

Sparse Portfolios and Benchmarking in Corporate Bond Markets

November 19, 2025

Abstract

We use detailed data on fixed income benchmark indexes in Canada and the United States to demonstrate the importance of benchmarks for corporate bond fund ownership and bond prices. Funds hold sparse portfolios, and index weights strongly predict active and passive fund ownership. We rationalize these patterns in a model with benchmarked managers who face portfolio management costs, which characterizes the assets managers optimally include in their portfolios. In the model, a bond's price increases with its *sparse benchmarking intensity (BMI)*—which captures its cumulative weight in all benchmarks and portfolio sparsity. Exploiting within-bond changes in sparse BMI driven by reallocations of benchmarked capital, we find that an increase in sparse BMI lowers bond yields and increases institutional ownership.

JEL Classification: G11, G12, G23

Keywords: Benchmark, corporate bonds, sampling, index rebalancing, mutual funds, ETFs

1 Introduction

An important trend in the corporate bond market is the increasing presence of mutual funds and exchange-traded funds (ETFs). For example, in the United States, mutual funds and ETFs have increasingly replaced households as significant domestic holders of corporate bonds (see Appendix Figure A1). Academics and policymakers have raised concerns that this may contribute to increased financial fragility in bond markets (see, e.g., Cai, Helmke, Mosk, and Suntheim (2025)), underscoring the importance of understanding the behavior of these investment vehicles. It is common in the literature to model the asset allocation decisions of fixed-income funds using a standard portfolio optimization framework. Yet, funds face unique incentives and constraints, influencing both their portfolio composition and, ultimately, asset prices. In this paper, we highlight two key aspects of fixed-income portfolio management. First, fund performance is typically evaluated relative to a benchmark index, directly influencing portfolio choices made by fund managers. Second, due to the large size of the fixed-income investment universe—generally much larger than that of equities—it is impractical for funds to hold every available bond. As a result, funds typically construct sparse portfolios comprising only a limited subset of bonds, a practice known as sampling.

Our research question is to assess the importance of benchmarks for investment grade corporate bond fund ownership and bond prices. Using data covering most fixed income benchmarks in Canada and the United States, we document that fund managers hold sparse portfolios and that benchmark weights strongly predict bond ownership of both active and passive funds. We rationalize these findings in a model in which benchmarked managers face portfolio management costs proportional to portfolio size. When adding an asset, a manager balances the marginal cost of managing a larger portfolio against the benefit of reducing tracking error variance and increasing expected return. The model delivers a measure of the inelastic demand a bond attracts from fund managers, which we term sparse benchmarking intensity (sparse BMI). Sparse BMI is calculated as a bond’s cumulative weight in all benchmarks, weighted by assets following each benchmark and divided by the bond’s market value, adjusted for sparsity. Exploiting within-bond changes in sparse BMI driven by reallocations

of benchmarked capital, such as those around bond maturity cutoffs, we provide empirical evidence consistent with the model’s predictions: higher sparse BMI leads to reductions in bond yields and increases in fund ownership.

We first present stylized facts on fixed income benchmark indexes for the United States and Canada adopted by mutual funds and ETFs. We find that in both countries, there is an increasing heterogeneity of benchmarks, catering to different investment mandates and styles. In the United States, most fixed-income assets are benchmarked to Bloomberg indexes, while Canadian funds primarily use FTSE benchmarks. To our knowledge, ours is the first study to systematically describe benchmarks in these markets. In contrast to equity benchmarks, bond benchmarks so far have received limited attention in the literature. We document that they generally contain far more assets than equity benchmarks and exhibit significantly higher turnover. The two-way turnover is over 40% annually for the Bloomberg U.S. Aggregate Bond Index compared to approximately 10% for the S&P 500. This high turnover imposes considerable rebalancing costs, making it prohibitively expensive for fund managers to hold every bond in their benchmark index. Despite the high turnover levels, fund managers steer performance of their funds close to their benchmark indexes, as evidence by their tracking error levels.

Our next set of stylized facts concerns fund portfolios. We document that these portfolios are highly sparse: U.S. fixed income funds hold only 7% of the bonds included in their benchmarks, whereas Canadian funds hold approximately 15%. Even passive funds and ETFs hold only a fraction of the assets in their benchmarks—up to 26% in the U.S. and 46% in Canada. Canadian corporate bond market is much narrower, yet there is still significant sparsity. These figures imply a level of sparsity substantially greater than what is typically observed in equity portfolios. Additionally, the sparsity of fund holdings is highly persistent.

We then show that a bond’s benchmark membership and weight strongly predict its likelihood of inclusion in portfolios, underscoring the central role of benchmarks in portfolio selection. This pattern is not confined to passive vehicles but holds across both active

and passive funds. Both types of funds tend to hold more bonds with higher benchmark weights. However, since managers typically seek exposure across all relevant bond categories (e.g., maturity and credit rating), they hold at least some bonds from the entire range of benchmark weights. We document that even within narrowly defined categories, there is sparsity.

In the second part of the paper, we propose a model that rationalizes these new stylized facts and derives testable asset pricing implications in an equilibrium with sparse portfolios. The model features fund managers who are benchmarked to an index and face portfolio management costs, which we interpret as rebalancing, monitoring, and other portfolio management costs. These costs increase with the number of assets included in the portfolio, preventing managers from holding all benchmark constituents and resulting in sparse portfolios. When adding an asset, a manager trades off the marginal cost of managing a larger portfolio against the benefit of reducing tracking error variance and increasing expected return. Accordingly, the manager excludes assets with low benchmark weights, given their small impact on tracking error.

The model highlights asset substitutability as another determinant of which assets are excluded. Specifically, managers are more likely to exclude assets for which there are many close substitutes. We find that, within each category of closely substitutable assets, managers are likely to hold ones with higher benchmark weights, larger market cap, and higher idiosyncratic risk.

When deriving optimal portfolios, we demonstrate that a portion of fund managers' demand is inelastic—even among active managers. This matters for equilibrium asset prices. Aggregating the inelastic demand across all managers yields a measure we define as *sparse benchmarking intensity* (sparse BMI). Sparse BMI thus captures the total inelastic demand at the asset level. In equilibrium, assets with higher sparse BMIs command higher prices and offer lower expected returns. Our model explicitly accounts for portfolio sparsity in deriving sparse BMI—this feature is particularly relevant to our corporate bond market application. For example, bonds with low benchmark weights are predicted to have low or zero sparse

BMI, and mutual fund flows have little impact on their prices.

In the third part of the paper, we empirically test these predictions in Canada and the United States. We first construct a measure of sparse BMI. Consistent with our theoretical framework, the measure includes both passive and active assets under management (AUM). In our sample, active AUM substantially exceeds passive AUM—particularly in Canada, where it accounts for nearly all AUM. A bond’s total BMI is calculated as its cumulative weight in all benchmarks, weighted by assets following each benchmark, and divided by the market value of the bond. To compute the bond’s sparse BMI, we add a sparsity correction that excludes benchmarks in which the bond’s index weight rank is not high enough. We consider several variants of the sparse BMI measure, each applying such a sparsity threshold but to a different benchmark segment. For instance, the threshold may be defined at the benchmark level (e.g., top-half index weights) or within benchmark rating buckets (e.g., top-half index weights within each rating bucket in the benchmark).

Importantly, as time passes, bonds roll down the maturity spectrum and transition from long- to intermediate- and short-term market indexes. Because these indexes contain different amounts of benchmarked capital, this roll-down produces sizeable changes in sparse BMI. However, these transitions are mechanically determined by a bond’s amortization schedule and are therefore predictable. The identifying variation in our empirical design instead comes primarily from reallocations of benchmarked capital across maturity segments, captured by changes in assets following each benchmark relative to its market value. We implement a panel regression framework that relates such plausibly exogenous changes in sparse BMI to subsequent changes in bond prices and fund ownership. Consistent with the model, we find that increases in sparse BMI are followed by declines in bond yields and credit spreads in both Canada and the United States.

The data on bond ownership lend further support to our mechanism. First, we document that long-term funds reduce their holdings of bonds crossing relevant maturity cutoffs, while short-term funds increase theirs. Second, and importantly, this is true both for active and passive funds. Whereas the result is somewhat expected for passive funds

and ETFs, the literature to date has not provided strong evidence that active fixed income funds buy additions and sell deletions to their benchmark indexes. Third, consistent with our model, our sparse BMI measure predicts changes in bond ownership for both active and passive funds.

Related Literature. Our paper contributes to several strands of literature. Our model is related to the literature on benchmarking, mandates, and habitats. Like us, [Brennan \(1993\)](#), [Basak and Pavlova \(2013\)](#), [He and Xiong \(2013\)](#), [Buffa, Vayanos, and Woolley \(2022\)](#), and [Pavlova and Sikorskaya \(2023\)](#) study the implications of benchmarking for portfolio management and asset prices. However, none of these models produces portfolio sparsity and explores how it affects asset prices. Furthermore, our model is related to the literature on portfolio underdiversification, as sparsity may result from underdiversification. Previous research proposes various mechanisms to explain underdiversification, such as ambiguity aversion ([Dow and Werlang, 1992](#)) or limited information-processing capacity ([Van Nieuwerburgh and Veldkamp, 2010](#)). Our mechanism is distinct. The paper closest to ours in this strand is [Merton \(1987\)](#), which assumes that investors exogenously neglect certain stocks and studies the resulting implications for asset prices. In contrast, our contribution is to endogenously determine which assets managers exclude from their portfolios and analyze how benchmarking considerations influence this decision.

There are several empirical papers that discuss sparsity in holdings. [Kojien and Yogo \(2019\)](#) introduce a demand system approach to asset pricing and discuss how it should be adjusted to handle sparse portfolios. [Graves \(2024\)](#) investigates the origins of portfolio sparsity of institutional investors and attributes it to binding short-sale constraints on assets they choose not to hold. The mechanism generating sparsity that is closest to ours is discussed in the ETFs context by [Shim and Todorov \(2022\)](#), [Brogaard, Heath, and Huang \(2024\)](#), [Koont, Ma, Pastor, and Zeng \(2025\)](#), who note that ETF creation/redemption baskets exclude some (illiquid) assets and attribute this to a tradeoff between transaction costs and tracking error. They argue that basket choices may contribute to financial fragility. We study fund portfolios and argue that benchmarking and portfolio management costs affect the allocations of

both passive and active bond fund managers.

More generally, our work connects to the emerging literature on institutional demand in the corporate bond market. [Dick-Nielsen and Rossi \(2019\)](#), [Bretscher, Schmid, Sen, and Sharma \(2023\)](#), [Bretscher, Schmid, and Ye \(2024\)](#), [Chaudhary, Fu, and Li \(2024\)](#), [Gabaix, Koijen, Richmond, and Yogo \(2025\)](#) shed light on the nature of institutional investors' portfolios in this market, but do not discuss the importance of fixed income benchmarks in asset allocation. One exception is [Ottonello \(2019\)](#), whose focus is on high-yield bond prices and funds' liquidity management.¹ We document results consistent with Ottonello's findings but utilize a fundamentally different identification strategy based on sparse BMI of investment-grade bonds. Several papers in this literature examine corporate bond mutual funds and ETFs. [Dannhauser and Hoseinzade \(2022\)](#), and [Ma, Xiao, and Zeng \(2022\)](#) examine active choices of these investors and explore how these choices may contribute to financial fragility. The closest paper to ours in this literature is [Bretscher, Schmid, and Ye \(2024\)](#) that attributes bond price changes around index maturity cutoffs to inelastic demand of passive funds. In contrast, we argue that active funds contribute to such demand as well. Supporting this claim, we find ownership changes for both passive and active funds around maturity cutoffs in line with their benchmark indexes. Additionally, instead of using a maturity-cutoff dummy as an instrument for fund ownership, we propose to use sparse BMI. It is a continuous measure with more variation than the dummy. Our measure offers a way to quantify inelastic demand in bond markets stemming from fund manager mandates.

2 Stylized Facts

In this section we document new stylized facts about corporate bond fund portfolios and benchmark indexes in Canada and the United States.

¹Recent evidence in market microstructure highlights the growing importance of index-tracking trades by mutual funds and ETFs in the corporate bond market (see, e.g., [Li, O'Hara, Rapp, and Zhou, 2025](#); [Shin, Zhou, and Zhu, 2025](#)).

2.1 Corporate bond investment management

Mutual funds and exchange-traded funds (ETFs) have become increasingly important holders of corporate bonds, particularly in the investment-grade (IG) segment. Their share of outstanding corporate debt has risen steadily over the past two decades (see Appendix Figure A1), making these funds key intermediaries in the transmission of demand shocks and in the provision of secondary market liquidity. Fund managers in this sector almost universally operate against benchmark indexes, which serve two purposes: they define the investment universe for eligible holdings and provide the yardstick for performance evaluation (Ma, Tang, and Gomez (2019)).

