase in the incumbent ETF’s total shares outstanding (indicating net new issuance of ETF shares). Second, consistent with an increase in secondary market demand, we find a significant increase in incumbent ETF trading volume. Our trading volume results confirm those in Li (2024), which provides the first evidence of increases in incumbent ETF trading volume following new ETF introductions using a staggered DiD and a Callaway and Sant’Anna (2021)-style DiD. The findings in our study and Li (2024) broadly support a complementarity effect between pairs of nearly identical ETFs.

One plausible alternative explanation for the increase in shares outstanding and trading volume after ETF introductions is that these introductions coincide with a surge in investor demand for low-cost, diversified portfolios (i.e., Agapova 2011). To the extent that ETF competitors introduce near-identical products when a comparable incumbent product has been especially successful, the increased trading volume may not be because of tax loss harvesting or wash sale avoidance. Instead, the increased activity may simply reflect exceptionally strong growth in the underlying demand for an incumbent ETF product. Furthermore, contemporaneous research has identified other tax advantages to *holding* ETFs, including their ability to avoid distributing capital gains (Moussawi et al. 2025). To discriminate between these alternative explanations, we analyze institutional investor activities at the micro level.<sup>3</sup>

In the second stage of our analysis, we use granular trading data from Abel Noser (formerly Ancerno) to document the existence of institutional investor same-day, same-fund swapping of highly correlated ETFs. This database contains all trades made by Abel Noser’s sizeable client base from January 2001 to December 2010. Prior research (e.g., Puckett and Yan 2011, Hu et al. 2018) shows that the characteristics of stocks owned and traded by Abel Noser’s clients do not significantly differ from the characteristics of stocks owned and traded by the average 13F-filing institution.

Our evidence shows that even during the early days of ETF availability (i.e., 2001–2010), institutional investors engaged in direct, same-day swap trades using highly correlated ETFs. Focusing on highly correlated ETFs traded by the same fund in opposite directions on the same trading day, we find that the Abel Noser database contains many such trades. Over our sample, the frequency of near-identical, same-day ETF swaps increases from a single trade in 2001 to more than 1,400 trades in 2010. The existence of these highly correlated ETF swap trades supports the wash sale hypothesis and corroborates the macro-level empirical evidence. Institutional investors incurred trading costs to swap between ETFs with return correlations of 99% or greater, with funds often tracking identical indices. We

are unaware of any economic rationale other than tax loss harvesting that could explain this phenomenon.

Although the highly granular Abel Noser data are useful in establishing the *existence* of wash sale transactions among institutional trades, the limited coverage of these data (in terms of sample, institutional coverage, the lack of holdings data, and our inability to track investors over a time series) prevents us from linking ETF swaps directly to funds' loss recognition activities. To establish this link more directly and to provide an estimate of the scope of tax avoidance through such activities, we turn to institutional investors' quarterly filings.

In the third stage of our analysis, we use Form ADV (Uniform Application for Investment Adviser Registration and Report by Exempt Reporting Advisers) and WhaleWisdom 13F data to examine the extent to which institutional investors use ETF wash sales for tax avoidance. We classify institutional investors as tax sensitive if they report non-zero high-net-worth (HNW) clients on Form ADV. Tax-insensitive institutions are identified as either those reporting zero high-net-worth clients on Form ADV or those identified as a pension plan or an endowment via Brian Bushee's Institutional Investor Classification Database.<sup>4</sup> The institutional investors that we cannot classify using these techniques are excluded from the sample.<sup>5</sup> WhaleWisdom contains every position that each 13F filer reports to the Securities and Exchange Commission (SEC). We interact 13F filers' tax sensitivity with measures of highly correlated ETF availability in ordinary least squares (OLS) regressions with filer fixed effects. Our results show that tax-sensitive institutions hold a greater number of highly correlated ETFs and have a significantly larger portion of their assets under management (AUM) in these instruments. Furthermore, tax-sensitive institutions' investment in highly correlated ETFs has proliferated in recent years, exceeding 25% of AUM. These findings are broadly consistent with highly correlated ETFs offering significant tax benefits to institutional investors.

To further understand whether the increased ETF holdings are driven by incentives to avoid the wash sale rule, we examine the extent to which investors swap highly correlated ETFs. We proxy for institutional trades using quarterly changes in 13F holdings. ETF swaps are identified as the offsetting quarterly buys and sells in separate ETFs with prices correlated at  $\geq 99\%$ . We find that tax-sensitive institutions engage in substantial swapping. These buys and sells of highly correlated ETFs seem to lack economic substance beyond harvesting capital losses. Since 2001, tax-sensitive institutional investors in our sample have swapped approximately \$417 billion between highly correlated ETFs and \$106 billion in 2022 alone.

Finally, to determine whether ETF swapping is indeed associated with the realization of capital losses, we follow Sikes (2014) and Blouin et al. (2017) to

estimate institutional investors' basis in each holding and calculate their realized losses, unrealized losses (UL), and the proportion of losses realized (*PLR*) in a given quarter (Odean 1998). Our results show a strong positive association between highly correlated ETF swaps and the *PLR*, with the tax-sensitive institutions' trading activity being the main driver. On average, a tax-sensitive institution that incurs a standard deviation increase in swapping activity recognizes 1.2% more of their available losses each quarter, a significant increase considering that the median quarterly *PLR* in our tax-sensitive institutional sample is 17.3%.

Our estimate of the extent of incremental capital loss recognition attributable to ETF swaps is affected by at least four sources of potential noise. First, to the extent that *different funds* within the same filing entity buy and sell offsetting amounts of highly correlated ETFs by happenstance, our estimate of swapping activity will be noisy and may be biased upward or downward. Second, our measure of quarterly changes in holdings will not capture intraquarter swapping of highly correlated ETFs. This source of noise will lead to a downward bias in our estimate of swapping. Third, to identify near-identical ETFs, we impose a high threshold for their correlation in returns ( $\geq 99\%$ ). If institutional investors use a less stringent threshold, our estimate of tax-related swaps will again be biased downward.<sup>6</sup> Finally, given the increase in ETF swapping observed throughout our sample period, the averages that we compute are likely to understate future swapping activities significantly. Considering all four sources of noise, we believe that our estimate of the capital loss recognition attributable to ETF swaps is likely a lower bound on the institutional investors' tax avoidance.

Our study contributes to several lines of existing research. First, we add to the literature on how institutional investors incorporate tax strategies into their active portfolio management (Sialm and Starks 2012, Blouin et al. 2017). We are unaware of other studies that examine how institutional investors exploit ETFs to harvest capital losses while maintaining a nearly identical portfolio position. Empirical evidence also documents that investors realize losses too late and inefficiently (Odean 1998, Barber and Odean 2004). Our evidence suggests that institutional investors utilize ETFs to circumvent trading frictions imposed by the wash sale rule to capture available capital losses.

Second, we add to the literature on the increasing importance of ETFs in U.S. capital markets. In the past 20 years, the global ETF market has exploded from around 300 ETFs representing \$200 billion in 2003 to more than 8,700 ETFs representing \$9.5 trillion in 2022 (Statista 2023). Prior research emphasizes the role that ETFs play in passive investment (see Ben-David et al. 2017 and Liebi 2020 for recent reviews) and financial analysis (Lundholm and Zheng 2025). Concurrent

work by Moussawi et al. (2025) illustrates the favorable tax treatment of *holding* ETFs rather than mutual funds. Whereas Moussawi et al. (2025) sheds light on how ETF managers avoid *ETF-level* taxes in the primary market, our work shows how institutional investors avoid *investor-level* taxes by trading near-identical ETFs in the secondary market. In particular, we present novel evidence that the accumulation and aggressive swapping of ETFs are concentrated among tax-sensitive investors.

Our work complements a contemporaneous working paper by Li (2024). Like Li (2024), we use ETF-level data to document an increase in trading volume of incumbent ETFs after the introduction of a highly correlated ETF. Both studies also find that the relationship is negatively associated with past returns. However, beyond these common findings, the two papers take different approaches to develop their inferences. Li (2024) develops an analytical model and uses this model to estimate the tax revenue loss and tax alpha from tax-loss harvesting. By “tuning model parameters to match the empirical estimation of the tax-loss trading volume,” Li (2024, p. 6) infers an annual tax revenue loss of “about 0.52% of AUM of highly correlated ETFs, equivalent to 25 billion USD.”

In contrast, we draw our inferences from multiple data sources. We first use ETF-level data to show that new near-identical ETF entries affect incumbent ETF demand in both the primary market and the secondary market. We then use two additional institutional trading databases, one at the trade level and one at the 13F filer level, to estimate the degree to which institutions engage in swapping highly correlated ETFs. In conducting these tests, our research design focuses on partitioning institutions into tax-sensitive and tax-insensitive types. This design allows us to (a) discriminate between higher trading volume because of increased demand for ETF products and higher volume because of tax-loss harvesting and (b) provide a direct estimate of the magnitude of the *tax-related* swapping activity. Overall, we estimate that 13F filers realized \$84 billion of losses from swapping highly correlated ETFs during our sample. The findings in Li (2024) and our study are complementary, and together, they provide robust evidence that highly correlated ETFs are increasingly used to harvest significant losses that would otherwise be disallowed by the wash sale rule.

Finally, we inform the policy debate around loss harvesting. Resource-constrained regulators consider success likelihood when deciding to challenge tax positions (Nessa et al. 2020), and the IRS weighs enforcement costs when deciding whether to clarify tax law (e.g., how to audit whether two ETFs are “too correlated” or “too similar in economic substance”). The IRS has not defined the term “substantially identical” in its wash sale provision, which fosters uncertainty about whether swapping highly correlated ETFs violates the wash sale rule among

legal scholars and practitioners (Fischer 2010, Ganel 2024). Our analysis and estimate of the magnitude of the trading activity seemingly driven by institutional investors’ loss harvesting give regulators a lower bound on ETF swapping activity as they weigh continued silence regarding asset similarity.

## 2. Wash Sale Background

An optimal investment strategy is to realize capital losses immediately and defer capital gains until the event of a forced liquidation (i.e., Constantinides 1983), all while maintaining the same risk/return exposures in the portfolio. However, the disallowance of losses generated via a wash sale creates a meaningful friction to this “realize and rebalance” strategy. The U.S. Federal Government has taken the position that although the goal of tax minimization is not inherently problematic, avoiding taxes with activities that otherwise lack economic purpose will not be permitted.

The U.S. Congress passed the Revenue Act of 1921 (the Act) in November 1921 (Blakey 1922). The Act was projected to “considerably increase the revenue by preventing taxpayers from taking [recognizable] losses in wash sales and other fictitious exchanges” and specifies that no deduction is permitted for losses on a sale “where it appears that within 30 days after such a sale the taxpayer purchases identical securities” (U.S. Senate 1921, p. 14). The congressional prohibition on wash sales coincided with the growing notoriety of a tax-loss-harvesting-via-wash-sale strategy suggested by the popular press (Wall Street Journal 1920). Enforcement of the law fell to the Treasury Department and the Internal Revenue Service. The IRS interprets wash sales via its Publication 550:

A wash sale occurs when you sell or trade stock or securities at a loss and within 30 days before or after the sale you: (1) buy substantially identical stock or securities, (2) acquire substantially identical stock or security in a fully taxable trade, (3) acquire a contract or option to buy substantially identical stock or securities, or (4) acquire substantially identical stock for your individual retirement arrangement or Roth IRA. (Internal Revenue Service 2021)

Thus, investors wishing to recognize a capital loss must sell the security in question and not purchase a “substantially identical” security within 30 days—before or after—the sale. This rule forces investors to choose between recognizing losses and maintaining an optimal portfolio (Jensen and Marekwica 2011). The IRS clarified in Revenue Ruling 2008-5 that investors cannot claim a tax loss on a substantially identical security that is purchased through an individual retirement account (IRA) within 30 days (Silow 2008) and has sanctioned various individuals and entities for violations of the wash sale rule (e.g., Internal Revenue Service 1983 (*Estate of Estroff*

*v. Commissioner*), 1985 (Rev. Ruling 85-87), 2017 (Technical Advice Memorandum 201737011)).

To effectively circumvent the wash sale rule, an asset must be readily exchangeable for a highly correlated but not identical asset. However, considerable ambiguity remains on what constitutes a “substantially identical” security (i.e., Fischer 2010). This ambiguity has led investors to consider alternative strategies to circumvent the trading frictions imposed by the wash sale rule. With the proliferation of near-identical ETFs, it has become extremely easy for investors to circumvent the wash sale rule by selling and immediately buying back highly correlated securities. But, what exactly does “substantially identical” mean in the context of ETFs?

To better understand the current guidance and rules of thumb available to investors, we searched both practitioner and academic sources. Our results, summarized in Figure 1 in the Online Appendix, reveal a wide range of interpretations. Some practitioners employ strict interpretations of what constitutes substantially identical and caution against using highly correlated ETFs to harvest tax losses (Goodman 2015, Matthews 2016, Holt 2023). These authors typically believe that two ETFs are substantially identical if their return correlation and/or holdings overlap exceed a certain threshold. For example, Matthews (2016, p. 12) proposes that “ETFs are substantially identical if they are reasonably expected to: (1) generate returns that mirror each other within a fraction of a percent; and (2) hold portfolios that overlap by 99% or more.”