However, as we demonstrate in Section 2.3, the corporate bond benchmarks themselves are extremely broad, often spanning thousands of individual securities across maturities, sectors, and ratings buckets. Fully replicating these indexes is operationally infeasible and would impose prohibitively high portfolio management costs. Consequently, managers face a trade-off between tracking error and operational complexity: they hold relatively sparse subsets of benchmark constituents while seeking to remain close to benchmark performance. This trade-off is saliently illustrated in Bloomberg’s portfolio allocation tools, snapshots of which we report in Appendix A.2. These tools highlight how portfolio managers balance tracking error targets with the desire to minimize portfolio management costs, resulting in sparse but benchmark-oriented bond portfolios.

To operationalize this sparsity, many fixed-income index fund managers employ stratified sampling techniques—selecting a representative subset of benchmark securities that aligns with the portfolio’s key risk factors such as duration, credit quality, sector weights, and convexity—while keeping portfolio management costs manageable. For example, Vanguard describes this method in their white papers and fund materials, emphasizing that their bond index funds use stratified sampling strategies to approximate index exposure without the need to hold every constituent security.²

²See Vanguard research note https://corporate.vanguard.com/content/dam/corp/research/pdf/a_bond_index_funds_balancing_act_tracking_error_and_cost.pdf.

2.2 Data

We construct our dataset from multiple sources. For Canada, we obtain benchmark constituent weights for all Canada Domestic Fixed Income Indexes from FTSE Russell (London Stock Exchange Group) from July 2003 to June 2024. The FTSE data also include daily bond characteristics and pricing data as well as daily index returns. We complement the bond-level data using Refinitiv to further include variables such as Return Index and bid-ask spread. We use Bank of Canada’s bond yield curves to construct maturity-matched credit spreads.³ We collect mutual fund and ETF holdings and fund static and dynamic characteristics from Morningstar, including a snapshot of fund primary prospectus benchmarks and fund returns. Because we observe all data on Canadian mutual funds, ETFs, and separate accounts in Morningstar, we treat all these fund structures as separate funds as long as they have different portfolio identifiers. We use USD-CAD exchange rates from FRED to align the fund asset currency (USD) with bond market value currency (CAD). Finally, we classify funds into active and passive using Morningstar’s ‘index fund’ and ‘investment type’ fields. Table A1 in the Appendix characterizes our sample of Canadian funds.

For the U.S., we obtain benchmark constituent weights for all Fixed Income Indexes from Bloomberg Aggregate index family as well as the bond characteristics and pricing data in 2012–2023. We get additional bond-level variables, including bid-ask spread and issue date, from WRDS bond return database. The mutual fund and ETF holdings as well as their static and dynamic characteristics come from CRSP Mutual Fund database. We augment the assets of funds in our sample by the size of their matched separate accounts reported in Morningstar, as those separate accounts hold virtually identical portfolios and significantly contribute to the AUM (Huang, Lu, Song, and Xiang (2023)). We use CRSP to classify funds into active and passive, as described in Appendix A.8.2. We further extract a dynamic panel of each funds’ primary prospectus benchmark over time directly from regulatory filings reported on U.S. SEC EDGAR and augmented with a Morningstar snapshot and U.S. SEC Mutual Fund Risk and Return database. More details and validation exercises on this

³See <https://www.bankofcanada.ca/rates/interest-rates/bond-yield-curves/>.

are reported in Appendix [A.8](#). Importantly, by tracking fund benchmark histories and benchmark composition histories, we are able to measure shifts in benchmarking intensity over time. Table [A2](#) in the Appendix characterizes our sample of U.S. funds.

In Canada, we only consider corporate bonds by selecting ‘corporate’ FTSE industry sector. Within that sector, there are 7 industry groups (communication, energy, financial, industrial, infrastructure, real estate, and securitization) and 21 finer industry subgroups. In the U.S., we filter for corporate bonds by setting Bloomberg’s `bclass1` to ‘corporates.’ Within this class, there are 3 industry groups (financial institutions, industrial, and utility) that include 18 finer industry subgroups and 58 sectors. In our sparse BMI construction, we refer to industry groups as segments, industry subgroups as industries, and industry sectors as sectors.

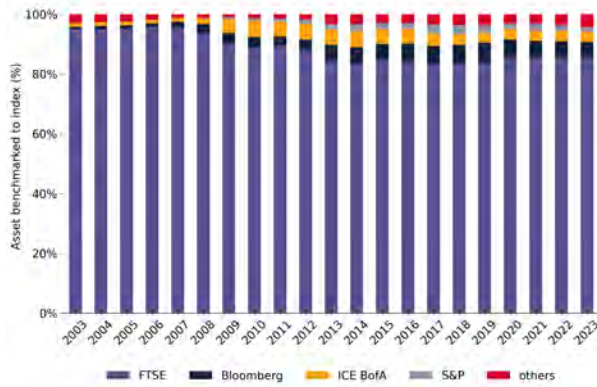
We report the descriptive statistics of the main variables used in analysis in Tables [A3](#) and [A4](#) in the Appendix.

2.3 Stylized facts

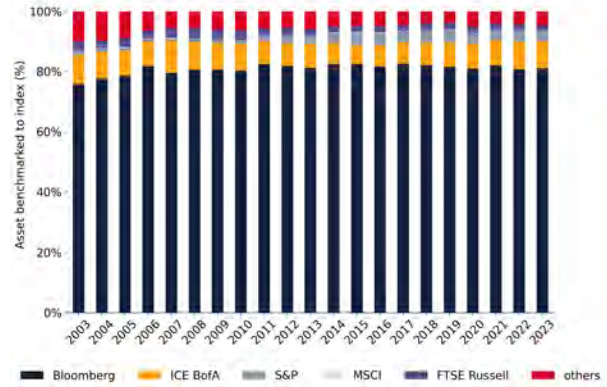
In this section, we document new stylized facts on corporate bond benchmarks and funds portfolios in Canada and the United States.

Among the universe of fixed income benchmarks adopted by mutual funds and ETFs in Canada, over 90% of total assets is benchmarked to FTSE indexes (Figure [1a](#)). Therefore, our detailed benchmark data from FTSE covers most of the assets of those institutional investors. Importantly, within FTSE Canada bond index family, there is an entire suite of indexes—such as, for example, short-term, mid-term, long-term, and aggregate—catering to different investment objectives. The assets under management benchmarked to each of them are shown in Figure [1c](#). In the US, on the other hand, FTSE Russell is very small, and the market is dominated by Bloomberg indexes. Despite slightly more heterogeneity than demonstrated in Canada, Bloomberg still accounts for the dominating share of more than 80% of the total assets (Figure [1b](#)). Within Bloomberg indexes, there are also indexes covering narrower market segments, such as short-term credit and 5–10Y corporate bonds.

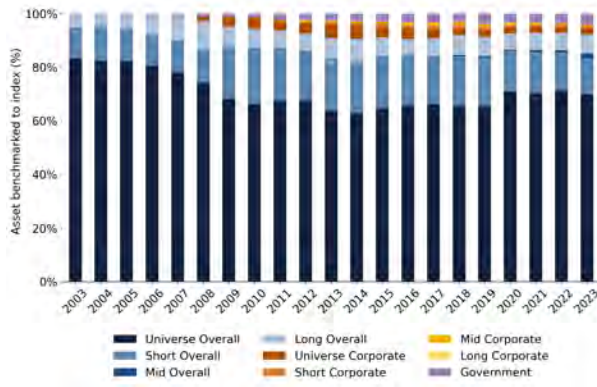
Figure 1: Assets benchmarked to indexes



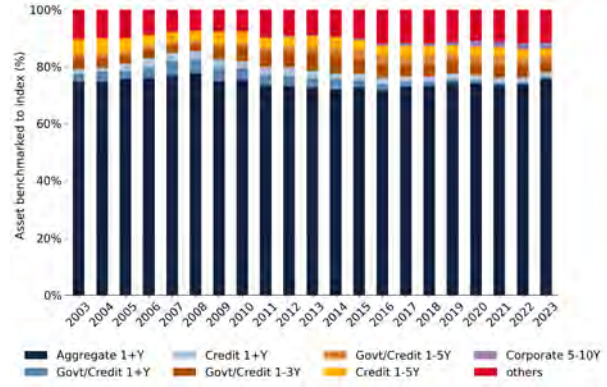
(a) Assets benchmarked to index providers (Canada)



(b) Assets benchmarked to index providers (US)



(c) Assets benchmarked to FTSE indexes (Canada)



(d) Assets benchmarked to Bloomberg indexes (US)

This figure illustrates the evolution of the share of benchmark groups in the total asset under management of Canada and US domestic fixed income mutual funds and ETFs.

Our first stylized fact concerns index turnover: fixed income indexes typically have significantly higher turnover than equity indexes. Bonds issuance and maturity create mechanical changes in bond market indexes, especially for the shorter segments of the market. This results in an average two-way turnover of 46% (71%), on average, for the Canadian (US) bond indexes in our sample, which is a magnitude difference from the turnover of around 10% for S&P 500 index, as shown in Table 1. We argue that this high turnover significantly increases portfolio management costs for corporate bond portfolios.

Despite the high turnover levels, fund managers in our sample manage to steer performance of their funds close to their benchmark indexes. The average annualized tracking

Table 1: Annual Turnover of FTSE Canada and Bloomberg US Fixed Income Indexes

Benchmark	Number of bond members	Additions	Deletions	Turnover rate (%)
Panel A: Canada FTSE Indexes				
Universe Overall Bond Index	1271	171	131	24.13
Short Overall Bond Index	488	141	126	55.01
Mid Overall Bond Index	306	113	104	71.33
Long Overall Bond Index	477	67	49	26.00
Universe Corporate Bond Index	809	126	96	28.23
Short Corporate Bond Index	348	103	92	56.45
Mid Corporate Bond Index	184	79	73	83.38
Long Corporate Bond Index	277	41	28	27.78
Universe Corporate BBB Bond Index	319	66	46	38.40
Short Mid Corporate BBB Bond Index	229	56	41	45.88
<i>Average</i>	471	96	79	45.66
Panel B: US Bloomberg indexes				
Aggregate	9430	2292	1893	45.55
Aggregate 1-3 year	2459	1694	1611	134.28
Aggregate 1-5 year	4468	1908	1766	82.58
Aggregate 3-5 year	2010	1534	1475	150.14
Aggregate 1-10 year	7513	2067	1804	52.32
Aggregate 1-10 year Ex-BAA	5706	1757	1581	59.78
Corporate	4828	960	686	34.41
Corporate 1-3 year	893	523	474	111.40
Corporate 1-5 year	1758	619	528	64.81
Corporate 1-10 year	3283	763	607	41.57
Corporate 5-10 year	1526	476	412	58.13
Corporate 10+ year	1544	320	202	36.14
Credit	5457	1084	782	34.61
Credit 1-3 year	1035	602	549	111.04
Credit 1-5 year	2009	705	606	64.94
Credit 1-10 year	3667	859	690	42.07
Credit 5-10 year	1658	521	451	58.61
Credit 10+ year	1791	362	228	35.49
Govt/Credit	6298	1534	1232	45.83
Govt/Credit 1-3 year	1411	943	887	131.85
Govt/Credit 1-3 year Ex-BAA	996	722	692	143.17
Govt/Credit 1-5 year	2591	1117	1016	84.12
Govt/Credit 1-10 year	4405	1299	1130	56.52
Govt/Credit 1-10 year A+	2624	949	873	70.08
Govt/Credit 5-10 year	1814	606	539	63.80
Govt/Credit 10+ year	1893	383	250	36.22
<i>Average</i>	3195	1023	883	71.13

This table reports number of bond additions and deletions from Bloomberg fixed income indexes and the turnover rates. The Bloomberg US Aggregate indexes are rebalanced at monthly frequency, and the FTSE Canada indexes are rebalanced at daily frequency. The annual number of addition and deletions and turnovers reported are calculated by aggregating all the rebalances over a year. The numbers are calculated for the full sample of benchmark index data, which is 2005 to 2023 in the U.S., and 2004 to 2023 in Canada.

error of passive funds in our Canadian (US) sample is well below 1% at 59 bps (92 bps). Even for active funds, tracking errors are well below their equity fund counterparts at 116 bps (207 bps) for Canada (US). Within all types of funds, there is considerable cross-sectional variation, with standard deviations of tracking error always exceeding the averages we report here. Details of our tracking error calculation as well as additional statistics are reported in

Table 2: Transition matrix of bond fund holdings

	1 Quarter		2 Quarters		1 Year		3 Years	
	Not held	Held	Not held	Held	Not held	Held	Not held	Held
Panel A: Canada								
Not held in last period	98.0%	2.0%	96.4%	3.6%	94.0%	6.0%	89.8%	10.2%
Held in last period	7.2%	92.8%	13.0%	87.0%	21.5%	78.5%	38.7%	61.3%
Panel B: US								
Not held in last period	99.4%	0.6%	98.9%	1.1%	98.2%	1.8%	96.6%	3.4%
Held in last period	8.0%	92.0%	13.7%	86.3%	22.7%	77.3%	45.8%	54.2%

This table reports transition probabilities for corporate bond holdings of mutual funds and ETFs (Canada and U.S.). Rows are conditioned on last period’s holding status; columns show the status after the indicated horizon. The reference universe is the broadest benchmark available in each geography, i.e., Bloomberg Aggregate index in the U.S. and Universe Overall Bond Index in Canada.