Other sources adopt more lenient interpretations. One common view is that two ETFs are likely substantially identical if they track the same underlying index. For example, BlackRock (2024, p. 1) states that swapping “ETF or Index Mutual Fund with same index” has a higher “risk of wash sale classification.” Dong (2022) argues that substantially identical requires two ETFs to track the same index: “VOO tracks the Standard and Poor’s (S&P) 500 Index, while VTI tracks the Center for Research in Security Prices (CRSP) U.S. Total Market Index. This makes them sufficiently differentiated in the eyes of the IRS, even though VTI has a 0.99 correlation to VOO.”<sup>7</sup> Finally, some authors, such as Mezziani and Yang (2012), argue that even tracking the same index does not necessarily make two ETFs substantially identical if they are issued by different trustees and/or are structured differently.

### 3. Sample and Descriptive Statistics

#### 3.1. Sample

We conduct analyses at both the ETF level and the institutional investor level. We identify ETFs in CRSP, with index ETFs identified as those with share code = 73 in the CRSP monthly stock database and index fund

flag = “B” or “D” in the CRSP mutual fund database. We utilize monthly returns, shares outstanding, and trading volume from 1993 to 2022. Removing monthly ETF observations with missing transaction volume, observations where we cannot compute spreads, and observations where the cumulative factor to adjust share prices/shares equals zero leaves us with a sample of approximately 234,000 ETF-month observations (panel A of Table 1 in the Online Appendix).

Our fund-level trading data span from 2001 to 2010 and come from Abel Noser. The data include granular trading activity for Abel Noser’s anonymous institutional clients (Hu et al. 2018 provides a comprehensive discussion of the Abel Noser database). Different vintages of the Abel Noser data have slightly different client and trade information (e.g., the reported numbers of yearly institutions and trades differ in the descriptive statistics of Puckett and Yan 2011 and Hu et al. 2018).

Our institutional-investor-level holdings data come from WhaleWisdom, which provides 13F filing data since 2001 for institutions with AUM of over \$100 million.<sup>8</sup> Institutional investors must disclose long positions valued at more than \$200,000 via Form 13F, which allows us to infer trading activity and wash sale rule avoidance. Observations in WhaleWisdom are at the investor-quarter-ticker level. After requiring identifying information, a match in CRSP, and sufficient information to determine a basis, we have a sample of 143 million investor-quarter-ticker observations. We measure trading activity using differences in holdings from one quarter to the next. For example, if we observe that Filer A holds 5,000 shares of Security B in a quarter and in the next quarter, Filer A reports holding only 1,000 shares of Security B, we infer that Filer A sold 4,000 shares of Security B.<sup>9</sup> We merge security prices from CRSP with the WhaleWisdom data set, and then, we estimate the cost basis of securities traded following Sikes (2014) and Blouin et al. (2017).

We identify the tax sensitivity of institutional investors via reported clientele on Form ADV. Form ADV is a required filing for professional investment advisors that provides information on the institution’s investment approach, employees, and clientele. We match Form ADV filings to 13F filers and define a tax-sensitive filer as one who serves high net worth individuals as clients. Tax-insensitive institutional advisors are those who either report zero HNW individual clients on Form ADV or those identified as pensions or endowments via Brian Bushee’s Institutional Investor Classification Database. Investors who file a Form 13F but not a Form ADV and are not identified as pensions or endowments remain unclassified and do not appear in our sample.

Panel B of Table 1 in the Online Appendix details our sample selection at the 13F filer level, beginning with investor positions available in 13F filings from 2001 to

2022. To conduct our tests, we require non-missing information from WhaleWisdom, CRSP, and Form ADV as well as investor-level AUM and trading activities above zero over our sample. We progressively collapse the sample toward measures aggregated at the filer-quarter level, which is the unit of observation for our second set of empirical tests. We are left with 208,252 filer-quarter observations.

### 3.2. Descriptive Statistics

Panel A of Table 1 presents our descriptive statistics at the ETF-month level. The median (mean) monthly trading volume is 709,000 (12.9 million) shares. The median (mean) ETF market value is \$140.3 million (\$1.67 billion). Roughly 30% of our ETF-month observations have a highly correlated peer ETF (i.e.,  $\geq 99\%$  return correlation).

Moving to our institutional investor data in panel B of Table 1, we classify 70% of filer-quarters as tax sensitive. These institutional investors have median (mean) assets under management of \$349 million (\$3.1 billion) and hold highly correlated ETFs equal to 1% (13%) of total AUM. They also exhibit a greater average propensity to realize losses than gains (28.7% versus 15.5%). Panel C of Table 1 presents the differences in means between tax-sensitive and tax-insensitive institutions. Tax-sensitive filers tend to have less AUM (\$2,520 million versus \$4,635 million) but hold more highly correlated ETF tickers (18 versus 4) and devote more of their portfolios to highly correlated ETFs (16% versus 7%). These differences are consistent with ETFs providing tax advantages for tax-sensitive filers.<sup>10</sup>

## 4. Empirical Design and Results

We predict that the availability of ETFs mitigates a trading friction facing investors: harvesting losses disrupts portfolios. We start our analyses at the macro level by exploring whether the introduction of a near-identical ETF increases or decreases demand for the incumbent ETF.

### 4.1. Do Emerging ETFs Act as Complements or Substitutes for Incumbent ETFs?

An ETF is only useful for avoiding the wash sale rule if at least one other highly correlated ETF is available to swap. Specifically, in this setting, a “treated” ETF would be a security with a near-identical peer available such that investors can exit their current ETF position and enter a new one without meaningfully changing their existing aggregate portfolio.

Li (2024), in an independently developed working paper, provides evidence that the trading volume of incumbent ETFs increases upon the introduction of a highly similar peer ETF using a staggered DiD as the main specification. There is a range of practitioner and legal opinions regarding the legality of tax loss

harvesting with highly correlated ETFs, with several opinions concentrating at a 99% correlation threshold (e.g., Jennings et al. 2020, Wealthfront 2021, Dong 2022, Holt 2023). Other academic work considers ETF similarity based on return correlation thresholds ranging from 90% (Khomyn et al. 2024) to 99% (Brown et al. 2022, Li 2024). We use a 99% threshold as a starting point to examine how incumbent ETFs respond to the introduction of a highly correlated ETF.

In contrast to Li (2024), we study the complementary versus substitutive effects of the introduction of highly correlated peers in the primary market (i.e., share creation) as well as the secondary market (i.e., trading volume). We implement a “stacked regression” approach (Gormley and Matsa 2011, Baker et al. 2022, Dambra et al. 2024) that estimates average treatment effects within a cohort. The stacked regression allows us to obtain clean counterfactuals by “stacking” cohort-specific data sets to avoid contamination from comparing newly treated ETFs with those already treated. Specifically, we create a cohort for each year-month in which incumbent ETFs have near-identical peers introduced.<sup>11</sup> Over our sample period, 110 cohorts are established, and we analyze ETF volume over the 36 months before and after the introduction of a paired ETF. We retain “never treated” ETFs that existed during the same window as control observations.<sup>12</sup> This creates a  $2 \times 2$  data set for each cohort (Baker et al. 2022).

Using this stacked data set, we conduct an analysis similar to that in Li (2024). Our focus is on the introduction of the first highly correlated ETF, meaning that an ETF is treated no more than once. For each  $2 \times 2$  data set, we create a data set-specific identifier (e.g., cohort  $j$ ). We then estimate

$$[Outcome]_{it}^j = \beta_1 Treat_i^j \times Post_t^j + \sum \beta Controls + \alpha_{ij} + \lambda_{tj} + \varepsilon_{it}, \quad (1)$$

where the dependent variable in Equation (1),  $[Outcome]$ , denotes measures of incumbent ETF demand. We estimate primary market demand via the number of ETF shares outstanding (e.g., Brown et al. 2021), where  $\ln(SharesOut)$  is the natural log-transformed monthly shares outstanding of each ETF  $i$  in our sample for month  $t$  in cohort  $j$ . Secondary market activity is captured by  $\ln(Volume)$ , the natural log-transformed monthly trading volume of each ETF  $i$  in our sample for month  $t$  in cohort  $j$ .  $Treat$  equals one for ETF  $i$  if a new ETF is introduced in cohort  $j$  with returns correlated at  $\geq 99\%$  with ETF  $i$  and is zero otherwise. As in Li (2024) and Khomyn et al. (2024), we measure the monthly return correlations between ETFs. We require at least 12 months of overlap to generate the return correlations, but we utilize the entire overlapping life of the ETFs.

**Table 1.** Descriptive Statistics

Panel A: CRSP ETFs—Market-level statistics								
	N	Mean	Standard deviation	P10	P25	Median	P75	P90
<i>ln(Volume)</i>	234,073	8.90	2.55	5.66	7.20	8.87	10.57	12.23
<i>Volume</i>	234,073	12.93	51.73	0.03	0.13	0.71	3.90	20.51
<i>ln(SharesOut)</i>	234,073	8.31	2.13	5.30	6.80	8.31	9.83	11.19
<i>SharesOut</i>	234,073	28.08	70.70	0.20	0.90	4.05	18.65	72.10
<i>HC ETF Peer</i>	234,073	0.30	0.46	0.00	0.00	0.00	1.00	1.00
<i>Spread</i>	234,073	0.00	0.01	0.00	0.00	0.00	0.00	0.01
<i>ln(Market Value)</i>	234,073	18.81	2.32	15.70	17.08	18.76	20.43	21.98
<i>Market Value</i>	234,073	1,665	5,411	6.60	26.22	140.30	743.70	3,502
<i>Volatility</i>	234,073	0.05	0.03	0.01	0.03	0.04	0.06	0.09
<i>B&amp;H Trailing 12</i>	234,073	0.06	0.22	−0.18	−0.05	0.05	0.17	0.30

Panel B: Overall 13F filer sample								
	N	Mean	Standard deviation	P10	P25	Median	P75	P90
<i>AUM</i>	208,252	3,146	10,988	89.60	153.76	349.22	1,198	5,161
<i>HC ETF AUM</i>	208,252	130.83	536.19	0.00	0.00	3.49	63.26	232.84
<i>Volume</i>	208,252	1,106	3,952	11.88	30.70	101.79	436.19	1,869
<i>Turnover</i>	208,252	0.43	0.50	0.06	0.13	0.26	0.55	0.99
<i>Realized Loss</i>	208,252	30.47	142.12	0.01	0.21	1.47	8.94	46.31
<i>Realized Gain</i>	208,252	66.56	245.07	0.12	0.83	4.24	23.27	115.65
<i>Unrealized Loss</i>	208,252	80.84	374.75	0.27	1.41	5.79	25.24	124.42
<i>Unrealized Gain</i>	208,252	683.47	2,851	3.56	14.01	50.41	195.53	897.04
<i>Realized Loss/Vol</i>	208,252	0.04	0.07	0.00	0.00	0.02	0.04	0.09
<i>Unrealized Loss/Vol</i>	208,252	0.25	0.75	0.00	0.01	0.06	0.19	0.55
<i>Unique HC ETFs</i>	208,252	13.43	27.09	0.00	0.00	2.00	15.00	39.00
<i>HC ETF Intensity</i>	208,252	13.46	22.98	0.00	0.00	0.61	17.24	50.81
<i>Tax Sensitivity</i>	208,252	0.70	0.46	0.00	0.00	1.00	1.00	1.00
<i>HC ETF Swaps</i>	208,252	0.20	0.60	0.00	0.00	0.00	0.00	0.57
<i>Available HC ETFs</i>	208,252	362.02	206.70	68.00	175.00	330.00	554.00	658.00
<i>PLR</i>	204,110	28.69	27.33	1.05	6.42	19.93	44.01	73.15
<i>PGR</i>	206,442	15.48	19.73	0.68	2.39	7.49	20.34	43.31
<i>HC ETF PLR</i>	88,595	33.10	37.24	0.00	0.94	14.99	62.66	100.00
<i>HC ETF PGR</i>	115,571	14.51	26.98	0.00	0.10	1.87	12.65	54.91

Panel C: Differences between tax-sensitive and tax-insensitive 13F filers			
Variable	Tax sensitive Mean	Tax insensitive Mean	Standardized difference [(TS − TINS)/σ]
<i>AUM \$</i>	2,520	4,635	−0.19***
<i>HC ETF AUM \$</i>	126.90	140.19	−0.02
<i>Volume \$</i>	804.60	1,822	−0.26***
<i>Turnover</i>	0.34	0.64	−0.60***
<i>Realized Loss \$</i>	22.21	50.13	−0.20***
<i>Realized Gain \$</i>	48.99	108.37	−0.24***
<i>Unrealized Loss \$</i>	60.87	128.40	−0.18***
<i>Unrealized Gain \$</i>	563.08	970.13	−0.14***
<i>Realized Loss/Vol</i>	0.04	0.04	−0.10***
<i>Unrealized Loss/Vol</i>	0.25	0.26	−0.01
<i>Unique HC ETFs</i>	17.53	3.66	0.51***
<i>HC ETF Intensity</i>	16.05	7.31	0.38***
<i>HC ETF Swaps</i>	0.25	0.07	0.29***
<i>PLR</i>	25.77	35.89	−0.37***
<i>PGR</i>	11.28	25.66	−0.73***
<i>HC ETF PLR</i>	30.95	45.51	−0.39***
<i>HC ETF PGR</i>	11.25	29.50	−0.68***

Notes. Panel A presents summary statistics for the CRSP ETF data. The unit of analysis is ETF-month. Pre-transformed *Volume*, *SharesOut*, and *Market Value* are provided for convenience (in millions). Panel B presents the summary statistics for the WhaleWisdom 13F data. The unit of observation is filer-quarter. Variables are defined in Appendix B. Dollar values are in millions. *PLR*, *PGR*, and *HC ETF Intensity* are percentages. Panel C presents variable means for tax-sensitive and tax-insensitive filers separately. The rightmost column presents the difference in means scaled by the population standard deviation. The *t*-tests in panel C are based on filer and year-quarter clustered standard errors.

\*\*\*Significance at 1%.

This construction uses future realizations to determine highly correlated ETFs. To the extent that investors are uncertain of how closely two index-based ETFs will comove immediately upon the introduction of a new ETF, this construction should bias against detecting any share or volume changes in the original ETF. We set  $Post$  equal to one for observations that occur after the introduction of a highly correlated ETF peer in cohort  $j$  in the  $t$  months after the paired ETF introduction and is zero otherwise.

The coefficient of interest in Equation (1) is  $\beta_1$ . To the extent that the introduction of a near-identical ETF acts as a substitute for an incumbent ETF (as economic theory would predict upon the introduction of a nearly identical asset), we expect  $\beta_1 < 0$ . Alternatively, to the extent that introducing a near-identical ETF provides a convenient investment tool to harvest losses and maintain near-identical portfolio allocations, we would expect  $\beta_1 > 0$ . Following prior ETF literature, our control matrix includes the size of the ETF, the bid-ask spread, volatility, and the prior year's buy-and-hold performance (Li 2024, Moussawi et al. 2025). We include

ETF-cohort ( $\alpha_{ij}$ ) and month-cohort ( $\lambda_{tj}$ ) fixed effects, which control for ETF- and time-invariant factors that may affect ETF trading activity. Given the inclusion of fund-cohort and month-cohort fixed effects, the individual  $Treat_i^j$  and  $Post_t^j$  terms are omitted from Equation (1). Non-discrete variables in our regression analyses are winsorized by year at the 1st and 99th percentiles, and we cluster our standard errors by ETF.

We start our regression analyses excluding control variables as time-varying covariates in a stacked regression partially capture treatment effects and can confound inferences (Gao and Huang 2020). In column (1) in Table 2, we observe a 20% increase in shares outstanding for the incumbent ETF in the 36 months after the introduction of a highly correlated paired ETF (coefficient = 0.186,  $p < 0.01$ ).<sup>13</sup>

We expect ETF swapping demand to be more prominent when investors generate capital losses on their investment holdings. In column (2) in Table 2, we interact  $Treat \times Post$  with the trailing 12-month buy-and-hold return ( $B\&H$  Trailing 12). A 12-month return window coincides with tax rules for long-term capital gains

**Table 2.** Effects of Introducing a Highly Correlated Partner ETF

Dependent variable	Primary Market Activity		Secondary Market Activity	
	(1) ln(SharesOut)	(2) ln(SharesOut)	(3) ln(Volume)	(4) ln(Volume)
$Treat \times Post$	0.186*** (3.27)	0.213*** (3.57)	0.269*** (3.90)	0.103** (2.50)
Spread		-14.150*** (-12.98)		-7.249*** (-8.76)
Volatility		1.774*** (4.29)		4.989*** (14.08)
ln(Market Value)				0.842*** (62.34)
$B\&H$ Trailing 12		0.233*** (5.03)		0.411*** (9.11)
$B\&H$ Trailing 12 $\times$ $Treat$		0.059 (0.68)		-0.292*** (-3.69)
$B\&H$ Trailing 12 $\times$ $Post$		0.146*** (3.94)		-0.267*** (-7.05)
$B\&H$ Trailing 12 $\times$ $Treat$ $\times$ $Post$		-0.291*** (-2.59)		-0.234** (-2.34)
Observations	4,690,132	4,690,132	4,690,132	4,690,132
Fixed effects	Cohort $\times$ date, cohort $\times$ ETF	Cohort $\times$ date, cohort $\times$ ETF	Cohort $\times$ date, cohort $\times$ ETF	Cohort $\times$ date, cohort $\times$ ETF
Adjusted $R^2$	0.912	0.914	0.852	0.903

*Notes.* This table presents results of stacked DiD regressions that examine monthly ETF primary and secondary market activity surrounding the introduction of a highly correlated peer.  $\ln(\text{SharesOut})$  is the natural log of month-end shares outstanding.  $\ln(\text{Volume})$  is the natural log of monthly trading volume.  $Treat$  is equal to one for an ETF incumbent if another ETF within a cohort has returns that are correlated at 99% or more and is equal to zero otherwise.  $Post$  is equal to one for observations that occur in the 36 months following the highly correlated peer ETF introduction and is equal to zero otherwise. In column (2), we omit  $\ln(\text{Market Value})$  to avoid multicollinearity between market value and shares outstanding. Cohorts are created with never treated control observations. Standard errors are clustered by ETF, and  $t$ -statistics are reported in parentheses. Appendix B contains detailed variable definitions.

\*\*Significance at 5%; \*\*\*significance at 1%.

and losses, and it is consistent with the window considered in Brown et al. (2022) and Li (2024).<sup>14</sup> We find that the main effect of high prior returns is increased shares outstanding. However, the interaction term indicates that shares outstanding following the introduction of a paired ETF increase (decrease) as a function of the incumbent ETF's negative (positive) stock returns is symptomatic of traders being more likely to exit (hold) their existing positions to generate capital losses (delay recognition of capital gains). We further explore the relation between ETF holdings and 13F filer loss recognition in Section 4.5.

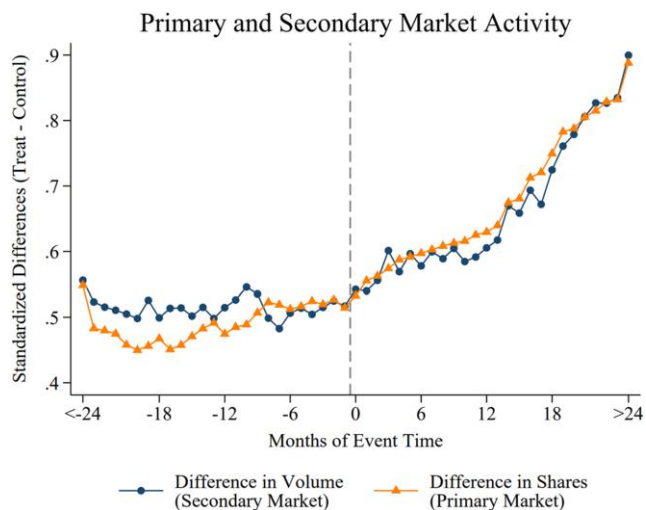
We find similar evidence to Li (2024) in the secondary market with  $\ln(\text{Volume})$  as the dependent variable. The coefficient on  $Treat \times Post$  in column (3) in Table 2 represents a 31% increase in ETF trading volume in the postperiod. Collectively, our results suggest that the introductions of near-identical ETFs act as complements to incumbent ETFs. To further analyze time-series changes in shares outstanding and trading volume for newly paired ETFs, we analyze the differential trends in shares outstanding and trading volume between treatment ETFs (i.e., incumbents) and control ETFs (i.e., funds where a highly correlated fund is not introduced) in the 36 months surrounding the introduction of a near-identical ETF. In the pre-ETF introduction period, we observe relatively stable shares outstanding and volume differentials between treatment and control ETFs. Post-ETF introduction, we observe a marked increase in the difference in both shares outstanding and trading volume between treatment and control ETFs (see Figure 1). This illustrative evidence supports our conjecture that the demand for a newly paired ETF is permanent and that the increase in ETF differential shares outstanding and trading activity only occurs *after* the introduction of a new ETF.

The results thus far are highly suggestive of ETF wash sale activities. However, to the extent that ETF competitors introduce near-identical products when a comparable incumbent product has been especially successful, our results may not be because of tax loss harvesting. If these ETF introductions coincide with strong surges in underlying demand for indexed ETF products, the increased trading volume and share creation may reflect a strong growth trend in the sector rather than tax loss harvesting. Furthermore, increases in ETF trading volume and share creation alone would not link institutional investors to wash sale avoidance. For these reasons, we turn to institutional trading data to provide additional evidence.

#### 4.2. Granular Evidence of Swap Trades

The complementary relation between incumbent and new-entry ETFs suggests that investors use near-identical assets to harvest tax losses. If that is what drives the increased demand in redundant ETFs, we

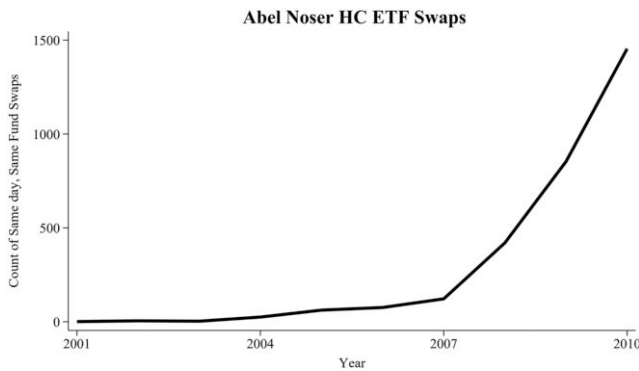
**Figure 1.** (Color online) Event Time Differences in ETFs with Highly Correlated Partners



*Notes.* This figure presents a graph of the ETF market in event time, where period 0 is the month in which an ETF has a highly correlated peer ETF introduced. Treated ETFs are those that have a highly correlated partner, and control ETFs are those that do not. The differences in both primary market activity and secondary market activity are plotted (i.e., standardized shares outstanding or trading volume for treated firms minus standardized shares outstanding or trading volume for control firms) for each month  $t$  individually for months between  $t = -24$  and  $t = 24$  and jointly for months  $-36 \leq t \leq 25$  and  $25 \leq t \leq 36$ .

should observe institutional investors swapping between highly correlated ETFs within the wash sale window. To test this conjecture, we exploit proprietary daily institutional investor trading data from Abel Noser. We start by counting the number of transactions where institutions engage in same-day, same-fund transactions where buys and sells of HC ETFs occur. In Figure 2, we observe that the frequency of HC ETF swaps over our sample has increased exponentially in our Abel Noser sample from a handful in the early 2000s to more than 1,400 in 2010. Notably, our tranche of Abel Noser data predates the recent ETF boom and only includes 8%–15% of the total dollar amount of U.S. institutional investor trades (Puckett and Yan 2011, Hu et al. 2018). Nevertheless, we observe a marked increase in institutional investor transactions consistent with harvesting tax losses as more HC ETFs come online through the early 2000s.

For illustrative purposes, in Appendix A, we provide examples of several ETF swaps across different Abel Noser clients. These trades are executed on the same day by the same fund. For example, Abel Noser client #732 sold \$3,615,739 of IVV and purchased \$3,563,270 of SPY on December 21, 2004. The correlation between IVV and SPY monthly price movements in the year leading up to the swap trade was 1.000, consistent with the client expecting tax advantages from the trade rather than differences in pretax returns. Thus, client

**Figure 2.** Institutional Investors and Highly Correlated ETFs

Notes. This figure plots the frequency of same day, same fund trades between highly correlated ETFs in the Abel Noser data. The *ClientMgrCode* data field from Abel Noser is used to identify intrafund activity. Only investment managers (*ClientTypeCode* = 2) are included in the chart.

#732 engaged in trades totaling \$7.2 million in dollar volume while shifting its economic position by only \$52,000 (less than 1% of the dollar volume). We are unaware of any economic rationale other than tax loss harvesting that explains these trades.<sup>15</sup>

#### 4.3. Estimating the Magnitude of ETF-Based Wash Sale Avoidance

Abel Noser data are not well suited to estimating the magnitude of loss harvesting. First, Abel Noser trading data extend from 2001 to 2010 when ETF trading volume averaged only 25% of 2022 volume. Second, Abel Noser's fund manager identifier (*ClientMgrCode*) is not stable across time, limiting our ability to identify time-series variation in institutional investor trading (see Hu et al. 2018 for a review of Abel Noser data features).<sup>16</sup> Finally, the data include trades by institutions but not their historical holdings, making it difficult to estimate investment bases and calculate gains and losses. To address these limitations and to more broadly estimate the magnitude of tax-motivated ETF holdings, trading volume, and capital loss recognition across a comprehensive panel of tax-sensitive institutional traders, we turn to 13F filings. We start these analyses by first examining whether tax-sensitive investors relative to tax-insensitive investors are more likely to invest in highly correlated ETFs. We estimate the following model:

$$\begin{aligned}
 [HC\ ETF\ Preference]_{it} = & \beta_1 Tax\ Sensitivity_i \\
 & + \beta_2 Available\ HC\ ETFs_t \\
 & + \beta_3 Tax\ Sensitivity_i \\
 & \times Available\ HC\ ETFs_t + \theta_i + \varepsilon_{it}.
 \end{aligned}
 \tag{2}$$

We measure *HC ETF Preference* as either *Unique HC ETFs* (the distinct number of highly correlated ETFs