Appendix A.5.

Second, corporate bond fund portfolios are extremely sparse. Canadian funds on average hold only 15% of bonds in their benchmark indexes. Passive bond funds tend to hold more bonds, with an average of 46%, but still fewer than passive equity funds would. Active bond funds, on the other hand, hold very concentrated portfolios with only 10% of the benchmark bonds. As Table 1 illustrates, the corporate bond universe in the U.S. is a magnitude larger than Canada, so we see even more sparsity. US funds on average hold only 7% of bonds in their benchmark indexes. Passive funds hold 26%, while active funds hold only around 4% of benchmark bonds.

This sparsity is also very persistent. As shown in the transition matrix in Table 2, an asset that is not held by a fund on average has over 96% (89%) chance of not being held over the next 3 years by US (Canadian) funds. On the other hand, if an asset is held, it is likely to be held in the next quarter. However, such holding probabilities drop much quicker, with 54–61% of bonds still being held at 3-year horizon, which likely reflects high turnover documented above.

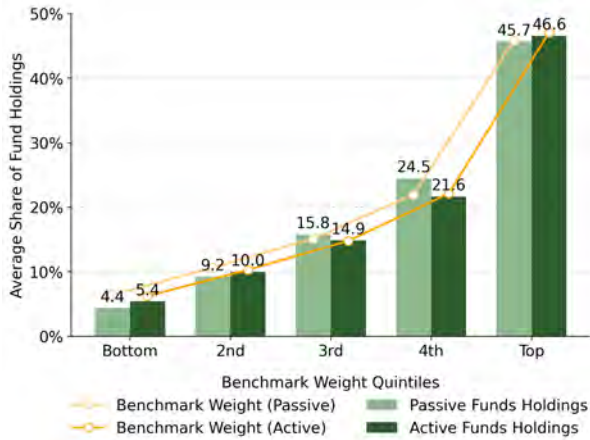
Next, we describe how benchmark index weights affect bond portfolio holdings and their sparsity. Figure 2 characterizes asset allocation across benchmark weight quintiles. We see that funds typically allocate larger shares of assets to bonds with larger index weights, with over 40% of fund assets allocated to the bonds in the top quintile of benchmark weight

across the two geographies. Asset distribution across benchmark weight quintiles is similar to the distribution of benchmark weights for passive funds. Active funds deviate from benchmark weights (and hold more cash and bonds outside of the benchmark) but the distribution across benchmark weight quintiles is roughly similar. Importantly, we see that sparsity also follows this monotonic pattern. As shown in Figure 2c, both active and passive funds hold the smaller fractions of benchmark bonds in the lower benchmark weight quintiles. Moreover, benchmark weights are as important within credit rating buckets. As shown in Appendix Figure A3, funds are more likely to hold bonds in the top quintiles, particularly for bonds with higher ratings.

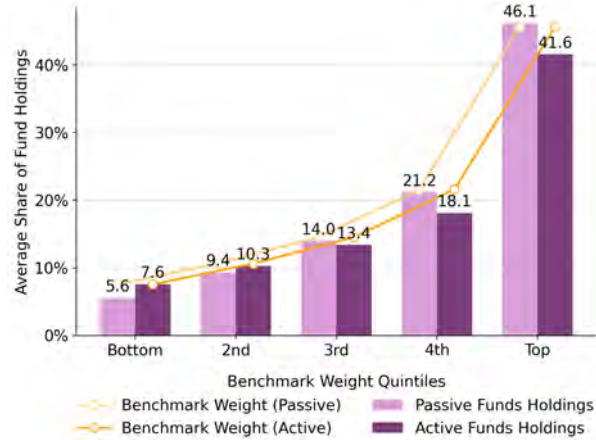
To shed more light on the relationship between benchmark index weights and fund portfolio holdings, we study their conditional correlations. Specifically, Table 3 reports the results of panel regressions of fund holdings on benchmark membership dummy variable and benchmark index weights. In all the regressions, we include both fund-by-date and bond-by-date fixed effects, which remove all the variation in fund or bond characteristics over time, so we only use the variation across funds in holdings of the same bond at the same time. We study both the intensive margin of holdings (portfolio weights) and the extensive margin (propensity to hold the bond). We take the broadest benchmark index bonds as a reference universe for the extensive margin analysis, that is, all corporate bonds in Bloomberg Aggregate Bond Index in the U.S. and all corporate bonds in FTSE Canada Overall Bond Index in Canada.

Columns (1) to (3) of Table 3 show that both Canadian and US funds put higher weights on bonds that are within their benchmarks and have a higher benchmark weight. Similarly, according to columns (4) to (6), funds are more likely to hold a bond (weight above zero) if it is in their benchmark index, especially if it has a high weight in the benchmark. Moreover, these variables contribute significantly to the R^2 , explaining a considerable share of variation in fund holdings especially on the extensive margin (columns (4)–(6)). These findings are overall very similar between the two geographies. We have also verified that these results are unchanged if, instead of including both fund-by-date and bond-by-date

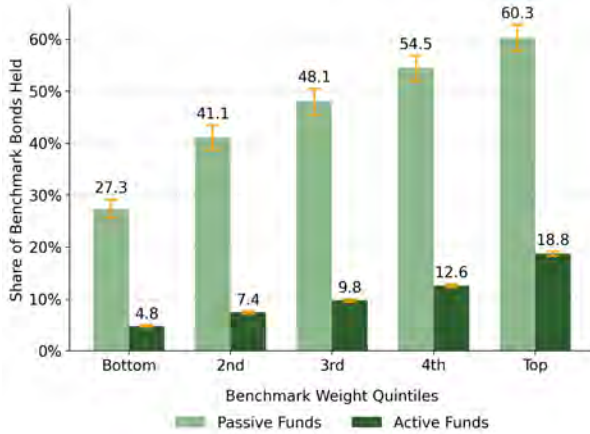
Figure 2: Fund holdings across benchmark weight quintiles



(a) Share of fund holdings across benchmark weight quintiles (Canada)



(b) Share of fund holdings across benchmark weight quintiles (US)



(c) Share of benchmark bonds held across benchmark weight quintiles (Canada)



(d) Share of benchmark bonds held across benchmark weight quintiles (US)

This figure illustrates allocation of active and passive fund assets across bonds in different quintiles of benchmark index weight. Passive funds include index mutual funds and ETFs. We only consider corporate bond part of fund and benchmark portfolios.

fixed effect, we include a comprehensive set of control variables, such as bond size, rating, liquidity, duration, convexity, price, yield, as well as fund-by-date fixed effects and bond fixed effects.

Importantly, we see that these correlations are not driven by passive funds only. Appendix Table A6 reports the estimates in the subsamples of active and passive funds. Even though benchmark weights and especially benchmark membership explain a larger share of variation in passive fund portfolios, these results highlight the importance of benchmarks in

Table 3: Regressions of fund holdings on benchmark index membership and weights for Canada and US

	Fund weight fund weight $\neq 0$			Dummy(fund weight > 0)		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Canada						
Dummy(within benchmark)		0.003*** (0.0004)	0.001 (0.0004)		0.269*** (0.009)	0.158*** (0.009)
Benchmark weight			0.695*** (0.064)			81.84*** (3.84)
Observations	436,171	436,171	436,171	3,623,740	3,623,740	3,623,740
R^2	75.0%	75.0%	75.1%	26.9%	30.8%	31.3%
Within R^2	-	0.07%	0.25%	-	5.3%	6.0%
Panel B: US						
Dummy(within benchmark)		0.085*** (0.005)	0.039*** (0.007)		0.158*** (0.003)	0.128*** (0.004)
Benchmark weight			0.573*** (0.040)			1.646*** (0.050)
Observations	5,070,391	5,070,391	5,070,391	114,805,106	114,805,106	114,805,106
R^2	97.5%	97.5%	97.5%	14.9%	17.9%	18.0%
Within R^2	-	0.04%	0.07%	-	3.5%	3.6%

This table reports estimates of the conditional correlation between fund weight or propensity to hold and benchmark membership dummy variable and benchmark index weight. Panel A reports the results for Canada, and Panel B for the US. The reference universe is the broadest benchmark available in each geography, i.e., Bloomberg Aggregate index in the U.S. and Universe Overall Bond Index in Canada. All regressions are saturated by including fund-by-date and bond-by-date fixed effects. Standard errors clustered at the bond and year-month levels are presented in parentheses. Significance levels are marked as: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

asset allocation of active funds as well.

Overall, the stylized facts we documented above suggest that benchmark membership and benchmark weight are important contributors to the funds' portfolio choice decisions, on both whether or not to include a bond in the portfolio and how much of it to hold. This underscores an interaction between benchmarking and portfolio management costs in corporate bond markets, which we will explore theoretically in the following section.

3 Model

The model features two periods, $t = 0, 1$. The financial market consists of a riskless asset with an exogenous interest rate normalized to zero (e.g., a storage technology), and N

risky assets that pay cash flows D_i in period 1, for $i = 1, \dots, N$. The per-share cash flows of the risky assets are given by

$$D_i = \bar{D}_i + \beta_{1,i}Z_1 + \beta_{2,i}Z_2 + \epsilon_i, \quad \beta_{m,i} \geq 0, \quad i = 1, \dots, N, \quad (1)$$

where $Z_j \sim N(0, \sigma_m^2)$ are common shocks (factors) and $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ is an idiosyncratic one. All shocks are orthogonal to each other. The $N \times 1$ vectors $D \equiv (D_1, \dots, D_N)'$ and $S \equiv (S_1, \dots, S_N)'$ denote vectors of period-1 cash flows and period-0 risky asset prices, respectively. Period-1 risky asset prices equal D . The risky assets are in fixed supply of $\bar{\theta} \equiv (\bar{\theta}_1, \dots, \bar{\theta}_N)'$ shares, which we interpret as the total supply of each asset net of the holdings of buy-and-hold investors (e.g., pension funds and insurance companies in the case of corporate bonds). We denote $\bar{D} \equiv (\bar{D}_1, \dots, \bar{D}_N)'$ and $\beta_m \equiv (\beta_{m,1}, \dots, \beta_{m,N})'$.

There are two types of investors: direct investors and fund managers. Direct investors, who represent a fraction λ_D of the investor population, manage their own portfolios. In contrast, fund managers manage portfolios on behalf of fund investors. Fund investors can directly purchase the riskless asset but do not trade stocks themselves; instead, they delegate portfolio selection to fund managers, who receive compensation for this service. A manager's performance is evaluated relative to the benchmark index. The benchmark index is a portfolio of $\omega \equiv (\omega_1, \dots, \omega_N)'$ shares of the assets described above. For realism, we constrain the components of ω to be nonnegative; some may be zero. The assets are ordered by their benchmark weight ω_i , with the first asset having the highest weight. All investors have constant absolute risk aversion (CARA) utility over terminal wealth (or compensation), given by $U(W) = -\exp(-\gamma W)$, where γ is the absolute risk aversion coefficient.

The terminal wealth of a direct investor is given by $W = W_0 + \theta_D'(D - S)$, where the $N \times 1$ vector θ_D denotes the number of shares held by the direct investor, and W_0 is the investor's initial wealth. The direct investor chooses the portfolio θ_D to maximize their utility $U(W)$. The assets under management delegated to the fund manager are denoted by W_F , and the manager selects a portfolio consisting of θ shares to maximize her utility $U(w)$, where

w is her compensation. The fund manager’s compensation w consists of three components: the first is a linear payout based on the fund’s absolute performance, the second depends on the fund’s performance relative to a benchmark portfolio, and the third component (c) is independent of performance. Specifically,

$$w = aR + b(R - B) + c - Cn, \quad a \geq 0, \quad b > 0, \quad n \leq N, \quad (2)$$

where $R \equiv \theta'(D-S)$ represents the fund portfolio’s absolute performance, and $B \equiv \omega'(D-S)$ denotes the benchmark portfolio’s performance. The term Cn in (2) represents portfolio management costs that scale with the number of assets (n) included in the fund’s portfolio. We interpret these costs as encompassing transaction and rebalancing expenses, monitoring costs, and other portfolio management overhead. Larger portfolios naturally entail higher costs. These costs need not arise immediately; they may reflect, for instance, expected future trading or rebalancing expenditures. For tractability, we model them as fixed—independent of the number of shares traded. Such costs are presumably lower for index funds than for actively managed funds, given the lower frequency of trades and reduced monitoring requirements. Finally, the parameters a and b capture the manager’s compensation sensitivity to absolute and relative portfolio performance, respectively.