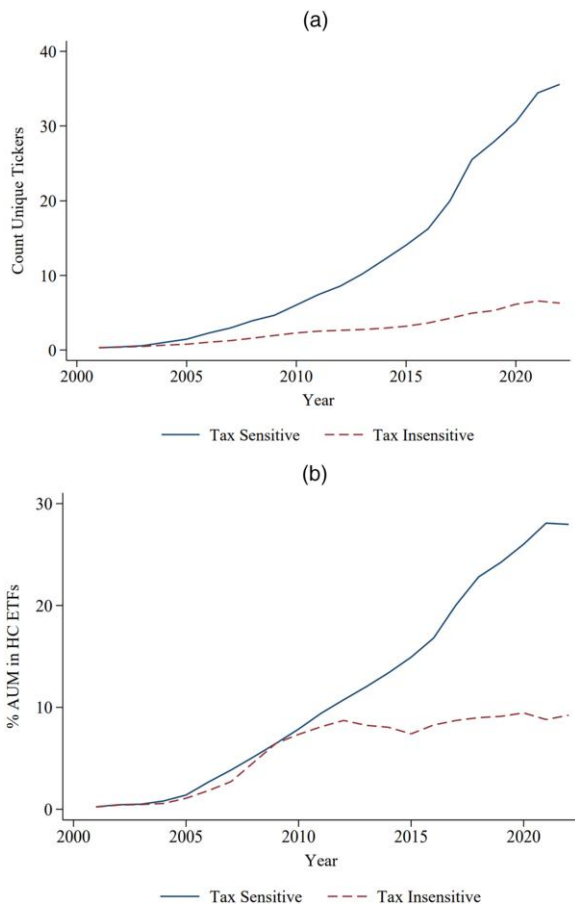
held by investor  $i$  in quarter  $t$ ) or *HC ETF Intensity* (highly correlated ETF AUM holdings over total AUM holdings for investor  $i$  in quarter  $t$ ). Both are measures of a preference for and adoption of highly correlated ETFs. With quarterly 13F data, our unit of observation is at the filer-quarter level. We define *Tax Sensitivity* as an indicator equal to one when a 13F filer reports non-zero HNW clients on Form ADV and zero if there are no HNW clients or if the institution is classified as a pension or endowment. Investment advisors file Form ADV with the SEC and must disclose information about their organization and clientele, including how many of their clients fall into various categorizations (e.g., individuals, high-net-worth individuals, pensions, and governments). We collected Form ADVs from 2009 to 2022. We merge the time series of ADV classifications with the WhaleWisdom data using a fuzzy match of multiple identifiers (year, quarter, Central Index Key (CIK), name, and phone number).<sup>17</sup> We then identify the modal value of *Tax Sensitivity* for each institution and apply that classification for all periods.<sup>18</sup> We predict that, ceteris paribus, tax-sensitive institutional investors will hold more near-identical ETFs ( $\beta_1 > 0$ ).

We further predict that tax-sensitive institutional investors will become more active in accumulating near-identical ETFs as the number of available highly correlated ETFs increases. We measure *Available HC ETFs* as the standardized count of unique highly correlated ETFs available in CRSP each quarter  $t$  of our sample. The interaction between *Tax Sensitivity* of institution  $i$  and *ETF HC Availability* in quarter  $t$  captures our empirical prediction. A  $\beta_3 > 0$  suggests that tax-sensitive institutions intensify their ETF holdings more rapidly than their peers as ETF availability increases. We also include filer-level fixed effects ( $\theta_i$ ) in some specifications.<sup>19</sup>

We start our empirical analysis by illustrating time-series changes in institutional investor holdings between tax-sensitive and tax-insensitive investors. Panel (a) of Figure 3 shows that *Unique HC ETFs* have increased significantly more rapidly for tax-sensitive investors than tax-insensitive investors. In panel (b) of Figure 3, we show that that *HC ETF Intensity* increased rapidly for tax-sensitive investors over the past 20 years, whereas *HC ETF Intensity* has remained largely flat for tax-insensitive investors since 2010. Although this provides descriptive evidence of tax-sensitive investors' preference for highly correlated ETFs, we explore this prediction further in our empirical analyses.

Table 3 presents the results of estimating Equation (2) with *Unique HC ETFs* as the dependent variable. In column (1) in Table 3, we find that tax-sensitive investors hold 13 more unique HC ETFs than tax-insensitive institutional investors (coefficient = 13.488,  $p < 0.01$ ). In column (2) in Table 3, we interact *Tax Sensitivity* with *Available HC ETFs*. We find that as more HC ETFs become available, tax-sensitive institutions are incrementally

**Figure 3.** (Color online) Institutional Investors and Highly Correlated ETFs



*Notes.* This figure presents descriptive graphs of the ETF market for each year of our sample. Panel (a) shows the change in average quarterly count of HC ETFs held for both investor types. Panel (b) shows the change in *HC ETF Intensity* (defined as *HC ETF AUM/AUM*) for both investor types over time. Data are from CRSP and WhaleWisdom. (a) Average quarterly count of unique HC ETFs held. (b) Average quarterly HC ETF intensity.

more likely to own a greater number of distinct highly correlated ETF tickers. The second and fourth specifications include a filer fixed effect and thus, omit the static *Tax Sensitivity* measure. Column (3) in Table 3 presents results with *HC ETF Intensity* as the dependent variable. Our main effect is consistent with tax-sensitive institutions holding 8.4 percentage points more of their AUM in highly correlated ETFs ( $p < 0.01$ ). The interaction of *Tax Sensitivity* and *Available HC ETFs* is positive but statistically insignificant for the full sample ( $t$ -statistic: 1.50). Panel (b) of Figure 3 shows that the intensity of the tax-sensitive institutional holdings accelerates relative to the tax-insensitive institutions only in the latter half of the sample. We confirm this by reperforming the test from column (4) in Table 3 after excluding the first five years of the sample and find positive and significant results (untabulated) consistent with tax-sensitive institutions

having an incremental preference for highly correlated ETFs.

#### 4.4. Establishing ETF Swapping as an Institutional Investor Tax Strategy

Although simple, the results in Table 3 and illustrations in Figure 3 establish an important empirical fact. There is a starkly different appetite for highly correlated ETFs based on the tax sensitivity of an institutional investor. It can be inferred that ETFs offer something to tax-sensitive institutions that is not appealing to tax-insensitive investors. However, the relation between ETF holdings and investor type may be driven by investor clientele effects unrelated to tax sensitivity. For instance, Bhattacharya et al. (2017) find that retail investors tend to return chase using ETFs, and the strong market performance of the S&P 500 in recent years could drive the uptick in ETF AUM among tax-sensitive investors.

Alternatively, the increase in ETF holdings could be driven by tax minimization preferences for ETFs unrelated to skirting wash sale rules. Moussawi et al. (2025) argue that the increase in tax-sensitive institutional investor holdings is influenced by the ability of ETFs to avoid within-ETF capital gains taxes from portfolio position changes. ETFs can use “heartbeat trades” (i.e., creation and in-kind redemption in the primary market) to avoid distributing taxable capital gains from trading appreciated securities (e.g., around index reconstitutions).

In this section, we examine whether 13F filers swap ETFs that are highly correlated with one another. In other words, we identify instances where an investor sells one ETF (e.g., VOO) while simultaneously buying an offsetting amount of another ETF (e.g., IVV). We focus on secondary market trading activity to better understand whether the increase in ETF holdings documented above can be attributable to avoiding the wash sale rule. Secondary market swapping of highly correlated ETFs is unrelated to the other tax advantages documented by Moussawi et al. (2025). Using the following empirical design, we examine ETF trading activity in the secondary market:

$$\begin{aligned}
 HC\ ETF\ Swaps_{it} = & \beta_1 Tax\ Sensitivity_i \\
 & + \beta_2 HC\ ETF\ Intensity_{it} \\
 & + \beta_3 Tax\ Sensitivity_i \\
 & \times HC\ ETF\ Intensity_{it} \\
 & + \sum \beta Controls + \psi_t + \theta_i + \varepsilon_{it}.
 \end{aligned}
 \tag{3}$$

*HC ETF Swaps* is the natural log of one plus the dollar volume of offsetting purchases and sales of highly correlated ETFs for investor  $i$  in quarter  $t$ .<sup>20</sup> We identify ETF swap activity by observing quarters in which an

**Table 3.** Unique HC ETFs Held and HC ETF Intensity

Dependent variable	(1) <i>Unique HC ETFs</i>	(2) <i>Unique HC ETFs</i>	(3) <i>HC ETF Intensity</i>	(4) <i>HC ETF Intensity</i>
<i>Tax Sensitivity</i>	13.488*** (10.32)		8.441*** (8.51)	
<i>Available HC ETFs</i>	9.343*** (43.28)	2.847*** (6.59)	7.375*** (34.24)	2.059*** (6.38)
<i>Tax Sensitivity</i> × <i>Available HC ETFs</i>		5.777*** (9.75)		0.479 (1.50)
Constant	3.931*** (4.44)		7.520*** (10.08)	
Observations	208,252	208,140	208,252	208,140
Fixed effects	None	Filer	None	Filer
Adjusted R <sup>2</sup>	0.174	0.803	0.133	0.845

Notes. This table presents results of regressing *Unique HC ETFs* and *HC ETF Intensity* on *Tax Sensitivity* and *Available HC ETFs*. The dependent variable *Unique HC ETFs* is the count of distinct ETFs held at the end of the quarter. The dependent variable *HC ETF Intensity* is equal to  $(HC\ ETF\ AUM/Total\ AUM) \times 100$  at the end of each quarter. We measure *Tax Sensitivity* as an indicator equal to one if an institution serves high-net-worth individuals. We measure *Available HC ETFs* as the standardized count of unique HC ETFs available in CRSP each quarter from 2001 to 2022. Standard errors are clustered by filer and by year-quarter, and *t*-statistics are reported in parentheses.

\*\*\*Significance at 1%.

investor sold an ETF while simultaneously buying one or more ETFs that are correlated with the sold ETF at 99% or greater. In other words, when we observe offsetting purchases and sales of two highly correlated ETFs in a given fiscal quarter (call them ETF *b* and ETF *c*) such that  $\Delta ETF_{bt} < 0$  and  $\Delta ETF_{ct, c \neq b} > 0$ , then we define the *HC ETF Swaps* as the dollar volume offset between the purchases and sales transactions:  $\min[(|\Delta ETF_{bt}| \times Price_{bt}), (|\Delta ETF_{ct}| \times Price_{ct})]$ .

We define *HC ETF Intensity* as noted above. Our coefficients of interest in Equation (3) are  $\beta_1$  and  $\beta_3$ , the interaction of *Tax Sensitivity* and *HC ETF Intensity*. We expect that tax-sensitive institutional investors will engage in more ETF swaps than tax-insensitive investors (i.e.,  $\beta_1 > 0$ ). In addition, we expect that tax-sensitive institutional investors will engage in more ETF swaps as their ETF holdings increase ( $\beta_3 > 0$ ). Our control matrix follows prior literature as we control for turnover, fourth-quarter selling, and the level of realized gains, realized losses, unrealized gains (UGs), and unrealized losses (Sikes 2014).

In column (1) in Table 4, we regress *HC ETF Swaps* on *Tax Sensitivity* and year and quarter fixed effects. Consistent with our prediction that swapping between HC ETFs serves a tax purpose, we find that institutions classified as tax-sensitive have significantly higher *HC ETF Swaps* (coefficient = 0.165;  $p < 0.01$ ).

Next, we include *HC ETF Intensity* and the interaction of *Tax Sensitivity* and *HC ETF Intensity* in column (2) in Table 4. Consistent with *Tax Sensitivity* being especially important to harvesting ETF losses in the presence of increasing *HC ETF Intensity*, we find that the coefficient on the interaction is significant at the 1% level (coefficient = 0.006). In column (3) in Table 4, we

estimate our full Equation (3) model, including additional controls and investor fixed effects, and we again find a positive coefficient on *Tax Sensitivity* × *HC ETF Intensity* (coefficient = 0.002;  $p < 0.01$ ). Notably, we observe that the ratio of realized losses over volume (*Realized Loss/Vol*) is increasing in *HC ETF Swaps*. In other words, more swapping is significantly associated with increased realization of losses, a finding that we examine more directly in Section 4.5. Thus, consistent with our prediction, we find that tax-sensitive institutions engage in more swapping of highly correlated ETFs, that this propensity is an increasing function of AUM devoted to highly correlated ETFs, and that the swapping activity is associated with more realized losses.

Figure 4 plots the data by tax sensitivity when scaling the amount of *HC ETF Swaps* by total volume traded. Among tax-sensitive institutional investors, ETF swaps increased from \$280 million in 2001 to \$106 billion in 2022. Tax-sensitive institutional investors have amassed \$417 billion of ETF swaps since 2001.<sup>21</sup>

#### 4.5. Realizing Losses and Gains

Although simultaneously entering and exiting economically identical ETFs seem to have no purpose beyond tax avoidance, institutions may act for other reasons. For example, they could be swapping into ETFs with lower management fees (Brown et al. 2022) or ETFs managed by personal acquaintances. To further validate that tax-sensitive institutions swapping ETFs are harvesting losses, we employ the proportion of losses realized measure from prior literature (Odean 1998, Sikes 2014). We estimate Equation (4) with *PLR* as the

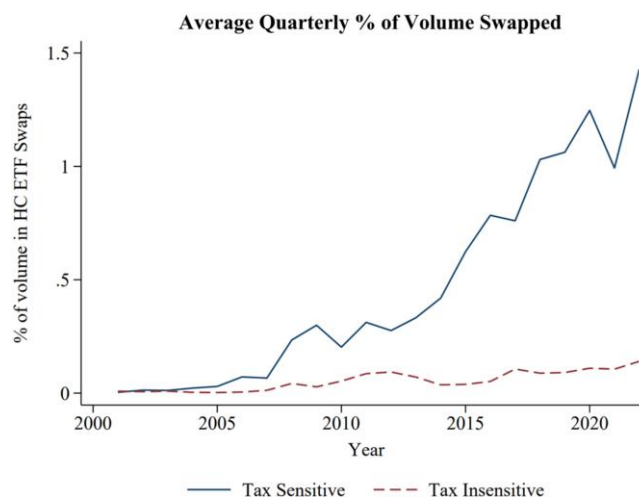
**Table 4.** Swapping Activity

Dependent variable = <i>HC ETF Swaps</i>	(1)	(2)	(3)
<i>Tax Sensitivity</i>	0.165*** (8.10)	0.040*** (3.76)	
<i>HC ETF Intensity</i>		0.003*** (5.98)	0.001*** (3.87)
<i>Tax Sensitivity</i> × <i>HC ETF Intensity</i>		0.006*** (8.52)	0.002*** (3.91)
<i>Unique HC ETFs</i>			0.014*** (27.35)
<i>Fourth Quarter</i>			0.011 (1.07)
<i>Realized Loss/Vol</i>			0.036** (2.00)
<i>Unrealized Loss/Vol</i>			−0.004 (−1.60)
<i>Realized Gain/Vol</i>			−0.022 (−1.07)
<i>Unrealized Gain/Vol</i>			−0.007*** (−9.13)
<i>Turnover</i>			0.010*** (2.69)
Observations	208,252	208,252	208,140
Fixed effects	Year, quarter	Year, quarter	Year, filer
Adjusted $R^2$	0.074	0.168	0.628

Notes. This table presents results of regressing dollar volume of highly correlated ETF swaps on *Tax Sensitivity* and *HC ETF Intensity*. The dependent variable, *HC ETF Swaps*, equals the natural log of one plus the dollar summation of offsetting buys and sells in HC ETFs for each filer in each quarter. We measure *Tax Sensitivity* as an indicator equal to one if an institution serves high-net-worth individuals. We measure *HC ETF Intensity* as equal to  $(HC\ ETF\ AUM / Total\ AUM) \times 100$  at the end of each quarter. We measure *Unique HC ETFs* as the count of distinct HC ETFs held at the end of the quarter. We measure *Fourth Quarter* as an indicator variable equal to one for the fourth calendar quarter. We measure *Realized Loss (Gain)/Vol* as equal to realized losses (gains)/volume. We measure *Unrealized Loss (Gain)/Vol* as equal to unrealized losses (gains)/volume. We measure *Turnover* as equal to volume/total AUM. Standard errors are clustered by filer and by year-quarter, and *t*-statistics are reported in parentheses.

\*\*Significance at 5%; \*\*\*significance at 1%.

**Figure 4.** (Color online) Average Quarterly Trading Volume Dedicated to *HC ETF Swaps*



Notes. This figure shows the dollar volume of *HC ETF Swaps* scaled by the dollar volume of total trades during quarter *t*. We separately plot tax-sensitive and tax-insensitive filers. Data are from CRSP and WhaleWisdom.

dependent variable:

$$PLR_{it} = \beta_1 HC\ ETF\ Swaps_{it} + \beta_2 Tax\ Sensitivity_{it} \\ \times HC\ ETF\ Swaps_{it} + \sum \beta Controls \\ + \psi_t + \theta_i + \varepsilon_{it}. \quad (4)$$

Calculating realized gains and losses requires a cost basis at the time of sale. We follow a similar approach to prior literature to proxy for cost basis, valuation, and trading behavior (Jin 2006, Sikes 2014, Blouin et al. 2017). Securities held during the first quarter of 2001 (the first year with quarterly 13Fs available from WhaleWisdom) are treated as purchased during that quarter. From then on, quarter-to-quarter increases in each security are treated as “buys,” and quarter-to-quarter decreases are treated as “sells.” If an institution owns multiple lots, we assume a highest in, first out (HIFO) approach to calculate realized gains and losses.<sup>22</sup> After establishing each position’s basis, sales price, and implied gain or loss, we aggregate gains and losses to the institution-quarter level. We measure *PLR* as the proportion of realized losses for institution *i* in time *t* scaled by their total realized losses and ULs in that quarter  $\left(\frac{RL_{it}}{RL_{it}+UL_{it}}\right)$ .

As a baseline test, we first examine the filer’s overall *PLR* as our dependent variable in Equation (4). We do so to address our fundamental research question:

whether ETFs mitigate frictions imposed by the wash sale rule in harvesting capital losses. Our second set of tests restricts the *PLR* to be related only to highly correlated ETFs (*HC ETF PLR*) to more directly tie the ETF swapping activity to losses harvested in the filer’s ETF portfolio. We anticipate that *HC ETF Swaps* will lead to higher *PLRs* for tax-sensitive investors ( $\beta_2 > 0$ ). We additionally control for the proportion of gains realized (*PGR*) to account for the tax incentive to realize losses as offsets for concurrently realized gains. We construct *PGR* as the proportion of realized capital gains (relative to total RGs and UGs in a quarter  $\left(\frac{RG_{it}}{RG_{it}+UG_{it}}\right)$ ). In addition to *PGR*, we include an indicator to capture fourth-quarter trading and 13F filer turnover (i.e., trading volume scaled by AUM) as well as filer fixed effects to identify within-investor effects and limit correlated omitted variable concerns.<sup>23</sup> Appendix B provides detailed variable definitions.

Table 5 presents the results of regressing *PLR* and *ETF PLR* on measures of *Tax Sensitivity* and *HC ETF Swaps*. In column (1) in Table 5, we see that *PLR* is an increasing function of *HC ETF Swaps* (coefficient = 2.584;  $p < 0.01$ ). In other words, institutions that engage in more swapping between highly correlated ETFs capture a larger proportion of their losses. Thus, our evidence suggests that ETF swapping appears to ameliorate the trading frictions imposed by the wash

**Table 5.** Proportion of Losses Realized and ETF Proportion of Losses Realized

Dependent variable	(1) <i>PLR</i>	(2) <i>PLR</i>	(3) <i>PLR</i>	(4) <i>HC ETF PLR</i>	(5) <i>HC ETF PLR</i>	(6) <i>HC ETF PLR</i>
<i>HC ETF Swaps</i>	2.584*** (9.53)	0.505 (1.47)	0.772** (2.48)	4.226*** (12.39)	2.039*** (3.90)	2.215*** (4.16)
<i>Tax Sensitivity</i> $\times$ <i>HC ETF Swaps</i>		2.578*** (5.13)	1.806*** (4.01)		2.668*** (4.45)	2.364*** (4.23)
<i>Unique HC ETFs</i>			−0.007 (−1.16)			−0.070*** (−5.82)
<i>Fourth Quarter</i>			3.759*** (8.01)			0.934 (0.93)
<i>PGR</i>			−0.029** (−2.24)			
<i>HC ETF PGR</i>						0.132*** (10.48)
<i>Turnover</i>			26.969*** (26.61)			22.406*** (23.01)
Observations	203,986	203,986	202,168	88,281	88,281	77,864
Fixed effects	Year, quarter, filer	Year, quarter, filer	Year, filer	Year, quarter, filer	Year, quarter, filer	Year, filer
Adjusted $R^2$	0.281	0.281	0.421	0.143	0.144	0.208

*Notes.* This table presents results of regressing *PLR* (*HC ETF PLR*) on *HC ETF Swaps* and *Tax Sensitivity*. The dependent variable *PLR* is defined as  $(Realized\ Losses / All\ Losses) \times 100$ . We define *HC ETF PLR* similarly but only for *HC ETF* losses. We measure *HC ETF Swaps* as the log of one plus the dollar summation of offsetting sells and buys in highly correlated ETFs. We measure *Tax Sensitivity* as an indicator equal to one if an institution serves high-net-worth individuals. We measure *Unique HC ETFs* as the count of distinct *HC ETFs* held at the end of the quarter. We measure *Fourth Quarter* as an indicator variable equal to one for the fourth calendar quarter. We measure *PGR* (*HC ETF PGR*) in the same manner as *PLR* (*HC ETF PLR*), but for gains rather than losses. We define *Turnover* as equal to volume/total AUM. Standard errors are clustered by filer and by year-quarter, and *t*-statistics are reported in parentheses.

\*\*Significance at 5%; \*\*\*significance at 1%.

sale rule (e.g., Jensen and Marekwica 2011) in that investors maintain a similar portfolio allocation and harvest tax losses. Column (2) in Table 5 makes clear that tax-sensitive institutions drive the *HC ETF Swaps* as the size and significance of the effect are concentrated in the interaction of *Tax Sensitivity* and *HC ETF Swaps*. In column (3) in Table 5, we include control variables for turnover, the fourth quarter, and the count of unique ETFs held. Consistent with prior literature, we find that significantly greater proportions of losses are realized in the fourth calendar quarter (Sikes 2014). A one-standard-deviation increase in *HC ETF Swaps* for tax-sensitive institutional investors (0.65) is thus associated with a 1.2% increase in *PLR* ( $1.806 \times 0.65$ ). Given that the sum of the mean realized losses and unrealized losses for our tax-sensitive subsample is \$83.08 million ( $22.21 + 60.87$ ), a one-standard-deviation increase in swapping activity equates to \$972,000 of additional losses realized per tax-sensitive filer-quarter. Aggregated over our sample, 70% of which are tax sensitive, a one-standard-deviation increase equates to \$138 billion ( $\$972,000 \times 70\% \times 202,168$  observations).<sup>24</sup>

We also find that, unsurprisingly, *Turnover* is the primary driver of *PLR*.<sup>25</sup> The relationship is mechanical as more trading implies more realization of capital gains and losses. The coefficient on the interaction of *Tax Sensitivity* and *HC ETF Swaps* remains positive and significant, indicating that tax-sensitive investors' ETF swaps are a driver of harvesting capital losses.

Columns (4)–(6) in Table 5 present a similar empirical design but with *ETF PLR* as the dependent variable. The sample size for this test is smaller than that of the first three specifications because *HC ETF PLR* is missing for fund-quarters with no highly correlated ETF holdings.<sup>26</sup> Again, we find results consistent with our predictions. Notably, in column (6) in Table 5, we observe that the coefficient on the *Fourth Quarter* indicator is not significantly different from zero. This suggests that harvesting losses among ETFs is less seasonally driven than among traditional equities (as documented in Sikes 2014), consistent with ETF swapping being a viable strategy throughout the calendar year.<sup>27</sup> Matthews (2016) corroborates this empirical finding, noting that

institutions face no tracking error and minimum transaction costs when swapping ETFs intrayear, which makes year-end-only ETF loss harvesting inefficient.

Another approach to estimating losses harvested via wash sale avoidance is to sum realized losses associated with HC ETF swapping activity. Table 6 does so for different definitions of “highly correlated” ranging from 95% to identical index ETFs. Using our  $\geq 99\%$  correlation threshold, we estimate that 13F filers realized \$84 billion of losses over our sample period from HC ETFs that we identify as being involved in swapping activity. Loss realizations are higher in bear market years (e.g., 2002, 2008–2009, 2018, and 2022) and in later years (untabulated). If the losses offset short-term capital gains or income, this represents \$31 billion in tax revenue. For scale, the IRS 2023 enacted budget was \$12.3 billion. Because of the assumptions required to generate these summations (i.e., assumptions regarding acquisition and disposal dates and prices, initial basis assignments, intraquarter activity, and intrafund activity) as well as the likelihood that investor behavior would change with rule clarification (i.e., if regulators specified that swapping HC ETFs violated the wash sale rule, the extent of swapping behavior would likely decline), we caution against interpreting these estimates as the likely increase in treasury revenue if the wash sale rule was clarified to disallow losses from swapping HC ETFs. Rather, we view these figures as indicative of the significant and growing size of this activity absent intervention.

#### 4.6. A Separate Effect from Primary Market Tax Advantages

Research documents two other tax benefits of ETFs. First, many ETFs have low turnover in their holdings, which limits the frequency with which they realize taxable gains (e.g., Colon 2023). Second, investors in ETFs benefit from section 852(b)(6), which exempts in-kind distributions from taxation through a process referred to as “heartbeat trades” (Choi et al. 2025, Moussawi et al. 2025). Large inflows from authorized participants create new ETF shares followed within a few days by large outflows in the form of in-kind redemptions of

**Table 6.** Estimating Tax Revenue Losses

	Same index	99% correlation	95% correlation
Tax-sensitive dollars swapped, \$	41,749M	417,192M	1,276,258M
Tax-sensitive losses associated with swapping-involved HC ETFs, \$	6,756M	84,479M	163,618M
Potential tax revenue at a 37% short-term capital gains rate, \$	2,500M	31,257M	61,649M
Potential tax revenue at a 20% long-term capital gains rate, \$	1,351M	16,896M	33,342M

*Notes.* This table provides estimates of swapping activity and associated losses at alternative threshold definitions of “highly correlated.” The estimates are derived by (1) identifying same quarter swaps between HC ETFs, (2) estimating the cost basis of each swapped security, (3) flagging the HC ETFs involved in the swap that have realized losses, and (4) summing the realized losses for only institutions classified as tax sensitive. “Same index” ETFs are those that explicitly state that they track the same defined index (e.g., we would include swapping between two ETFs that both track the S&P 500 in this measure but not between two ETFs that both focus on “midcap growth equities” without specifying a specific index). M, millions.

appreciated shares on the primary market. In this way, appreciated securities are disbursed without any realized capital gains that would have been generated if the securities had been sold and cash had been disbursed. This second tax benefit accrues to investors for holding ETFs.