3.1 Investor Optimization Problems

The direct investor’s problem at time $t = 0$ is

$$\max_{\theta_D} E_0[-\exp\{-\gamma W\}]. \quad (3)$$

Substituting in the budget constraint $W_1 = W_0 + \theta'_D(D - S)$, and evaluating the expectation in (3), we can equivalently represent the direct investor's problem as follows:⁴

$$\max_{\theta_D} \left[-\exp\left\{-\gamma[W_0 + \theta'_D(\bar{D} - S) - \underbrace{\frac{1}{2}\gamma\theta'_D\Sigma\theta_D}_{\text{Variance}}]\right\} \right].$$

The solution to this mean-variance optimization problem is well-known and yields the following demand function:

$$\theta_D = \frac{1}{\gamma}\Sigma^{-1}(\bar{D} - S). \quad (5)$$

It is straightforward to show that the optimization problem of a fund manager can be represented as $\max_{\theta,n} E[-\exp\{-\gamma(a+b)\theta'(D-S) - b\omega'(D-S)\}] - Cn$, which is equivalent to the following mean-tracking-error-variance optimization problem

$$\max_{\theta} (a+b)\theta'(\bar{D} - S) - \frac{1}{2}\gamma \underbrace{((a+b)\theta - b\omega)'\Sigma((a+b)\theta - b\omega)}_{\text{Tracking error variance}} - Cn. \quad (6)$$

The first term implies that the manager should select assets with higher expected returns and lower tracking error variance. Unlike a direct investor, the manager is concerned with the fund's tracking error variance rather than its total variance. This is because the manager is benchmarked to an index, meaning that the zero-risk portfolio from her perspective is one that perfectly replicates the benchmark (specifically, $\theta = \frac{b}{a+b}\omega$). The final component of the manager's tradeoff is the portfolio management cost, which discourages including too many assets in the fund's portfolio.

⁴To evaluate the expectation in (3), we need the following property. Suppose $Y \sim N(E[Y], Var[Y])$ is an $N \times 1$ random vector, α is a (constant) scalar, and x is a constant vector. Then

$$Ee^{\alpha x'Y} = e^{\alpha x'E[Y] + \frac{\alpha^2}{2}x'Var[Y]x}. \quad (4)$$

3.2 Equilibrium with Uncorrelated Cash Flows

We begin by solving a simplified version of the model without systematic risk, i.e., $\beta_{m,i} = 0$. To sharpen our results in this section, we assume that the benchmark weights ω are proportional to the market capitalization of the assets $\bar{\theta}$ —a realistic assumption given that bond indexes are typically value-weighted. We normalize the weights in the fund's benchmark to be $\omega = 1/W_F\bar{\theta}$, where W_F denotes the fund's initial assets under management delegated by investors.

Equilibrium prices in this economy are derived from the market-clearing condition $\lambda_D\theta_D + \lambda_M\theta_M = \bar{\theta}$. The following proposition reports the portfolios of direct investors and fund managers and equilibrium asset prices.

Proposition 1 *The portfolios of a direct investor and of a fund manager are given by*

$$\theta_{D,i} = \frac{1}{\gamma\sigma_i^2}(\bar{D}_i - S_i), \quad (7)$$

$$\theta_{M,i} = \begin{cases} \frac{1}{\gamma\sigma_i^2(a+b)}(\bar{D}_i - S_i) + \frac{b}{a+b}\omega_i, & \text{if } i \leq n^* \quad (\text{held by managers}) \\ 0, & \text{if } n^* < i \leq N \quad (\text{not held by managers}), \end{cases} \quad (8)$$

and the equilibrium asset prices are

$$S_i = \begin{cases} \bar{D}_i - \gamma \frac{1}{\lambda_D + \frac{\lambda_F}{a+b}} \sigma_\epsilon^2 \left(\bar{\theta}_i - \frac{b}{a+b} \underbrace{\lambda_F \omega_i}_{BMI} \right), & i \leq n^* \quad (\text{held by managers}) \\ \bar{D}_i - \gamma \frac{1}{\lambda_D} \sigma_\epsilon^2 \bar{\theta}_i, & n^* < i \leq N \quad (\text{not held by managers}). \end{cases} \quad (9)$$

The optimal number of assets in the funds' portfolios, n^* , satisfies the following inequalities:

$$\frac{\gamma}{2} \sigma_\epsilon^2 \left(\frac{\bar{\theta}_{n^*}}{\lambda_D} + b\omega_{n^*} \right)^2 - C \geq 0, \quad (10)$$

$$\frac{\gamma}{2} \sigma_\epsilon^2 \left(\frac{\bar{\theta}_{n^*+1}}{\lambda_D} + b\omega_{n^*+1} \right)^2 - C < 0. \quad (11)$$

To develop an intuition for the structure of equilibrium reported in Proposition 1, it is

useful to consider a special case without portfolio management costs. In that special case, all assets are held by fund managers, i.e., $n^* = N$. Direct investors hold standard mean-variance portfolios. Fund managers' portfolios consist of a combination of the mean-variance portfolio and an additional portfolio that hedges the managers against their underperformance relative to the benchmark. That hedging portfolio is the benchmark index, and, as evident from (8), the demand for this portfolio is inelastic, i.e., it does not depend on asset prices and depends only on the strength of the absolute and relative performance sensitivities of the managers' pay.⁵ Equilibrium asset prices reflect the presence of both direct investors and fund managers. The presence of the latter gives rise to a term $b/(a+b)\lambda_F\omega_i$ in asset prices, which reflects the aggregate inelastic demand from fund managers. It implies that the higher the mass of asset managers (λ_F) and the higher the weight of an asset in the benchmark ω_i , the higher the price of the asset. [Pavlova and Sikorskaya \(2023\)](#) generalize these insights to a setting with heterogeneous benchmarks, in which case the aggregate inelastic demand of fund managers is proportional to $\sum_j \lambda_{F,j}\omega_{i,j}$, where j indexes benchmarks and $\lambda_{F,j}$ represents the mass of managers evaluated relative to benchmark j . They term this quantity Benchmarking Intensity (BMI), and we adopt the same terminology in this paper.

We now turn to the case with portfolio management costs, $C > 0$. In this case, it may no longer be optimal for fund managers to hold all available assets in their portfolios, as each additional asset contributes to the portfolio management cost. [Proposition 1](#) reveals that the managers hold *sparse* portfolios that exclude all assets with a rank above n^* . Recall from our assumption earlier in this section that these assets are the ones with the lowest benchmark weight. While excluding any assets from the portfolio increases its tracking error, assets with the lowest benchmark weight have the smallest effect on it.

To understand the nuances of the tradeoff the manager faces when deciding whether to add a marginal asset to the portfolio, it is useful to examine the value function of a fund

⁵This observation was first made in [Kashyap, Kovrijnykh, Li, and Pavlova \(2021\)](#).

manager whose portfolio size is fixed at n assets.

$$V(n) \equiv \max_{\theta} \underbrace{(a+b)\theta'(\bar{D}-S)}_{\text{Expected return}} - \frac{1}{2}\gamma \underbrace{((a+b)\theta - b\omega)' \Sigma ((a+b)\theta - b\omega)}_{\text{Tracking error variance}} - \underbrace{Cn}_{\text{Cost}} \quad (12)$$

When adding an asset, a manager balances the marginal benefit of increasing the portfolio's expected return and reducing its tracking error variance with the marginal cost of managing a larger portfolio. The inequalities determining the optimal number of assets n^* presented in (10)–(11) are equivalent to

$$V(n^*) - V(n^* - 1) = \frac{\gamma}{2}\sigma_{\epsilon}^2 \left(\frac{\bar{\theta}_{n^*}}{\lambda_D} + b\omega_{n^*} \right)^2 - C \geq 0, \quad (13)$$

$$V(n^* + 1) - V(n^*) = \frac{\gamma}{2}\sigma_{\epsilon}^2 \left(\frac{\bar{\theta}_{n^*+1}}{\lambda_D} + b\omega_{n^*+1} \right)^2 - C < 0, \quad (14)$$

respectively. The first two terms in (10)–(11) represent the marginal benefit of including an asset, while the last term captures the marginal cost, which is simply a per-asset portfolio management cost C . As (10)–(11) reveal, the marginal benefit depends positively on three asset characteristics: the asset's (i) weight in the benchmark ω_i , (ii) size $\bar{\theta}_i$, and (iii) idiosyncratic risk σ_{ϵ} . We have already discussed that assets with higher benchmark weights contribute more to reducing the fund's tracking error. Idiosyncratic risk acts in a similar way: if an asset has a high idiosyncratic risk, including it in the portfolio reduces the fund's tracking error.⁶ Furthermore, an asset's size is positively related to the likelihood of its inclusion in the portfolio because larger assets in our model have higher expected returns. Assets with greater market capitalization constitute a larger share of the market and thus have higher betas, making them riskier and requiring higher expected returns as compensation for risk.

We next discuss asset prices in the presence of portfolio sparsity. Since funds do not hold assets at the bottom of the benchmark, fund managers do not participate in pricing

⁶Although we assume for parsimony that all assets share the same idiosyncratic risk, equations (10)–(11) remain valid even without this assumption. One simply needs to replace σ_{ϵ} with σ_{ϵ_i} , where the latter denotes asset-specific idiosyncratic risk.

these assets. Instead, their prices are determined solely by direct investors, whereas the prices of the top n^* assets are determined by both direct investors and fund managers. This is the key intuition behind equation (9). The pricing difference between the top n^* assets and the remaining ones parallels that in Merton (1987), where certain groups of investors (exogenously) are unaware of and do not hold subsets of assets. Consequently, asset prices in our model depend on the *sparse BMI* ($sBMI$), defined as

$$sBMI_i = \begin{cases} BMI_i, & \text{if } i \leq n^*, \\ 0, & \text{if } n^* < i \leq N, \end{cases} \quad (15)$$

a novel measure that quantifies inelastic demand in asset markets stemming from fund manager mandates and model-implied portfolio sparsity.

We conclude the subsection by presenting several testable predictions of our model for the special case of uncorrelated cash flows.

Testable Prediction 1: Fund managers hold sparse portfolios.

Testable Prediction 2: Fund managers are more likely to hold assets with (i) higher benchmark weights, (ii) larger market capitalization (size), (iii) higher idiosyncratic risk.

Testable Prediction 3: Asset prices with higher sparse BMIs has higher prices and lower expected returns.

Testable Prediction 4: The index inclusion effect occurs only for assets with non-zero sparse BMI.

3.3 Equilibrium with Correlated Cash Flows

While the uncorrelated cash flows specification offers analytical tractability, it rules out asset substitutability—a key channel for our application. We therefore introduce common shocks (factors) to the cash flows in (1) by assuming $\beta_{m,i} > 0$.

The problem of solving the N -asset version of the model is a combinatorics problem. To highlight the three main insights from our model related to asset substitutability, it is not essential to present the full solution to the problem. Instead, we focus on characterizing the fund manager's decision to include an additional asset in a sparse portfolio.

We begin with a two-asset version of our model with one common shock (risk factor). For convenience, we adopt the following notation for the variance–covariance matrix Σ :

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \quad (16)$$

where the variances σ_i^2 , $i = 1, 2$, capture both the common and idiosyncratic components of the cash flows, and ρ_{12} denotes their correlation.

A manager facing no portfolio management costs optimally holds both assets. In an economy with portfolio management costs, the cost of holding both assets may be prohibitively high. We focus on the interesting and realistic scenario in which the cost is sufficiently high and portfolios are sparse. For illustration, we consider the case in which the manager opts for Asset 1 rather than Asset 2. Her (sparse) optimal portfolio is presented in the following lemma.