Similar to our findings, Moussawi et al. (2025) document that tax-sensitive institutions allocate larger portions of their portfolios to ETFs. The authors attribute this preference to the primary market tax advantages. In Table 7, we illustrate that tax-sensitive filers are significantly more likely to hold ETFs that engaged in large dollar volumes of heartbeat trades (column (1) in Table 7) and large numbers of heartbeat trades (column (2) in Table 7), consistent with Moussawi et al. (2025). An alternative explanation to our main findings is that they are simply a side effect of investors crowding into ETFs for the primary market tax advantages. To demonstrate that the effects that we document are both distinct from and incremental to those identified in Moussawi et al. (2025), we estimate the following model:

$$\begin{aligned} HC\ ETF\ Swaps_{it} = & \beta_1 Tax\ Sensitivity_{it} \\ & + \beta_2 HC\ ETF\ Intensity_{it} \\ & + \beta_3 Tax\ Sensitivity_{it} \\ & \times HC\ ETF\ Intensity_{it} \\ & + \beta_4 Heartbeat\ Intensity\ Quintile \\ & + \psi_t + \varepsilon_{it}. \end{aligned} \quad (5)$$

Variables are previously defined, and *Heartbeat Intensity Quintile* is a count variable that sorts each filer-quarter into its quintile of ETF heartbeat trade activity.

ETF heartbeat trade activity is identified by following the procedure described in Moussawi et al. (2025) and using the ETF Global database for ETF fund flows. If our main findings are subsumed by a preference for ETFs that offer heartbeat trades, we anticipate an insignificant  $\beta_3$  coefficient with a positive and significant  $\beta_4$  coefficient. On the other hand, if institutional investors swap highly correlated ETFs for reasons beyond the primary market tax advantages, we would predict  $\beta_3 > 0$ . Importantly, we do not believe that the primary market tax advantages and secondary market tax advantages are mutually exclusive; rather, we would anticipate that both play a meaningful role for tax-sensitive institutions.

Results from estimating Equation (5) are in column (3) in Table 7. We see a positive and significant coefficient on both  $\beta_3$  and  $\beta_4$ , suggesting that the secondary market wash sale avoidance that we document is incremental to tax-sensitive preferences for ETFs that engage in heartbeat trades. We also re-estimate Equation (5) on a subsample of filer-quarters that hold ETFs below the median level of heartbeat trade activity. Consistent results in this subsample would suggest that the effects are not primarily driven by swapping ETFs that engage in heartbeat trades. We find that the estimates are very similar to those in Table 4 in this subsample, and we report the results in column (4) in Table 7. Again, this finding indicates that the secondary market trading advantages are distinct from and incremental to the primary market advantages of holding ETFs.

#### 4.7. Alternative Measures of Tax Sensitivity

Our base measure of tax sensitivity is an objective cut-off between having and not having HNW individuals

**Table 7.** Heartbeat Trades and Swapping Activity

Dependent variable	(1) <i>Heartbeat Size Intensity</i>	(2) <i>Heartbeat Count Intensity</i>	(3) <i>HC ETF Swaps</i>	(4) <i>HC ETF Swaps</i>
<i>Tax Sensitivity</i>	68.707*** (5.26)	0.046*** (10.79)	-0.011 (-1.22)	0.001 (0.55)
<i>HC ETF Intensity</i>			0.003*** (5.15)	0.002*** (5.51)
<i>Tax Sensitivity</i> × <i>HC ETF Intensity</i>			0.005*** (7.04)	0.004*** (6.37)
<i>Heartbeat Count Intensity Quintile</i>			0.064*** (11.98)	
Observations	208,252	208,252	208,252	136,289
Fixed effects	Year, quarter	Year, quarter	Year, quarter	Year, quarter
Adjusted $R^2$	0.029	0.093	0.191	0.110

*Notes.* Columns (1) and (2) present results of regressing proxies for heartbeat trade activity on *Tax Sensitivity*. The dependent variable, *Heartbeat Size Intensity*, is the dollar volume of heartbeat trade activity scaled by the dollar value of ETFs held by the filer in each quarter. We measure *Heartbeat Count Intensity* as the number of heartbeat trades executed among the ETFs held by the filer in each quarter scaled by the number of ETFs held by the filer. In column (3) and (4), the dependent variable is *HC ETF Swaps* as previously defined. We define *Heartbeat Count Intensity Quintile* as a count variable sorting each filer-quarter into a quintile of *Heartbeat Count Intensity*. Column (4) limits the sample to only filer-quarters that are below the median of *Heartbeat Count Intensity*. Heartbeat trades are identified using ETF Global Fund Flows data according to the procedure defined in Moussawi et al. (2025). Standard errors are clustered by filer and by year-quarter, and  $t$ -statistics are reported in parentheses.

\*\*\*Significance at 1%.

as clients. Prior literature exploits differential cutoffs of HNW individuals on Form ADV. Table 8 displays the robustness of our main results to alternative measures of *Tax Sensitivity*. Column (1) in Table 8 provides our base results from column (3) in Table 5.<sup>28</sup> In columns (2)–(4) in Table 8, we utilize different HNW individual cutoffs and re-estimate Equation (4). In column (2) in Table 8 (column (3) in Table 8), we define a tax-sensitive institution as one in which at least 25% (50%) of its clients are HNW individuals and a tax-insensitive institution as one that has zero HNW clients, a pension fund, or an endowment. In column (4) in Table 8, we classify a tax-sensitive institution as one where at least 50% of its clients are HNW individuals and a tax-insensitive institution as one where either (1) an investment advisor’s combined charity, government, and pension clients exceed 50% of their total clients or (2) the institution is a corporate pension fund, a public pension fund, or an endowment. In column (5) in Table 8, we utilize the tax sensitivity measure in Blouin et al. (2017). This constrains our sample size given that (1) WhaleWisdom does not share a primary identifier with the Thomson Reuters data used in Blouin et al. (2017), necessitating a matching procedure from which we keep only high-quality matches, and that (2) the

measure is only available through 2018.<sup>29</sup> In column (6) in Table 8, we apply each filer’s modal value of tax sensitivity from Blouin et al. (2017) to all of its observations, carrying its classification forward through 2022. Finally, we create continuous and time-varying measures of tax sensitivity using the proportion of clients who are individuals (individuals + HNW individuals) and the proportion of AUM that is attributable to individual clients in columns (7) and (8) in Table 8.

In each column in Table 8, we observe statistically significant evidence that the relation between ETF swapping and loss recognition is incrementally stronger for tax-sensitive filers. These tests indicate that our results are not dictated by how we measure tax sensitivity.

## 5. Conclusion

The United States has disallowed tax losses from wash sales for over 100 years. Prior literature has posited that wash sales have prevented investors from maximizing tax-efficient returns in their investment portfolios (e.g., Jensen and Marekwica 2011). In recent years, the expansion of ETFs has provided investors with a new, low-cost tool, whereby capital losses can be realized without disturbing an optimal portfolio. We contribute to the growing ETF literature and document trading

**Table 8.** Tax Sensitivity Robustness

Dependent variable = <i>PLR</i>	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Tax Sensitivity</i> definition	Simple cut	25% HNW	50% HNW	50/50	Blouin et al. (2017)	Blouin et al. (2017) (modal)	Continuous client	Continuous dollars
<i>Tax Sensitivity</i>					−0.553 (−1.23)		−1.702* (−1.87)	−0.340 (−0.33)
<i>HC ETF Swaps</i>	0.772** (2.48)	0.981*** (3.84)	1.394*** (5.38)	1.027** (2.17)	0.766*** (3.17)	0.809*** (3.83)	0.896*** (3.47)	1.144*** (4.81)
<i>Tax Sensitivity</i> × <i>HC ETF Swaps</i>	1.806*** (4.01)	1.801*** (4.54)	1.527*** (4.03)	1.342** (2.14)	1.406*** (2.68)	1.616*** (4.51)	2.060*** (4.70)	2.074*** (4.57)
<i>Unique HC ETFs</i>	−0.007 (−1.16)	−0.005 (−0.69)	−0.003 (−0.40)	−0.007 (−0.72)	0.031*** (2.70)	0.000 (0.05)	−0.007 (−1.18)	−0.008 (−1.23)
<i>Fourth Quarter</i>	3.759*** (8.01)	3.884*** (8.25)	3.840*** (8.29)	4.726*** (7.47)	3.764*** (7.74)	3.745*** (8.13)	3.760*** (8.00)	3.770*** (8.02)
<i>PGR</i>	−0.029** (−2.24)	−0.033** (−2.39)	−0.033** (−2.42)	−0.035** (−2.54)	−0.059*** (−4.02)	−0.057*** (−4.37)	−0.029** (−2.23)	−0.029** (−2.22)
<i>Turnover</i>	26.969*** (26.61)	27.027*** (24.66)	26.571*** (23.82)	39.465*** (30.77)	32.769*** (32.09)	30.439*** (32.83)	26.992*** (26.50)	26.987*** (26.52)
Observations	202,168	189,970	176,404	100,654	127,056	177,851	197,932	197,932
Fixed effects	Year, filer	Year, filer	Year, filer	Year, filer	Year, filer	Year, filer	Year, filer	Year, filer
Adjusted <i>R</i> <sup>2</sup>	0.421	0.418	0.416	0.417	0.446	0.422	0.423	0.423

*Notes.* Each column presents the test from column (3) in Table 5 with a different definition of *Tax Sensitivity*. The dependent variable in all specifications is *PLR* as previously defined. Column (1) uses the same definition as in all other tables. Columns (2) and (3) define tax sensitives as filers that report that HNW individuals comprise >25% or >50% of their clientele, respectively. Column (4) defines tax sensitives as filers with >50% HNW clients and tax insensitives as those with >50% charity, government, or pension clients. Column (5) uses the tax sensitivity measure from Blouin et al. (2017). Column (6) uses the modal value of the measure from Blouin et al. (2017) for each filer applied to all sample years. Column (7) is a continuous measure created by summing the percentage of clients who are individuals or HNW individuals. Column (8) is similar to column (7) but uses the percentage of AUM rather than percentage of clients. Fixed effects are noted near the bottom of each specification. Standard errors are clustered by filer and by year-quarter, and *t*-statistics are reported in parentheses.

\*Significance at 10%; \*\*significance at 5%; \*\*\*significance at 1%.

behavior consistent with ETF wash sale avoidance and tax loss harvesting. First, we find that the introduction of a near-identical ETF leads to more demand (as measured by share creation) for the incumbent ETF. Similar to Li (2024), we also find that incumbent ETFs have increased secondary market trading volume subsequent to the introduction of a near-identical ETF. Next, we document direct evidence of institutional investors swapping highly correlated ETFs in same-day, same-fund trades. Finally, we find that tax-sensitive institutions hold a more diverse set of highly correlated ETFs, invest a larger portion of their AUM in these ETFs, engage in more swapping between near-identical ETFs, and capture more capital losses with this swapping. We estimate conservatively that capital loss recognition attributable to annual swapping among tax-sensitive institutional investors is in the tens of billions of dollars.

It seems clear to us that the direct swapping of highly correlated ETFs for the sole purpose of harvesting tax losses violates the spirit of the original wash sale rule. But, even as this behavior has become increasingly widespread, regulators have remained silent on how the term “substantially identical securities” should be applied to ETFs. Absent such guidance, it is not surprising that a wide range of opinions on the matter has emerged among taxpayers and their advisors. Considering the rapid proliferation of nearly identical ETFs in recent years, we believe that more detailed guidance on the applicability of the wash sale rule to ETFs would help align potential enforcement actions with the original intent of the rule.

One possible explanation for the regulatory inaction is that policy setters have not yet fully grasped the scale and prevalence of the tax-loss-harvesting activities

associated with highly correlated ETFs. If this is the case, we hope that by providing a conservative estimate of the magnitude of these activities, our analyses will help elevate this issue and bring it forward for regulatory consideration.

It may be that the reason for the regulatory inaction is more complicated than that. Other factors may be at play. Perhaps legal ambiguity and the practical challenges of implementing and enforcing new ETF rules are prohibitively high. Or, perhaps resistance from institutional investors and high-net-worth individuals makes such actions politically infeasible. Perhaps even regulators’ personal trading incentives have a role to play. To refrain from speculation, we leave these questions to future research.

### Acknowledgments

The authors acknowledge Mary Cowx (discussant), Robert Feldgarden, Wentao Li, Shaphan Ng (discussant), Ben Yost (discussant), and seminar and workshop participants at the 2025 Financial Accounting and Reporting Section Midyear Meeting, the 2025 Hawaii Accounting Research Conference, the 2024 Conference on Financial Economics and Accounting, the 2024 University of Georgia Fall Accounting Symposium, Baruch College, Georgetown University, the University at Buffalo, the University of Illinois Chicago, the University of Washington, the Brigham Young University Accounting Research Symposium, The Stanford Initiative for Business, Taxation, and Society Lab, the PhD Accounting Workshop Series, two anonymous reviewers, an anonymous associate editor, and Shiva Rajgopal (editor) for helpful suggestions. They also thank Douglas Laporte and Christina Zhu for assistance with the ETF Global and Abel Noser data. Suhani Aggarwal, Luke Alexander, Isha Bhansali, Stefanie Luo, and Justin Sialm provided excellent research assistance.