Lemma 1 *If the manager is restricted to hold only one asset, Asset 1, in her portfolio, her optimal portfolio is*

$$\theta_{M,i} = \begin{cases} \frac{\bar{D}_1 - S_1}{\gamma\sigma_1^2(a+b)} + \frac{b}{a+b}\omega_1 + \underbrace{\frac{b}{a+b} \frac{\rho_{12}\sigma_1\sigma_2}{\sigma_1^2}}_{\text{Additional demand}} \omega_2, & \text{if } i = 1, \\ 0, & \text{if } i = 2. \end{cases} \quad (17)$$

The optimal portfolio of a fund manager who holds both assets is given by

$$\theta = \frac{1}{\gamma(a+b)}\Sigma^{-1}(\bar{D} - S) + \frac{b}{a+b}\omega \quad (18)$$

Lemma 1 shows that if—due to portfolio management costs—the manager opts to

hold only Asset 1, she holds *more* of it than she would in the economy without portfolio management costs. With correlated assets, Asset 1 acts as a substitute for Asset 2, so by overweighting Asset 1, the manager gains exposure to Asset 2's common risk factor. This effect disappears when cash flows are uncorrelated: setting $\rho_{12} = 0$ eliminates it.

As in the uncorrelated cash flows case considered in Section 3.2, the prices of assets not included in managers' (sparse) portfolios are lower than in the economy without portfolio management costs. The following lemma reports asset prices in both economies.

Lemma 2 *If the manager is restricted to hold only one asset, Asset 1, in her portfolio, asset prices are given by*

$$\begin{aligned} S_1 &= \bar{D}_1 - \gamma A \left[(\sigma_1^2 \bar{\theta}_1 + \rho_{12} \sigma_1 \sigma_2 \bar{\theta}_2) - \frac{b\lambda_F}{a+b} (\sigma_1^2 \omega_1 + \rho_{12} \sigma_1 \sigma_2 \omega_2) \right], \\ S_2 &= \bar{D}_2 - \gamma A \left[(\rho_{12} \sigma_1 \sigma_2 \bar{\theta}_1 + \rho_{12}^2 \sigma_2^2 \bar{\theta}_2) - \frac{b\lambda_F}{a+b} (\rho_{12} \sigma_1 \sigma_2 \omega_1 + \rho_{12}^2 \sigma_2^2 \omega_2) \right] - \frac{\gamma}{\lambda_D} \sigma_2^2 (1 - \rho_{12}^2) \bar{\theta}_2, \end{aligned} \quad (19)$$

where $A \equiv \left[\lambda_D + \frac{\lambda_F}{a+b} \right]^{-1}$.

In the economy without portfolio management costs, where asset managers hold both assets, the price of Asset 1 is the same as in (19). The price of Asset 2 is given by

$$S_2 = \bar{D}_2 - \gamma A \left[(\rho_{12} \sigma_1 \sigma_2 \bar{\theta}_1 + \sigma_2^2 \bar{\theta}_2) - \frac{b\lambda_F}{a+b} (\rho_{12} \sigma_1 \sigma_2 \omega_1 + \sigma_2^2 \omega_2) \right].$$

By examining the expressions for S_2 in both economies, it is easy to see that Asset 2's price is lower in the economy with portfolio management costs, unless $\rho_{12} = 1$. Since the managers do not hold Asset 2, the demand for it is lower and so is its price.

It follows from Lemma 2 that, in this economy with portfolio management costs, sBMI is

$$sBMI_i = \begin{cases} BMI_i, & \text{if } i = 1, \\ \rho_{12}^2 BMI_i, & \text{if } i = 2, \end{cases} \quad (20)$$

where we define asset's i 's BMI as in the previous subsection, namely, $BMI_i = \lambda_F \omega_i$. One can see this by examining the sensitivities of Assets 1 and 2 to their BMIs. In contrast to

the economy of Section 3.2, in the economy with correlated cash flows, the price of Asset 2 depends on its BMI. The intuition is that the manager can reduce the tracking-error variance of her portfolio by substituting Asset 2 with Asset 1. This channel is reflected in the last term of (17), which depends on Asset 2's benchmark weight ω_2 . However, the sensitivity of the price to BMI is lower than in the economy without portfolio management costs. The less substitutable Assets 1 and 2 are, the lower the sensitivity. In the limit, when the cash flows of the two assets are uncorrelated (as in Section 3.2), the sensitivity vanishes.

We now turn to examining the manager's tradeoff involved in adding Asset 2 to her portfolio. With a slight abuse of notation, denote the value function of a fund manager who is restricted to hold only Asset 1 as $V(1)$ and that of a manager who holds both assets $V(2)$. The former is the value function as in equation (12), provided that the manager is restricted to hold only Asset 1. The marginal benefits and costs of adding Asset 2 can be assessed by comparing the two value functions. The following lemma reports the outcome of this comparison.

Lemma 3 *Let $V(1)$ be the value function of the manager who is restricted to hold only one asset, Asset 1, in her portfolio, and $V(2)$ the value function of the manager who holds both assets.*

$$V(2) - V(1) = (1 - \rho_{12}^2) \frac{\gamma}{2} \sigma_2^2 \left(\frac{\bar{\theta}_2}{\lambda_D} + b\omega_2 \right)^2 - C$$

The expression for the costs and benefits of adding Asset 2 to the manager's portfolio presented in Lemma 3 parallels that in equation (13) for the uncorrelated case flows. Specifically, the marginal benefit of including Asset 2 in the portfolio is higher if the asset has (i) higher benchmark weights, (ii) larger market capitalization, (iii) higher idiosyncratic risk. The marginal cost of including Asset 2 is C . The additional consideration that the correlated case brings is the influence of asset substitutability, as measured by the correlation ρ_{12} . Lemma 3 reveals that the more substitutable Assets 1 and 2 are, the lower the marginal benefit to the manager of adding Asset 2 to her portfolio. This is intuitive because

the manager can substitute for holding Asset 2 by holding more of Asset 1, which saves on the portfolio management costs, while still delivering a similar expected return and tracking error variance reduction.

For illustrative purposes, we have assumed in this subsection that fund managers first include Asset 1 in their portfolio and then Asset 2, provided that the benefits of doing so exceed the portfolio management costs. The case of managers first including Asset 2 and then Asset 1 is analyzed analogously. Solving the problem fully requires the comparison of the value functions in each of these cases and establishing which combination of assets yields the highest value function. Since our goal here is to illustrate the tradeoff facing the manager contemplating the inclusion of an additional asset, we do not present the full solution.

In practice, fund managers sample bonds from distinct categories, for example, credit rating or maturity buckets. To model this choice, we consider a 4-asset version of our model with two common shocks. For expositional simplicity, we assume that the two common shocks are orthogonal to each other (this can be achieved by appropriately redefining the shocks). Under this assumption, the variance-covariance matrix of the cash flows has a block-diagonal form, representing two distinct buckets:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & 0 & 0 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \hat{\sigma}_1^2 & \hat{\rho}_{12}\hat{\sigma}_1\hat{\sigma}_2 \\ 0 & 0 & \hat{\rho}_{12}\hat{\sigma}_1\hat{\sigma}_2 & \hat{\sigma}_2^2 \end{pmatrix}, \quad (21)$$

where the variances of the assets' cash flows σ_i^2 and $\hat{\sigma}_i^2$, $i = 1, 2$ capture both common and idiosyncratic variances of the cash flows, and ρ_{12} and $\hat{\rho}_{12}$ denote their correlations.

The intuition from the two-asset case extends directly. Fund managers prefer to add assets that are less correlated with their existing holdings. For example, suppose the portfolio currently contains Asset 1. With a block-diagonal covariance structure—payoffs correlated within a block but uncorrelated across blocks—the marginal benefit of adding Asset 2 (same block) is smaller, *ceteris paribus*, than the benefit of adding Asset 3 or Asset 4 (a different

block).

The analysis in this subsection generates several additional testable predictions of our model.

Testable Prediction 5: Managers select within “buckets” of substitutable assets.

Testable Prediction 6: Within each bucket, managers are likely to hold assets with (i) higher benchmark weights, (ii) larger market capitalization, (iii) higher idiosyncratic risk.

Testable Prediction 7: Sparse BMI coincides with the original BMI for assets that are held by managers and is lower for the assets that are not held.

4 Sparse BMI in Bond Markets: Empirical Analysis

In this section, we build a measure of sparse benchmarking intensity using data on Canada and U.S. domestic fixed income mutual funds and ETFs and their benchmarks, and test the model predictions in both U.S. and Canadian corporate bond markets.

4.1 Empirical measure of sparse benchmarking intensity

Guided by the model, we calculate the sparse benchmarking intensity (BMI) for bond i in month t as

$$sBMI_{i,t} = \frac{\sum_{j=1}^J \lambda_{j,t} \times \omega_{i,j,t} \times D_{i,j,t}}{MV_{i,t}}, \quad (22)$$

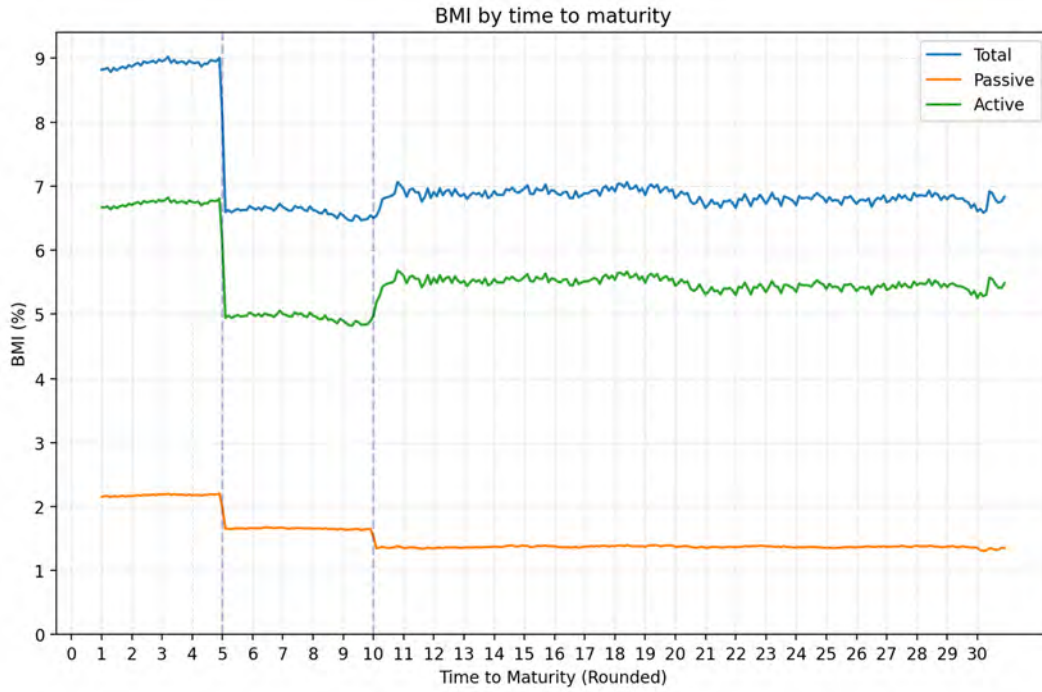
where $\lambda_{j,t}$ is the total assets under management (AUM) of mutual funds and ETFs benchmarked to index j in month t , $\omega_{i,j,t}$ is the weight of bond i in index j in month t , $D_{i,j,t}$ is a dummy variable corresponding to bond i 's rank in benchmark j in month t , and $MV_{i,t}$ is the market value of bond i in month t . Importantly, what makes this BMI sparse is that we only count the bond's weight in the index if the bond's rank in that index is high enough, as captured by $D_{i,j,t}$. For example, we consider the top half of each index the bond belongs to or the top half of each rating bucket of each index the bond belongs to. This measure

picks up how much capital is benchmarked against a bond. Yet, because it excludes indexes in which the bond has low weight—in line with our definition of BMI in the model—this measure also reflects portfolio management costs in the corporate bond market.