## Appendix A. Abel Noser Direct Washing Evidence

**Table A.1.** Abel Noser Direct Washing Evidence

Client	Date	Symbol	Traded, \$	$\rho$	$\Delta$ Position, \$	Volume traded, \$	$\Delta$ Position/volume, %
732	12/21/2004	IVV	(3,615,739)	1.000	52,469	7,179,009	0.7
		SPY	3,563,270				
1,166	1/11/2007	SPY	(28,493,020)	0.990	3,494,151	53,491,889	6.5
		OEF	24,998,869				
1,191	3/19/2007	VII	(2,798,431)	0.999	30,345	5,566,517	0.5
		IWV	2,768,086				
1,015	9/5/2007	EEM	(10,461,004)	0.990	183,209	20,738,799	0.9
		VWO	10,277,795				
732	5/13/2008	XLE	(6,714,513)	0.996	386,662	13,815,688	2.8
		IYE	7,101,175				
1,325	11/18/2008	VWO	(10,595,292)	0.990	61,510	21,252,094	0.3
		EEM	10,656,802				
1,325	1/28/2009	IYR	(4,621,272)	0.991	607,229	8,635,315	7.0
		ICF	4,014,043				

*Notes.* A small selection of same day, same fund trades in highly correlated ETFs in the Abel Noser data are shown. The  $\Delta$  Position column is computed as the net change in exposure to the assets underlying the ETFs (i.e.,  $\text{abs}[\text{dollars sold} - \text{dollars bought}]$ ). The final column indicates the relative proportion of the change in position to the volume traded. EEM, iShares MSCI Emerging Markets ETF; ICF, iShares Select U.S. REIT ETF; IWV, iShares Russell 3000 ETF; IYR, iShares U.S. Real Estate ETF; IYE, iShares U.S. Energy ETF; OEF, iShares S&P 100 ETF; VWO, Vanguard Emerging Markets Stock Index Fund; XLE, State Street Energy Select Sector SPDR ETF.

## Appendix B. Variable Definitions

**Table B.1.** Variable Definitions

Variable	Description	Source
$\ln(\text{Volume})$	The natural log of ETF shares (in hundreds) traded in a month. This is at the ETF level and is found in Table A.1 and Table 3.	CRSP
$\ln(\text{SharesOut})$	The natural log of split-adjusted ETF shares outstanding at the end of the month.	CRSP
<i>HC ETF Peer</i>	An indicator equal to the value of $Treat \times Post$ in Equation (1).	CRSP Mutual Funds
<i>Spread</i>	Bid-ask scaled by nominal price.	CRSP
<i>Volatility</i>	The standard deviation of the prior 12 monthly returns.	CRSP
$\ln(\text{Market Value})$	The natural log of (shares outstanding $\times$ price).	CRSP
<i>B&amp;H Trailing 12</i>	The buy and hold return over the trailing 12 months.	CRSP
<i>Tax Sensitivity</i>	Indicator equal to one for 13F filers that serve a non-zero proportion of high-net-worth clients and zero otherwise. High-net-worth clients determined via question 5D of Form ADV. We also classify pensions and endowments as tax insensitive (indicator equal to 0) per the “Type” variable in the Institutional Investor Classification Data available on Brian Bushee’s website.	IAPD (SEC), Institutional Investor Classification Data
<i>Assets Under Management (AUM)</i>	Computed as $\Sigma(\text{Average of quarterly month-end prices for security } i \text{ from CRSP} \times \text{Number of shares of security } i \text{ held})$ for quarter $t$ for all 13F securities available in CRSP.	WhaleWisdom 13F, CRSP
<i>HC ETF Intensity</i>	$(\text{HC ETF AUM in quarter } t / \text{AUM in quarter } t)$ with HC ETF AUM computed following the AUM calculation above but limited to securities with CRSP share code 73, which have a highly correlated peer available in a given quarter.	WhaleWisdom 13F, CRSP
<i>Realized Loss</i>	Realized losses occur when $(\text{Sale Proceeds for security } i - \text{Basis for security } i) < 0$ . Sales Proceeds are computed as the $(\text{Number of shares of security } i \text{ sold} \times \text{Average of quarterly month-end prices for security } i \text{ from CRSP})$ . Realized losses are aggregated and reported at the filer-quarter level. Basis is computed using a highest in, first out methodology as in Blouin et al. (2017).	WhaleWisdom 13F, CRSP
<i>Unrealized Loss</i>	Unrealized losses occur when $(\text{Market Value for security } i - \text{Basis for security } i) < 0$ . Market Value for security $i$ is computed as $(\text{Average of quarterly month-end prices for security } i \text{ from CRSP} \times \text{Number of shares of security } i \text{ held})$ . Unrealized losses are aggregated and reported at the filer-quarter level. Basis is computed using a highest in, first out methodology as in Blouin et al. (2017).	WhaleWisdom 13F, CRSP
<i>Realized Gains</i>	Calculations are the same as for realized losses but for $(\text{Sale Proceeds for security } i - \text{Basis for security } i) > 0$ .	WhaleWisdom 13F, CRSP
<i>Unrealized Gains</i>	Calculations are the same as for unrealized losses but for $(\text{Market Value for security } i - \text{Basis for security } i) > 0$ .	WhaleWisdom 13F, CRSP
<i>Proportion of Losses Realized (PLR)</i>	<i>Realized Losses</i> in quarter $t$ scaled by $(\text{Realized Losses} + \text{Unrealized Losses})$ in quarter $t$ .	WhaleWisdom 13F, CRSP
<i>Proportion of Gains Realized (PGR)</i>	<i>Realized Gains</i> for quarter $t / (\text{Realized Gains} + \text{Unrealized Gains})$ in quarter $t$ .	WhaleWisdom 13F, CRSP
<i>HC ETF PLR (HC ETF PGR)</i>	Calculations are the same as for <i>PLR (PGR)</i> but for only HC ETF securities identified with CRSP share code 73, which have a highly correlated peer available in a given quarter.	WhaleWisdom 13F, CRSP

Table B.1. (Continued)

Variable	Description	Source
Volume, \$	Estimated dollar volume of total buys and sells for quarter $t$ for each filer. Total buys are computed as $\Sigma(\text{Average of quarterly month-end prices for security } i \text{ from CRSP} \times \text{Increase in the number of shares of security } i \text{ held in quarter } t \text{ compared with } t - 1)$ . Total sells are computed as $\Sigma(\text{Average of quarterly month-end prices for security } i \text{ from CRSP} \times \text{Decrease in the number of shares of security } i \text{ held in quarter } t \text{ compared with } t - 1)$ .	WhaleWisdom 13F, CRSP
Realized Loss (Gain)/Vol	Realized Losses (Gains) for quarter $t$ /Volume for quarter $t$ .	WhaleWisdom 13F, CRSP
Unrealized Loss (Gain)/Vol	Unrealized Losses (Gains) for quarter $t$ /Volume for quarter $t$ .	WhaleWisdom 13F, CRSP
Turnover	Volume for quarter $t$ /AUM for quarter $t$ .	WhaleWisdom 13F, CRSP
Unique HC ETFs	The count of distinct HC ETFs held by the 13F filer for quarter $t$ .	WhaleWisdom 13F, CRSP
Available HC ETFs	The count of distinct HC ETFs available in CRSP for each quarter $t$ .	CRSP
Fourth Quarter	Indicator equal to 1 for the 13F filings associated with the fourth quarter of the calendar year and zero otherwise.	WhaleWisdom 13F
HC ETF Swaps	The natural log of one plus the summation of offsetting buys and sells in highly correlated ETFs. Total Buys and Total Sells are computed the same as for the Volume variable but limited to HC ETFs. We record an HC ETF swap if there is both a Buy and a Sell of two different but highly correlated ETFs for quarter $t$ . The offset is defined as $\min[(\text{Buy of ETF } 1), (\text{Sell of ETF } 2)]$ . We consider only one Sell for each Buy. We limit our calculation to the offsetting quantity as that is the only portion of the trade that the wash sale rule would apply to.	WhaleWisdom 13F, CRSP
Heartbeat Size Intensity	Dollar volume of heartbeat trade activity (as defined in Moussawi et al. 2025) scaled by the dollar value of ETFs held by the filer in quarter $t$ .	ETF Global
Heartbeat Count Intensity	Count of heartbeat trades among ETFs held by the filer in quarter $t$ scaled by the number of ETFs held.	ETF Global
Heartbeat Count Intensity Quintile	Count variable formed by sorting each filer-quarter into a quintile of Heartbeat Count Intensity.	ETF Global

Notes. This table provides variable definitions and data sources. All nondiscrete variables are winsorized by year in our regression analyses at the 1% and 99% levels. All quarterly measures are as of the end of quarter 13F filing unless noted otherwise. Investment Adviser Public Disclosure (IAPD).

## Endnotes

<sup>1</sup> Fischer (2010) cites the U.S. Court of Appeals' opinion from *Hanlin v. Commissioner* that (in the context of the wash sale rule) "the words 'substantially identical' indicate that something less than precise correspondence will suffice to make the transaction a wash sale." For the remainder of the paper, we reserve the phrase "substantially identical" for discussion of the legal definition. We use "near identical" and "highly correlated" interchangeably to denote ETFs with price movements correlated at 0.99 or greater.

<sup>2</sup> Although the 99% threshold is arbitrary, it is consistent with what some brokerage firms have suggested for their clients (e.g., Wealthfront 2021) as well as other academic work (e.g., Matthews 2016, Brown et al. 2022, Li 2024).

<sup>3</sup> We focus on institutional investors rather than individual investors for several reasons. First, because institutional investors are required to distribute capital gains regularly, they have an immediate need to harvest tax losses. In contrast, individual investors can avoid capital gains via a step-up in basis at death (Dambra et al.

2020). Second, institutional investors can be partitioned by tax sensitivity, providing a natural control group when evaluating tax-motivated ETF hypotheses. Third, institutional investors comprise a significant majority of trading volume in the United States, allowing us to understand better how prevalent ETF-based swapping has become.

<sup>4</sup> The data are available at <https://accounting-faculty.wharton.upenn.edu/bushee/>. WRDS provides a CIK-mgrno linking table, which we use to match our WhaleWisdom data to the Institutional Investor Classification Data. We retain only matches flagged by WRDS as high quality.

<sup>5</sup> We also re-examine our main results using a variety of alternate definitions of tax sensitivity and find them to be robust (see Section 4.7).

<sup>6</sup> For instance, when we employ a lower threshold of 95% return correlation to identify ETF pairs, the magnitude of swapping since 2001 balloons to \$1.3 trillion (versus \$417 billion of swapping using a 99% return correlation).

<sup>7</sup> Prior legal precedence exists in support of a more lenient interpretation in non-ETF settings. For example, a legal expert referred us to a Supreme Court decision on the deductibility of a loss by a savings and loan's (S&L) sale of an interest in multiple single-family mortgages when they simultaneously exchanged it with an equivalent interest in a different package of single-family mortgages to nearby S&Ls (*Cottage Savings Ass'n v. Commissioner*, 499 U.S. 554 1991). The Supreme Court stated that an exchange of property generates a tax loss deduction only if the properties exchanged are "materially" or "essentially" different. The Supreme Court concluded that because the interests exchanged by Cottage Savings and the other S&Ls derived from loans made to different obligors and secured by different homes, the exchanged interests embodied distinct entitlements. The expert argued that Cottage Savings creates doubt that the IRS could successfully challenge the ETF loss harvesting that we document.

<sup>8</sup> The literature typically uses Thomson Reuters 13F data, which allow researchers to start estimating the cost basis of positions beginning in the 1980s. Given that Thomson Reuters is missing many ETFs (WRDS Research 2017), we use another data provider, WhaleWisdom, that offers superior coverage of ETFs. We compare reported holdings for a small number of filers in original SEC filings, and we find that the WhaleWisdom database reflects the SEC filings more accurately than the Thomson Reuters 13F data. Further, few highly correlated ETFs existed before 2001.

<sup>9</sup> Although the 13F holdings data provide useful insight into institutional investor equity holdings, they have limitations. First, we lack insight into intraquarter trading as holdings data are only available quarterly. Second, 13F data omit short positions, obscuring our insight into more complicated portfolio strategies. Finally, 13F data are aggregated at the investor (not fund) level, which induces measurement error into our classifications of which funds are tax sensitive and which trades are intrafund. Our results should be interpreted with these caveats in mind.

<sup>10</sup> Interestingly, tax-sensitive investors have lower proportions of losses and gains realized (as measured following Sikes 2014) than tax-insensitive investors. However, the fact that tax-sensitive investors have lower realized capital gains and capital losses is a direct implication of the differential trading volume between the two subsets of investors. Turnover for tax-insensitive investors nearly doubles that of tax-sensitive investors as the dollar volumes of buys and sells are 34% and 64% of total AUM, respectively.

<sup>11</sup> To illustrate, consider cohort construction for SPY and IVV, two ETFs that track the S&P 500. Their price movements are correlated at close to 1.00. SPY was introduced first (1993) and is the "incumbent." IVV was introduced in May of 2000. A cohort is constructed with May of 2000 set as zero in event time, and ETFs that do not gain a highly correlated partner in the 72 months around May of 2000 are included as clean controls. This forms the cohort. "Always treated" ETFs are excluded (e.g., IVV does not appear in the cohort because it has no preperiod; it is introduced in a treated state). This cohort-forming process is repeated for the first introduction of a highly correlated partner for any ETF with pre- and posttreatment data.

<sup>12</sup> Results are robust to including both "never treated" and "not yet treated" observations as controls (untabulated).

<sup>13</sup> We find qualitatively similar evidence in column (1) in Table 2 with cumulative fund flows from the ETF Global database as an alternative dependent variable (untabulated).

<sup>14</sup> In defining highly correlated ETFs, Brown et al. (2022) state: "Return correlations are calculated using daily returns over the trailing 12 months, and we use correlation thresholds of 95% and 99%."

<sup>15</sup> We selected examples for Table A.1 in Appendix A based on trade size. The dollar volume of the typical swap tends to be lower

and without perfectly offsetting purchase and sale amounts, consistent with Abel Noser's clients harvesting tax losses as part of a broader investment strategy.