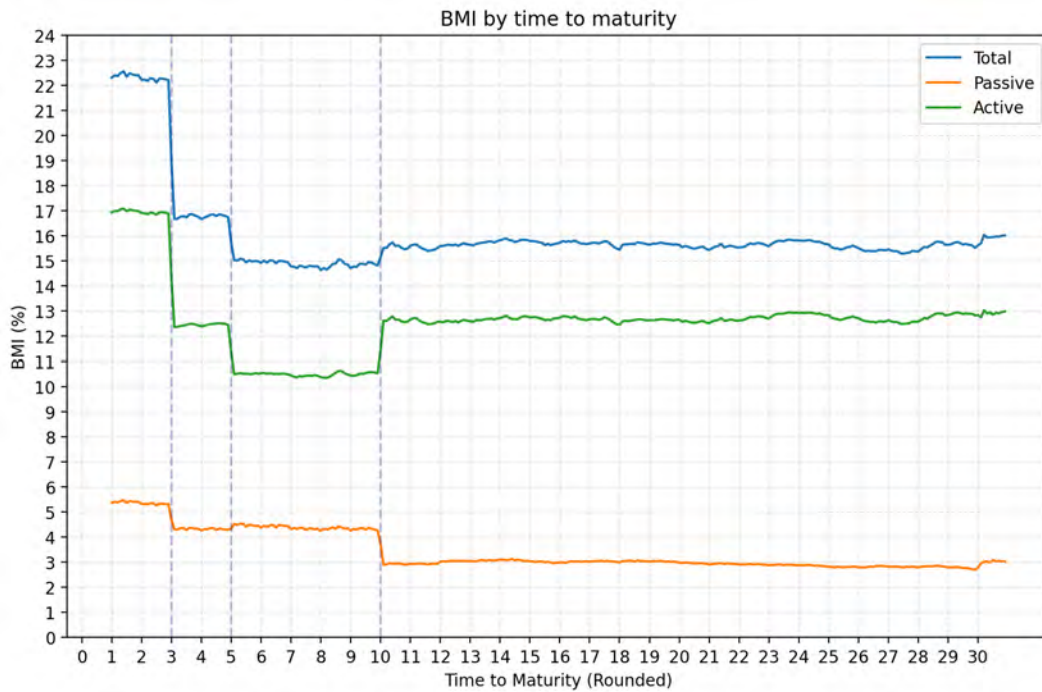
There is no clear theoretical guidance on the ‘right’ level of sparsity or the optimal granularity of bucketing (e.g., by rating, sector, or a combination). To explore this, we conduct a series of diagnostics in which we vary the sparsity cutoff and evaluate the model’s ability to predict whether a bond is held by any benchmarked fund (a binary ownership indicator). The results, presented as AUC (area under the curve) plots for both the U.S. and Canadian samples in Appendix Figure B1, show that the true–false positive trade-off shifts as the sparsity threshold changes. Across all bucketing approaches, sparsity cutoffs around deciles 4–6 yield economically similar predictive performance. We verified this by bootstrapping ownership and BMI data to generate distributions of the Youden J statistic and testing for differences across sparsity cutoffs. For simplicity and interpretability, we focus on an ‘above-median’ sparsity definition in the main analysis and use it consistently in subsequent sections. This choice introduces some measurement noise: alternative sparsity thresholds or bucketing schemes may classify marginal bonds differently. Importantly, this type of noise is likely to attenuate the estimated effects in the subsequent sections, meaning that the true economic impact of benchmarking could be even stronger than what we document.

Tables A3 and A4 report descriptive statistics in the full sample for the U.S. and Canada. The average total BMI in the U.S. is over 17%, while in Canada it is over 10%. In both samples, the average sparse BMIs are roughly half of the total ones, as we primarily use BMIs with sparsity below median. These descriptives mask considerable heterogeneity by bond maturity, as we show below. Because of the higher granularity and depth in the U.S. corporate bond market, which we discuss in Section 2.2, we will report the results for two additional sBMIs that exploit this granularity.

Figure 3: Benchmarking intensity (BMI) by bond maturity



(a) Canada



(b) US

This figure illustrates the average benchmarking intensity by bond's time to maturity in the corporate bond markets in (a) Canada and (b) US. The sample includes all bonds in the broadest benchmark in each geography.

Both in Canada and the U.S., the benchmarked AUM in the mutual fund and ETF sector varies by bond maturity. Figure 3 illustrates that this variation in assets generates discontinuities in BMI around maturity cutoffs, with different sign of discontinuities by fund type. Specifically, we see that in Canada, there are two cutoffs, at 5 years and 10 years to maturity. At the 5-year cutoff, both active and passive BMIs increase, while at the 10-year cutoff, there is a decline in active BMI and an increase in passive BMI. In the United States, there is an additional cutoff at 3 years to maturity, at which both active and passive BMIs increase. At the 5-year cutoff in the U.S., there is a sharp increase in active BMI, while passive BMI virtually does not change. Finally, at the 10-year cutoff, there is a decrease in active BMI and an increase in passive BMI. In Appendix B.2, we use the 5-year cutoff to illustrate that sparse BMIs experience the same jumps as the total BMI, while in Appendix B.3, we formally estimate the size of the discontinuities for all BMIs at the 3/5/10-year cutoffs by regressing the changes in (sparse) BMI on cutoff-specific switch dummy variables.

4.2 Identification using changes in sBMI

A key challenge in estimating the effect of sparse BMI on bond holdings or prices is that the level of sparse BMI is not random. Bonds with high sBMI tend to be larger, more liquid, and of higher credit quality, so cross-sectional regressions of prices or ownership on the level of sBMI risk conflating benchmarking with persistent fundamentals. To mitigate this concern, we focus on changes in sBMI and exploit within-bond variation over time.

The definition in Section 4.1 implies that, after cancelling out the bond’s own price from the index weight and the denominator, sBMI for bond i in month t can be written as

$$sBMI_{i,t} = \sum_j \frac{\lambda_{j,t}}{IndexMV_{j,t}} D_{i,j,t},$$

where $IndexMV_{j,t}$ is the market value of that index, while $D_{i,j,t}$ and $\lambda_{j,t}$ are defined around (22). Thus, for a given bond, variation in $sBMI_{i,t}$ over time arises from two sources: (i) changes in the ratio $\lambda_{j,t}/IndexMV_{j,t}$, which capture how much benchmarked capital is al-

located to each index relative to its market value, and (ii) changes in $D_{i,j,t}$, which reflect both changes in the bond’s index membership and its movements into and out of specific sparse segments of the index. In our setting, D is highly persistent and largely predictable from remaining maturity and issue size, both of which evolve slowly for investment-grade corporate bonds. Thus, although maturity-based transitions at the 3-, 5-, and 10-year cutoffs generate large discrete changes in $sBMI_{i,t}$, these shifts are predictable, and the economically meaningful variation arises primarily through changes in $\lambda_{j,t}/IndexMV_{j,t}$. We therefore do not rely on unanticipated changes in $D_{i,j,t}$.

Instead, our identification comes primarily from changes in the amount of benchmarked capital per unit of index market value, $\lambda_{j,t}/IndexMV_{j,t}$, for bonds with given maturity, size, and rating. Intuitively, once we control for these characteristics and include bond and year-month fixed effects, month-to-month changes in $sBMI_{i,t}$ largely reflect reallocations of benchmarked assets across maturity-rating segments driven by fund flows and benchmark rebalancing, rather than bond-specific news. Because the bond’s own price cancels out of the $sBMI_{i,t}$ formula, there is no mechanical channel through which changes in yields feed back into sparse benchmarking intensity.

Our empirical specification is therefore based on changes in sBMI. In the monthly panel of investment-grade corporate bonds, we estimate

$$\Delta Y_{i,t+h} = \beta \Delta sBMI_{i,t} + \zeta' \bar{X}_{i,t} + \nu_i + \mu_t + \epsilon_{i,t+h}, \quad (23)$$

where $\Delta Y_{i,t+h}$ is either the change in bond i ’s price over $h = \{1, \dots, 6\}$ months from t (measured with yield, yield spread, or total return) or the change in its total, passive, or active fund ownership from $t - h$ to $t + h$, with $h \in \{3, 6, 9, 12\}$ months. $\Delta sBMI_{i,t}$ is the change in sparse BMI in month t . Bond-level controls in $\bar{X}_{i,t}$ include logarithm of issue amount outstanding, numeric credit rating, bid-ask spread quintile, and time to maturity. ν_i and μ_t are bond and year-month fixed effects, and we cluster standard errors by issuer and year-month. Identification comes from within-bond variation in $\Delta sBMI_{i,t}$ after controlling

for these observables and fixed effects.

This design differs from [Bretscher, Schmid, and Ye \(2024\)](#), who focus on passive ownership and use indicator variables for crossing maturity cutoffs as their main source of variation. We instead use a continuous measure of sparse BMI that aggregates both active and passive benchmarked capital, and we exploit changes in sBMI for all bond-month observations rather than focusing on windows around maturity cutoffs. Under the assumption that relative reallocations of capital across benchmark indexes are orthogonal to bond i 's pricing and ownership, our estimates of β recover the causal effect of changes in sparse BMI on corporate bond yields and institutional holdings.

4.3 sBMI and bond pricing

In this section, we report and discuss the estimation results for pricing variables.

Table 4 shows how corporate bond yield spreads react to changes in the total and sparse BMIs. Consistent with model predictions, an increase in sparse BMI leads to a persistent decrease in bond yields in both Canada and the U.S. Importantly, as shown in column (1), total BMI cannot capture this effect, and in our model, that happens because of how portfolio management costs shape sparsity in benchmarked fund holdings. Tables B2 and B3 report the estimates for changes in yields and total bond returns, respectively. Similar to the results reported in this section, yields decrease with sBMI while total returns increase.

The estimated effects are economically modest but clearly detectable, particularly for the sparse BMI measures. Using the standard deviation of monthly changes in the sBMIs, we estimate that a one-standard-deviation increase in the rating-based sparse BMI reduces U.S. bond yield spreads by roughly 2–3 basis points over the $t - 1$ to $t + 1$ window. For example, the coefficient of -0.016 combined with a standard deviation of 1.5 implies a change of approximately -2.4 bps (around 6% of the monthly standard deviation of yield spread changes, which is about 39 bps). The corresponding effect in Canada is smaller, roughly -0.5 bps for the same horizon, or about 2% of the monthly standard deviation of Canadian spread changes (23 bps). Across specifications, the total BMI measure has

Table 4: Changes in BMIs and bond yield spreads (United States and Canada)

	<u>Total BMI</u>		<u>Sparse BMI</u>			
		Rating	Segment– Rating	Segment– Rating– Maturity	Sector– Rating	Sector– Rating– Maturity
Panel A: United States						
$t-1$ to $t+1$	–0.004** (0.001)	–0.016*** (0.004)	–0.014*** (0.003)	–0.010*** (0.003)	–0.007*** (0.002)	–0.004*** (0.001)
Observations	605,726	605,726	605,726	605,726	605,726	605,726
R-squared	0.492	0.492	0.492	0.492	0.492	0.492
$t-1$ to $t+3$	–0.002 (0.002)	–0.011*** (0.002)	–0.010*** (0.002)	–0.008*** (0.002)	–0.005*** (0.001)	–0.003*** (0.001)
Observations	589,024	589,024	589,024	589,024	589,024	589,024
R-squared	0.506	0.506	0.506	0.506	0.506	0.506
$t-1$ to $t+6$	0.008*** (0.002)	–0.008*** (0.002)	–0.007*** (0.002)	–0.005*** (0.001)	–0.003*** (0.001)	–0.002*** (0.000)
Observations	564,148	564,148	564,148	564,148	564,148	564,148
R-squared	0.510	0.510	0.510	0.510	0.510	0.510
Panel B: Canada						
$t-1$ to $t+1$	–0.008 (0.007)	–0.008*** (0.001)	–0.006*** (0.001)	–0.003*** (0.001)		
Observations	123,748	123,748	123,748	123,748		
R-squared	0.619	0.619	0.619	0.619		
$t-1$ to $t+3$	0.003 (0.010)	–0.005*** (0.001)	–0.003*** (0.001)	–0.003*** (0.001)		
Observations	123,402	123,402	123,402	123,402		
R-squared	0.682	0.682	0.682	0.682		
$t-1$ to $t+6$	0.003 (0.011)	–0.004** (0.002)	–0.002** (0.001)	–0.002*** (0.001)		
Observations	122,781	122,781	122,781	122,781		
R-squared	0.722	0.722	0.722	0.722		

This table reports the $\Delta sBMI$ coefficient estimates from regression (23) for the U.S. (Panel A) and Canadian (Panel B) samples. The dependent variable is the change in yield spread (relative to a maturity-matched Treasury bond) in percentage points. All regressions control for the logarithm of the issue amount outstanding, numeric credit rating, bid–ask spread quintile, and time to maturity, and include bond and year–month fixed effects. Standard errors are double-clustered by issuer and year–month and reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results in Panel B omit sector-based sparse BMIs due to the absence of this classification level in Canada (blank cells).

smaller predictive power, while the sparse BMIs—especially those based on ratings and rating–segment groupings—consistently show larger and more persistent effects, with the U.S. effects being roughly two to three times larger in magnitude than those in Canada.

We can also consider the magnitude of sBMI changes due to crossing the 5-year cutoff specifically. The switch results in the increase of 1.4 p.p. in the rating-based sparse BMI in

Canada (see Table B1), which implies a decrease in yield spreads of 1.12 bps. In the U.S., the switch results in an increase of 46 bps in the rating-based sparse BMI, corresponding to a decrease in yield spreads of 0.7 bps.

4.4 sBMI and fund holdings

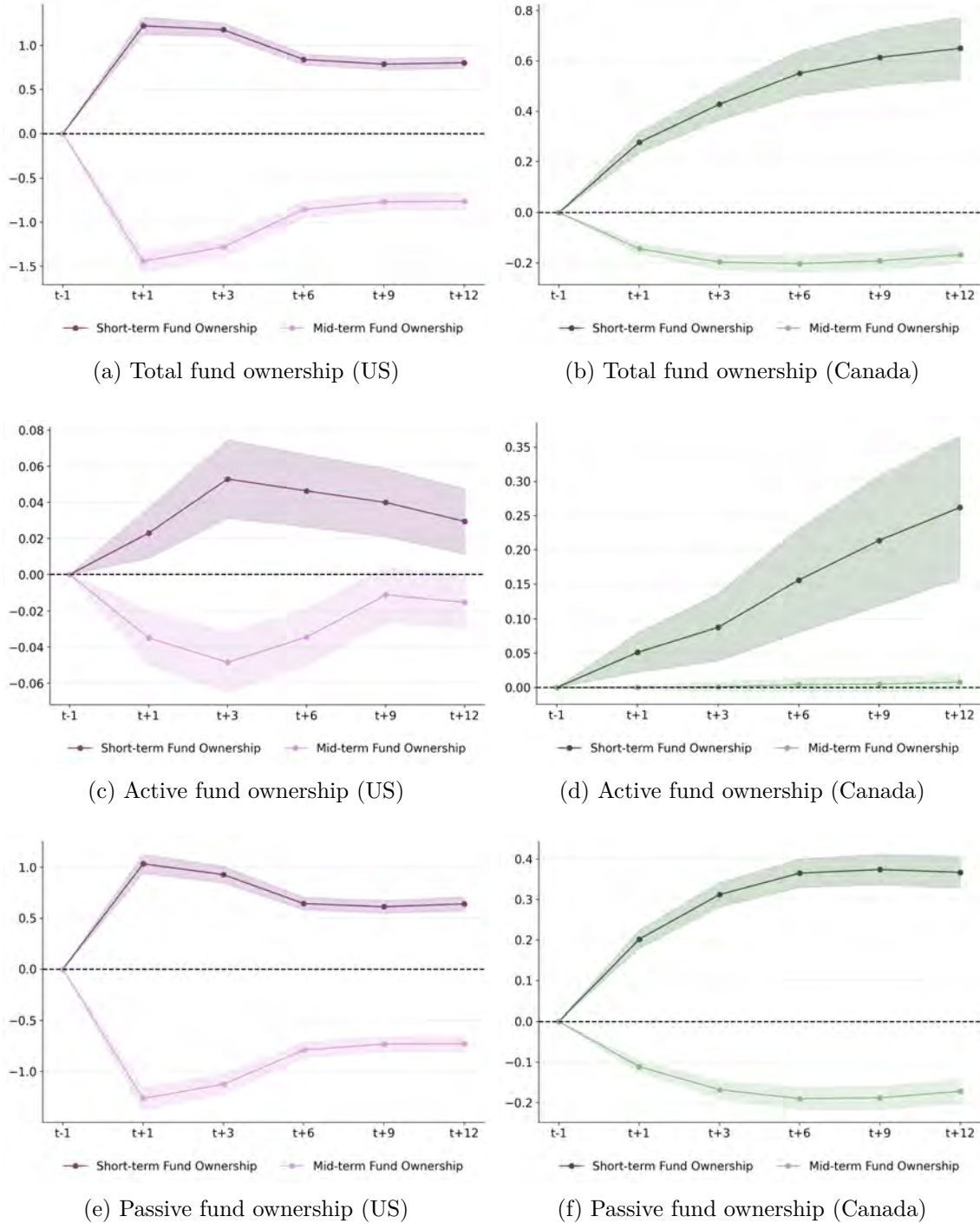
According to our model in Section 3, equilibrium yields are lower for bonds with larger sBMIs because funds hedge against underperforming their benchmark indexes. The model predicts that funds, both active and passive, rebalance in line with their mandates and, therefore, their assets contribute to BMI. In this section, to complement the stylized facts in Section 2, we document that active and passive fund ownership changes around maturity cutoffs as predicted by BMI and show that an increase in sBMI leads to an increase in both types of ownership in our sample.

First, the average changes in ownership around maturity cutoffs mask considerable heterogeneity stemming from maturity-based benchmarks. In Appendix B.5, we report how ownership changes for bonds crossing each of the cutoffs on average. Even when significantly different from zero, those average changes do not always go in the same direction as BMI changes, especially sparse ones. Our data allows us to characterize changes in ownership of funds with maturity-based benchmarks specifically. To do so, we estimate a counterpart of regression (23) around each maturity cutoff in both U.S. and Canada using benchmarked fund ownership as dependent variable. We present the estimates for the 5-year cutoff in Figure 4 and the corresponding figures for the 3- and 10-year cutoffs are reported in Appendix B.7. The figure highlights that, when a bond crosses the 5-year cutoff, it is sold by funds benchmarked to the mid-term term or 5–10 year indexes and it is purchased by funds benchmarked to short-term or 1–5 year indexes. Importantly, we see very similar picture for both active and passive funds in both geographies we study.⁷

Second, our model predicts that fund portfolios are sparse, and hence estimating

⁷An exception is that mid-term active funds in Canada seem to not necessarily sell bonds crossing the 5-year cutoff, which suggests that, in light of larger transaction costs in Canada, these funds may hold such bonds to maturity.

Figure 4: Changes in benchmarked ownership around the 5-year maturity cutoff



This figure illustrates how benchmarked fund ownership of corporate bonds changes around the 5-year maturity cutoffs in Canada and the United States. We report total fund ownership changes of funds with short (1–5 years) and mid-term (5–10 years) benchmark indexes, as well as their active and passive components. All estimates and confidence intervals are from estimating specification (23) around the 5-year cutoff. All y-axes correspond to p.p. changes in ownership in response to the move across the 5-year cutoff. We report 90% confidence intervals based on standard errors clustered by issuer and year-month.

the effect of benchmarking on rebalancing is challenging, as an average effect of crossing the cutoff across all bonds is attenuated to zero because some bonds are excluded from the portfolios. Studying how ownership changes with sparse BMI overcomes this challenge. Table 5 presents the results of estimating the regression (23) for total, active and passive fund ownership and documents that all ownership significantly increases with an increase in sBMI. Different sBMIs have varying success but sparsity is crucial to pick up rebalancing of active funds. The table reports the changes in the 6-month window to account for delayed rebalancing, and figures in Appendix B.8 illustrate the estimates for all horizons between 1 and 12 months.

Despite their statistical significance, the effects of sparse BMI shocks on ownership are quite modest economically. In the U.S., a one-standard-deviation increase in the rating-based sparse BMI (around 1.5 units) is associated with a roughly 4 bps increase in total ownership (0.028 by 1.5 p.p.) over the six-month window, relative to an average total ownership change standard deviation of about 5.2 p.p.. This translates to roughly 0.8% of the cross-sectional volatility in ownership changes. Breaking this down, roughly half comes from the increase in active ownership and half from the increase in passive ownership. The magnitudes are similar in Canada. For example, a one-standard-deviation increase in the rating-based sparse BMI (around 0.6 units) is associated with a 4 bps rise in total ownership (0.073 by 0.6 p.p.), which is primarily driven by the rise in active ownership.

4.5 Discussion

Our empirical design and findings contribute to the growing literature on institutional demand in fixed income markets, while offering several novel angles relative to prior work. In this section, we discuss these differences and also address several potential caveats in our research design.

Our approach builds on the maturity-based discontinuities in corporate bond markets first documented by [Bretscher, Schmid, and Ye \(2024\)](#), but introduces two key extensions. First, we exploit shock intensity, allowing us to recover both the direction and magnitude of

Table 5: Changes in ownership and sparse BMIs (United States and Canada)

	Total BMI		Sparse BMI			
		Rating	Segment– Rating	Segment– Rating– Maturity	Sector– Rating	Sector– Rating– Maturity
Panel A: United States						
<i>A1. Total ownership</i>						
$t-6$ to $t+6$	0.0669*** (0.0110)	0.0281*** (0.0060)	0.0294*** (0.0049)	0.0208*** (0.0041)	0.0180*** (0.0037)	0.0132*** (0.0026)
Observations	520,572	520,572	520,572	520,572	520,572	520,572
R-squared	0.204	0.204	0.204	0.204	0.204	0.204
<i>A2. Active ownership</i>						
$t-6$ to $t+6$	0.011 (0.0095)	0.0113** (0.0052)	0.0145*** (0.0042)	0.0101*** (0.0036)	0.0098*** (0.0034)	0.0085*** (0.0025)
Observations	520,572	520,572	520,572	520,572	520,572	520,572
R-squared	0.204	0.204	0.204	0.204	0.204	0.204
<i>A3. Passive ownership</i>						
$t-6$ to $t+6$	0.0561*** (0.0058)	0.0164*** (0.0021)	0.0147*** (0.0018)	0.0105*** (0.0014)	0.0081*** (0.0010)	0.0046*** (0.0006)
Observations	520,572	520,572	520,572	520,572	520,572	520,572
R-squared	0.287	0.286	0.286	0.286	0.285	0.285
Panel B: Canada						
<i>B1. Total ownership</i>						
$t-6$ to $t+6$	0.529*** (0.135)	0.0734*** (0.0221)	0.0561*** (0.0177)	0.0260** (0.0112)		
Observations	122,441	122,441	122,441	122,441		
R-squared	0.534	0.534	0.534	0.534		
<i>B2. Active ownership</i>						
$t-6$ to $t+6$	0.414*** (0.123)	0.0612*** (0.0198)	0.0454*** (0.0153)	0.0200** (0.0100)		
Observations	122,441	122,441	122,441	122,441		
R-squared	0.509	0.509	0.509	0.509		
<i>B3. Passive ownership</i>						
$t-6$ to $t+6$	0.113*** (0.0249)	0.0127*** (0.00361)	0.0112*** (0.00351)	0.00618*** (0.00208)		
Observations	122,441	122,441	122,441	122,441		
R-squared	0.348	0.347	0.347	0.347		

This table reports the $\Delta sBMI$ coefficient estimates from regression (23) for the U.S. (Panel A) and Canadian (Panel B) samples. The dependent variables are change in fund ownership (total, active, and passive) in percentage points. All regressions control for the logarithm of the issue amount outstanding, numeric credit rating, bid–ask spread quintile, and time to maturity, and include bond and year–month fixed effects. Standard errors, clustered by issuer and year–month, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results for Canada omit sector-based sparse BMIs (blank cells).

the demand shock, and capturing cases where the discontinuity differs across maturities, fund types, and time. Second, our analysis includes the Canadian corporate bond market, and we focus on funds benchmarked to indexes within Bloomberg Aggregate Bond Index family in the U.S., though we confirm in robustness checks that results remain similar when expanding the sample to include funds with other benchmarks. Finally, our analysis concentrates on the cross-section of corporate bonds within the Bloomberg aggregate family, which effectively excludes some smaller and less liquid bonds. This is the segment of the market for which we have the most comprehensive data on benchmarked capital tracking the bonds.

Unlike much of the existing literature that emphasizes passive ownership, our results point to a broader mechanism of benchmark-driven demand, where both passive and active funds rebalance in response to exogenous changes in their benchmark indexes. Our findings also highlight the role of portfolio sparsity: funds do not hold the full index, but their trading still co-moves with benchmark changes, making the effects of sparse benchmarking a central driver of price and ownership dynamics in our paper.

Although the estimated effects on prices and ownership are small in absolute terms, we are able to detect them precisely because the design leverages a large number of maturity-based shocks across bonds and time. Each time a bond crosses a maturity cutoff, the amount of capital benchmarked to it—and thus the corresponding sparse BMI—adjusts, creating thousands of incremental shocks across the panel. This rich cross-sectional and time-series variation, observed over more than a decade of monthly data, allows us to exploit the intensity of these maturity-driven shocks, even though the sparsity is measured imperfectly and subject to some noise. In other words, while any single maturity transition generates only a small shift in ownership, the number of these events provides enough statistical power to reveal consistent and economically meaningful effects and potentially allow for estimation of, for example, price elasticities of demand in corporate bond markets.

One striking feature of the demand shocks in our setting is that the timing of a maturity crossing is perfectly predictable. Market participants can observe a bond’s remaining maturity and anticipate when it will enter or exit a given maturity bucket. This stands in

contrast with the index inclusion effects in equity markets, where index additions and deletions need to be predicted. However, the magnitude of the demand shock is perhaps even harder to evaluate in advance, as it depends not only on the evolution of index AUM and weights but also on which bonds are included in the sparse portfolios of individual funds.

Corporate bond markets are characterized by limited liquidity and infrequent trading, which raises concerns about stale prices and delayed price discovery. To mitigate these issues, we (i) use wide event windows to capture gradual adjustments, and (ii) control for liquidity proxies such as bid–ask spread quintiles in all specifications. Using the level of spreads or omitting these controls produces nearly identical results, suggesting that our findings are not driven by liquidity effects. Consistent with industry practice, we rely on index provider prices from Bloomberg and FTSE, which are widely used for portfolio valuation by institutional investors and have been adopted in academic studies.