<sup>16</sup> We can only ensure that trades occurred intrafund by limiting our investigation to single-day trades. This limitation prevents a time-series analysis, but it allows us to identify intrafund swap trades in highly correlated ETFs.

<sup>17</sup> For unmatched observations, we further match to the institution "type" field from Institutional Investor Classification Database available on Brian Bushee's website. Matched observations with "type" of pension or endowment are classified as tax insensitive per Blouin et al. (2017).

<sup>18</sup> We utilize high-net-worth individuals—who account for a significant portion of assets under management on Form ADV—to classify institutions as tax sensitive in our main specification (e.g., Jin 2006, Blouin et al. 2017). Our measure, however, uses a coarser threshold than some others (e.g., Jin 2006, Sikes 2014, Blouin et al. 2017). Our "greater-than-zero" threshold has the advantage of classifying the most filer-quarters. We rely on the assumption that although investment advisors serving few high-net-worth individuals may have small tax-avoidance incentives, they unambiguously have more tax-avoidance incentives than institutions serving zero high-net-worth individuals. We assume that tax sensitivity is a mostly static trait, consistent with prior literature (Blouin et al. 2017).

<sup>19</sup> In untabulated robustness tests, our results are quantitatively similar in Equations (2)–(5) when we deploy a less conservative 95% return correlation threshold for classifying a highly correlated ETF. In empirical analyses with filer-level fixed effects, the time-invariant variable *Tax Sensitivity* is not reported.

<sup>20</sup> Results are robust to using a scaled measure of swapping (e.g., *ETF Swap Vol/Total Vol*). We do not consider leveraged ETFs to be highly correlated pairs (e.g., an S&P 500 fund and an S&P 500 2X fund) because they represent a different risk profile even if correlated at  $\geq 99\%$ . Further, the CRSP mutual fund database does not treat leveraged ETFs as an index (i.e., not index fund flag B or D).

<sup>21</sup> Expanding our definition of highly correlated ETFs to include those whose returns are correlated at 95% or more increases the magnitude of swapping activity to \$1.3 trillion.

<sup>22</sup> For an excellent numerical illustration of estimating quarterly gains and losses with FIFO, see Blouin et al. (2017, appendix A). Because institutions can use FIFO, their positions in large liquid ETFs are rarely zeroed out. Indeed, among "sell" transactions involved in an HC ETF swap, we see only 7% of instances in which the quarterly position is fully liquidated (untabulated). Holding multiple tranches of multiple HC ETFs provides more potential harvesting opportunities when faced with return volatility.

<sup>23</sup> Note that the filer fixed effect omits the main *Tax Sensitivity* effect as the variable does not vary within filer.

<sup>24</sup> Using an alternative threshold of a  $\geq 95\%$  return correlation increases the magnitude of a one-standard-deviation increase to \$192 billion.

<sup>25</sup> *Turnover* is also highly correlated with *PGR*, which is included in column (3) in Table 5. We document a negative association between *PGR* and *PLR* in column (3) in Table 5. In untabulated tests, we confirm that the negative relationship is because of the inclusion of both *Turnover* and *PGR* in the regression. Excluding *Turnover* yields a positive relation between *PGR* and *PLR*.

<sup>26</sup> Approximately 18% of our filers never hold a highly correlated ETF, which constitutes 12% of our filer-quarter observations. In untabulated results, we find that our results in Tables 3–5 are robust to excluding these observations.

<sup>27</sup> Our results cannot be explained by investors' propensity to window dress. When investors historically harvested losses, the stock with negative recent returns (i.e., the harvested stock) would be

replaced by a different stock with positive recent returns (i.e., the window dressing). However, with highly correlated ETFs (i.e.,  $\geq 99\%$ ), the ETF with negative recent returns (i.e., the harvested ETF) is replaced by an ETF that has nearly identical negative recent returns. Thus, in our setting, ETF swapping does not allow an institution to window dress.

<sup>28</sup> In columns (1)–(4) and (6) in Table 8, the alternative measures of *Tax Sensitivity* do not vary over time, and thus, the *Tax Sensitivity* base coefficients are subsumed by filer fixed effects.

<sup>29</sup> We use the WRDS CIK-mgrno linking table (wrds\_13f\_link), which flags high-quality matches.

## References

- Agapova A (2011) Conventional mutual index funds versus exchange-traded funds. *J. Financial Markets* 14(2):323–343.
- Baker AC, Larcker DF, Wang CC (2022) How much should we trust staggered difference-in-differences estimates? *J. Financial Econom.* 144(2):370–395.
- Barber BM, Odean T (2004) Are individual investors tax savvy? Evidence from retail and discount brokerage accounts. *J. Public Econom.* 88(1–2):419–442.
- Ben-David I, Franzoni F, Moussawi R (2017) Exchange-traded funds. *Annual Rev. Financial Econom.* 9(1):169–189.
- Ben-David I, Franzoni F, Moussawi R (2018) Do ETFs increase volatility? *J. Finance* 73(6):2471–2535.
- Ben-David I, Franzoni F, Kim B, Moussawi R (2023) Competition for attention in the ETF space. *Rev. Financial Stud.* 36(3):987–1042.
- Bhattacharya U, Loos B, Meyer S, Hackethal A (2017) Abusing ETFs. *Rev. Finance* 21(3):1217–1250.
- BlackRock (2024) Tax loss harvesting at a glance. Accessed October 21, 2024, <https://www.blackrock.com/us/individual/literature/brochure/loss-harvesting-and-wash-sale-rule-considerations-us.pdf>.
- Blakey RG (1922) The Revenue Act of 1921. *Amer. Econom. Rev.* 12(1):75–108.
- Blouin JL, Bushee BJ, Sikes SA (2017) Measuring tax-sensitive institutional investor ownership. *Accounting Rev.* 92(6):49–76.
- Brown DC, Cederburg S, Towner M (2022) Dominated ETFs. Working paper, University of Arizona, Tucson.
- Brown DC, Davies SW, Ringgenberg MC (2021) ETF arbitrage, non-fundamental demand, and return predictability. *Rev. Finance* 25(4):937–972.
- Callaway B, Sant’Anna PH (2021) Difference-in-differences with multiple time periods. *J. Econom.* 225(2):200–230.
- Choi K-S, Li X, Yost B, Zou Y (2025) The ETF tax advantage and firm payout policy. Working paper, Boston College, Boston, MA.
- Colon JM (2023) Unplugging heartbeat trades and reforming the taxation of ETFs. *Univ. Chicago Bus. Law Rev.* 2(1):53–136.
- Constantinides GM (1983) Capital market equilibrium with personal tax. *Econometrica* 51(3):611–636.
- Dambra M, Gustafson M, Quinn PJ (2020) Tax-advantaged trust use among IPO executives: Determinants and implications for valuation and future performance. *Accounting Rev.* 95(3):145–175.
- Dambra M, Mihov A, Sanz L (2024) Unintended real effects of EDGAR: Evidence from corporate innovation. *Accounting Rev.* 99(6):75–99.
- Dong T (2022) Tax-loss harvesting using ETFs: A guide. *ETF Central* (June 20), <https://www.etfcentral.com/news/tax-loss-harvesting-using-etfs-guide>.
- Fischer GM (2010) New twists on an old plot: Investors look to avoid the wash sale rule by harvesting tax losses with exchange-traded funds note. *Washington Univ. Law Rev.* 88(1):229–260.
- Ganel O (2024) Can you avoid the wash-sale rule with ETFs and this loophole? *WealthTender* (May 19), <https://wealthtender.com/insights/investing/can-you-avoid-the-wash-sale-rule-with-etfs-and-this-loophole/>.
- Gao M, Huang J (2020) Informing the market: The effect of modern information technologies on information production. *Rev. Financial Stud.* 33(4):1367–1411.
- Goodman B (2015) Maximize tax losses with ETFs. *Barron’s* (October 31), <https://www.barrons.com/articles/maximize-tax-losses-with-etfs-1446267602>.
- Gormley TA, Matsa DA (2011) Growing out of trouble? Corporate responses to liability risk. *Rev. Financial Stud.* 24(8):2781–2821.
- Grinblatt M, Keloharju M (2004) Tax-loss trading and wash sales. *J. Financial Econom.* 71(1):51–76.
- Holt J (2023) Tax-loss harvesting: What does it mean to be substantially identical? *BigLawInvestor* (April 3), <https://www.biglawinvestor.com/tax-loss-harvesting-substantially-identical/>.
- Hu G, Jo KM, Wang YA, Xie J (2018) Institutional trading and Abel Noser data. *J. Corporate Finance* 52 (October 2018):143–167.
- Huang JC, Guedj I (2009) Are ETFs replacing index mutual funds? Preprint, submitted March 15, <http://dx.doi.org/10.2139/ssrn.1108728>.
- Internal Revenue Service (1983) No loss allowed on stock sale and repurchase intended solely to obtain a tax loss. Accessed June 24, 2024, <https://www.taxnotes.com/research/federal/court-documents/court-opinions-and-orders/no-loss-allowed-on-stock-sale-and-repurchase-intended-solely/1p3zr7>.
- Internal Revenue Service (1985) Loss on corporate stock sale is not allowed if “in-the-money” put option is sold within 30 days and is likely to be exercised. Accessed June 24, 2024, <https://www.taxnotes.com/research/federal/irs-guidance/revenue-rulings/loss-on-corporate-stock-sale-is-not-allowed-if-in/dbsv>.
- Internal Revenue Service (2017) Internal Revenue Service national office technical advice memorandum. Accessed June 24, 2024, <https://www.irs.gov/pub/irs-wd/201737011.pdf>.
- Internal Revenue Service (2021) About Publication 550, investment income and expenses. Accessed March 16, 2023, <https://www.irs.gov/forms-pubs/about-publication-550>.
- Jennings WW, O’Malley TC, Payne BC (2020) Normal return gaps: Dispersion illuminates diversification. *J. Wealth Management* 23(2):18–35.
- Jensen B, Marekwica M (2011) Optimal portfolio choice with wash sale constraints. *J. Econom. Dynam. Control* 35(11):1916–1937.
- Jin L (2006) Capital gains tax overhang and price pressure. *J. Finance* 61(3):1399–1431.
- Khomyn M, Putnins T, Zoican M (2024) The value of ETF liquidity. *Rev. Financial Stud.* 37(10):3092–3148.
- Lasser JK (2011) Tax rules for ETF losses. Accessed October 21, 2024, <https://www.fidelity.com/learning-center/investment-products/etf/tax-rules-for-losses-etfs>.
- Li W (2024) Beating the index with ETFs. Preprint, submitted February 15, <https://ssrn.com/abstract=4595934>.
- Liebi LJ (2020) The effect of ETFs on financial markets: A literature review. *Financial Markets Portfolio Management* 34(2):165–178.
- Lundholm R, Zheng X (2025) The changing nature of financial analysis in the presence of ETFs. *Rev. Accounting Stud.* 30(1):860–898.
- Malkiel BG (2023) Celebrating 10 years of tax-loss harvesting. *Wealthfront* (April 14), <https://www.wealthfront.com/blog/10-years-of-tax-loss-harvesting/>.
- Matthews DW (2016) Tax loss harvesting: Can robo-advisers navigate wash sale rule? *Tax Notes* (December 12), <https://www.taxnotes.com/tax-notes-federal/financial-instruments/tax-loss-harvesting-can-robo-advisers-navigate-wash-sale-rule/2016/12/12/r3bw>.
- Meziani AS, Yang JGS (2012) Assessing the value of tax efficient rebalancing using ETFs: Is it always better than a tax deferred strategy? *Internat. Res. J. Appl. Finance* 3(9):1274–1284.
- Moussawi R, Shen K, Velthuis R (2025) The role of taxes in the rise of ETFs. Preprint, submitted April 28, <http://dx.doi.org/10.2139/ssrn.3744519>.

- Nessa M, Schwab CM, Stomberg B, Towery EM (2020) How do IRS resources affect the corporate audit process? *Accounting Rev.* 95(2):311–338.
- Odean T (1998) Are investors reluctant to realize their losses? *J. Finance* 53(5):1775–1798.
- Puckett A, Yan XS (2011) The interim trading skills of institutional investors. *J. Finance* 66(2):601–633.
- Sialm C, Starks L (2012) Mutual fund tax clienteles. *J. Finance* 67(4):1397–1422.
- Sikes SA (2014) The turn-of-the-year effect and tax-loss-selling by institutional investors. *J. Accounting Econom.* 57(1): 22–42.
- Silow ML (2008) Recent ruling revisits wash sale rules in internal revenue code. *Legal Intelligencer* (January 8).
- Statista (2023) Development of assets of global exchange traded funds (ETFs) from 2003 to 2022. Accessed March 16, 2023, <https://www.statista.com/statistics/224579/worldwide-etf-assets-under-management-since-1997/>.
- U.S. Senate (1921) Report 275 67th Session of Congress. Senate Finance Committee.
- Wall Street Journal (1920) Establishing loss on stocks. (November 19).
- Wealthfront (2021) Wealthfront tax-loss harvesting. Accessed March 15, 2023, <https://research.wealthfront.com/whitepapers/tax-loss-harvesting/>.
- WRDS Research (2017) Research note regarding Thomson Reuters ownership data issues. Accessed May 22, 2023, [https://wrds-www.wharton.upenn.edu/documents/533/Research\\_Note\\_-Thomson\\_S34\\_Data\\_Issues.pdf](https://wrds-www.wharton.upenn.edu/documents/533/Research_Note_-Thomson_S34_Data_Issues.pdf).