There are three potential caveats related to the breadth of our baseline panel. First, around the 10-year cutoff, insurance companies and other long-horizon investors often reduce holdings for portfolio management or regulatory reasons (see [Li \(2023\)](#) and [Chaudhary \(2024\)](#)), which could confound the interpretation of our findings. Second, at the 3-year cutoff, many bonds are already close to maturity and often held to maturity by institutional investors, reducing turnover and dampening the observed price and ownership response. Third, bonds at the extreme ends of the maturity spectrum, such as 2-year bonds and 30-year bonds, may be hard to compare in a clean way in a linear regression framework. In [Appendix B.6](#), we present results focusing specifically on bonds around the 5-year cutoff. Estimates are similar but statistically weaker, likely because of the reduced test power.

In [Appendix B.9](#), we also document sharp ownership discontinuities around benchmark size-eligibility cutoffs in both Canada and the United States. For the smallest bonds, there is a discrete jump in fund ownership exactly at the threshold for benchmark inclusion. This pattern, previously shown for passive funds in fixed-income markets (e.g., [Dathan and Davydenko \(2025\)](#)), appears to hold for active funds as well and provides additional evidence in support of our model.

5 Concluding Remarks

In this paper, we present evidence from Canada and the United States that corporate bond ownership by both active and passive corporate bond funds is strongly influenced by their mandates, and in particular their benchmark indexes. The tracking errors of active funds in our comprehensive sample of funds and their benchmarks are not substantially higher than those of passive, indicating that the two classes of funds still stay relatively close to their benchmarks. Accordingly, we find that benchmark weights strongly predict active and passive fund ownership.

We document that corporate bond funds' portfolios are sparse, with an average fund in Canada holding only 15% of bonds in its benchmark and in the U.S. only 7%. We hypothesize that the reason for this sparsity is a high turnover of bond indexes and high rebalancing costs, as well as other portfolio management costs. We introduce portfolio management costs and benchmarking considerations in a simple equilibrium asset pricing model and show that it produces sparse portfolios that include assets with higher benchmark weights, higher market capitalization, and higher idiosyncratic risk within each bond category. The model also implies that bond prices depend on sparse benchmarking intensity, a measure of inelastic benchmark-driven demand of fund managers. Exploiting plausibly exogenous within-bond changes in sparse BMI driven by reallocations of benchmarked capital, we document that increases in bonds' sparse BMIs lead to reductions in bond yields and increases in active and passive fund ownership, in line with the predictions of our model.

Whereas numerous studies in the literature documented index inclusion effects in equity markets, there are few corresponding studies for bonds. The lack of index effects in bonds is typically interpreted as evidence that bonds are more substitutable than equities. While we do not dispute this interpretation, we offer an alternative explanation for the difficulties in documenting index effects in bonds: funds' portfolio sparsity. According to our theory, index effects should be present only in bonds with non-zero sparse BMIs, and our empirical findings confirm this implication. Corporate bond markets are not the only markets in which we observe portfolio sparsity. Using sparse BMI for measuring index effects

in other markets, especially in markets with higher rebalancing costs, may help studies that aim to document various effects of institutional investors on asset prices.

A growing literature highlights the potential financial fragility arising from bond mutual funds, which engage in liquidity transformation—offering daily redemptions while holding illiquid assets—and may be forced to sell illiquid corporate bonds to meet investor outflows (see, e.g., [Goldstein, Jiang, and Ng, 2017](#); [Jiang, Li, Sun, and Wang, 2022](#)). Our findings add nuance to this concern. We show that mutual fund holdings are highly sparse and concentrated in bonds with higher sparse BMIs. These bonds—being more widely held across mutual funds—are precisely those most exposed to coordinated sales pressures during periods of outflows. In contrast, bonds with low sparse BMIs, which are largely held by investors less exposed to fund outflows, are less likely to be subject to such fire-sale dynamics. A natural next step in this agenda is to examine whether firms account for potential financial fragility when making their debt issuance decisions ([Mota and Siani, 2025](#)).

Our sparse BMI measure offers a way to quantify the inelastic demand in bond markets stemming from fund manager mandates for both active and passive funds. One of the methods to estimate price elasticities of demand has been to use changes in *passive* bond ownership around index rebalancing events as demand shifters. Our results demonstrate that demand of active corporate bond funds also contributes to these demand shifters. Given that active funds manage more assets—particularly in the earlier part of our sample—they represent an important component of benchmark-driven demand and thus contribute significantly to these aggregate effects. One future avenue for research is to revisit demand elasticity estimates in light of these findings. Furthermore, a growing strand of research in corporate finance shows that passive ownership can lower firms’ cost of bond financing and influence real economic outcomes (e.g., [Dathan and Davydenko, 2025](#)). Our findings indicate that active bond mutual funds behave similarly to passive funds in that they closely follow their mandates. As a result, it may be challenging to fully disentangle the effects of passive ownership from those attributable to mutual funds and ETFs more broadly.

References

- Basak, Suleyman, and Anna Pavlova, 2013, Asset Prices and Institutional Investors, *American Economic Review* 103, 1728–1758.
- Brennan, Michael, 1993, Agency and Asset Pricing, Working paper UCLA.
- Bretscher, Lorenzo, Lukas Schmid, Ishita Sen, and Varun Sharma, 2023, Institutional Demand and Corporate Bond Prices, Working Paper.
- Bretscher, Lorenzo, Lukas Schmid, and Tiange Ye, 2024, Passive Demand and Active Supply: Evidence from Maturity-mandated Corporate Bond Funds, Working Paper.
- Brogaard, Jonathan, Davidson Heath, and Da Huang, 2024, ETF sampling and index arbitrage, Working paper.
- Buffa, Andrea, Dimitri Vayanos, and Paul Woolley, 2022, Asset Management Contracts and Equilibrium Prices, *Journal of Political Economy* 130, 2148–2205.
- Cai, Yuhua, Anna Helmke, Benjamin Mosk, and Felix Suntheim, 2025, Fund Investor Types and Bond Market Volatility, Global Financial Stability Note, International Monetary Fund.
- Chaudhary, Manav, 2024, Regulator beliefs, .
- , Zhiyu Fu, and Jian Li, 2024, Corporate Bond Multipliers: Substitutes Matter, Working Paper.
- Dannhauser, Caitlin D., and Saeid Hoseinzade, 2022, The Unintended Consequences of Corporate Bond ETFs: Evidence from the Taper Tantrum, *Review of Financial Studies* 35, 51–90.
- Dathan, Michele, and Sergei Davydenko, 2025, Debt issuance in the era of passive investment, *Management Science*.
- Dick-Nielsen, Jens, and Marco Rossi, 2019, The Cost of Immediacy for Corporate Bonds, *Review of Financial Studies* 32, 1–41.
- Dow, James, and Sérgio R. da Costa Werlang, 1992, Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio, *Econometrica* 60, 197–204.
- Gabaix, Xavier, Ralph S. J. Koijen, Robert J. Richmond, and Motohiro Yogo, 2025, Up-

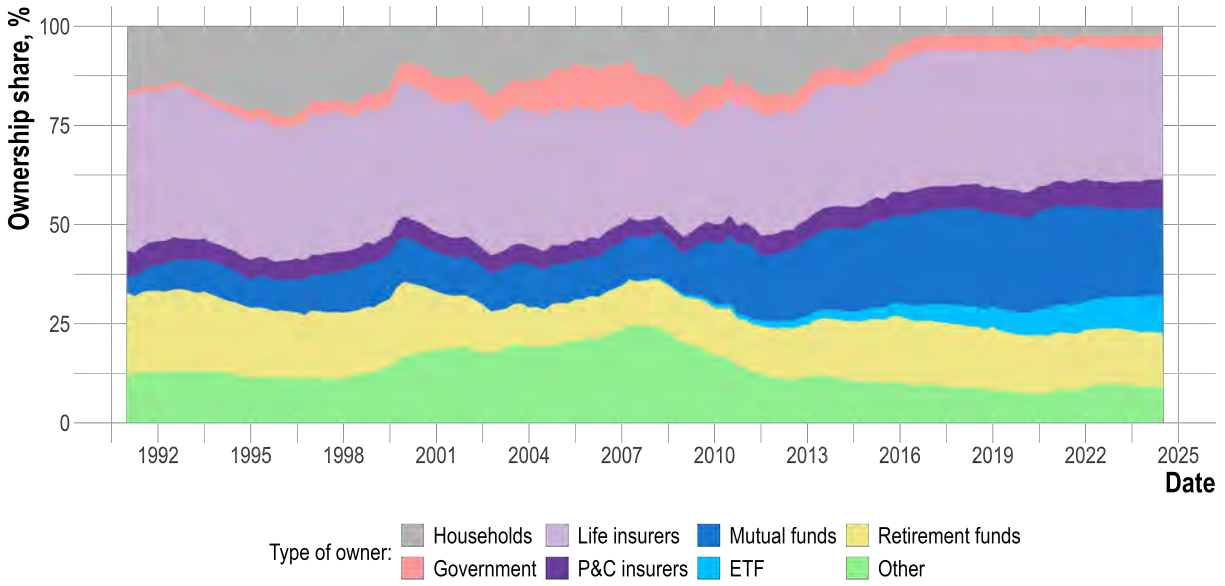
- grading Credit Pricing and Risk Assessment through Embeddings, Working paper.
- Goldstein, Itay, Hao Jiang, and David T. Ng, 2017, Investor flows and fragility in corporate bond funds, *Journal of Financial Economics* 126, 592–613.
- Graves, Daniel, 2024, What Lies Beneath Zero: Censoring, Demand Estimation, and Hidden Beliefs, Working Paper.
- He, Zhiguo, and Wei Xiong, 2013, Delegated Asset Management and Investment Mandates, *Journal of Finance* 68, 719–765.
- Huang, Shiyang, Xu Lu, Yang Song, and Hong Xiang, 2023, Remeasuring scale in active management, *SSRN Electronic Journal*.
- Jiang, Hao, Yi Li, Zheng Sun, and Ashley Wang, 2022, Does mutual fund illiquidity introduce fragility into asset prices? evidence from the corporate bond market, *Journal of Financial Economics* 143, 277–302.
- Kashyap, Anil, Natalia Kovrijnykh, Jian Li, and Anna Pavlova, 2021, The Benchmark Inclusion Subsidy, *Journal of Financial Economics*, 142, 756–774.
- Koijen, Ralph S.J., and Motohiro Yogo, 2019, A Demand System Approach to Asset Pricing, *Journal of Political Economy* 127, 1475–1515.
- Koont, Naz, Yiming Ma, Lubos Pastor, and Yao Zeng, 2025, Steering a Ship in Illiquid Waters: Active Management of Passive Funds, forthcoming, *Review of Financial Studies*.
- Li, Jessica S., Maureen O’Hara, Andreas C. Rapp, and Xing Alex Zhou, 2025, Bond market illiquidity: Is portfolio trading the solution?, Discussion paper, Working paper.
- Li, Ziang, 2023, The long-term interest rate and corporate bond credit spreads, *SSRN Electronic Journal*.
- Ma, Linlin, Yuehua Tang, and Juan-Pedro Gomez, 2019, Portfolio manager compensation in the u.s. mutual fund industry, *The Journal of Finance* 74, 587–638.
- Ma, Yiming, Kairong Xiao, and Yao Zeng, 2022, Mutual Fund Liquidity Transformation and Reverse Flight to Liquidity, *Review of Financial Studies* 35, 4674–4711.
- Merton, Robert C., 1987, A Simple Model of Capital Market Equilibrium with Incomplete Information, *Journal of Finance* 42, 483–510.

- Mota, Lira, and Kerry Y. Siani, 2025, Financially sophisticated firms, Working Paper, MIT Sloan School of Management.
- Ottonello, Giorgio, 2019, The Impact of Benchmarking in Fixed-Income Markets , Working Paper.
- Pavlova, Anna, and Taisiya Sikorskaya, 2023, Benchmarking Intensity, *Review of Financial Studies* 36, 859–903.
- Shim, J. J., and K. Todorov, 2022, ETFs, Illiquid Assets, and Fire Sales, Working paper.
- Shin, Sean Seunghun, Xing (Alex) Zhou, and Qifei Zhu, 2025, The Growing Index Effect in the Corporate Bond Market, Working paper.
- Van Nieuwerburgh, Stijn, and Laura Veldkamp, 2010, Information Acquisition and Under-Diversification, *Review of Economic Studies* 77, 779–805.

A Appendix: Stylized Facts

A.1 Ownership of U.S. Corporate Bonds

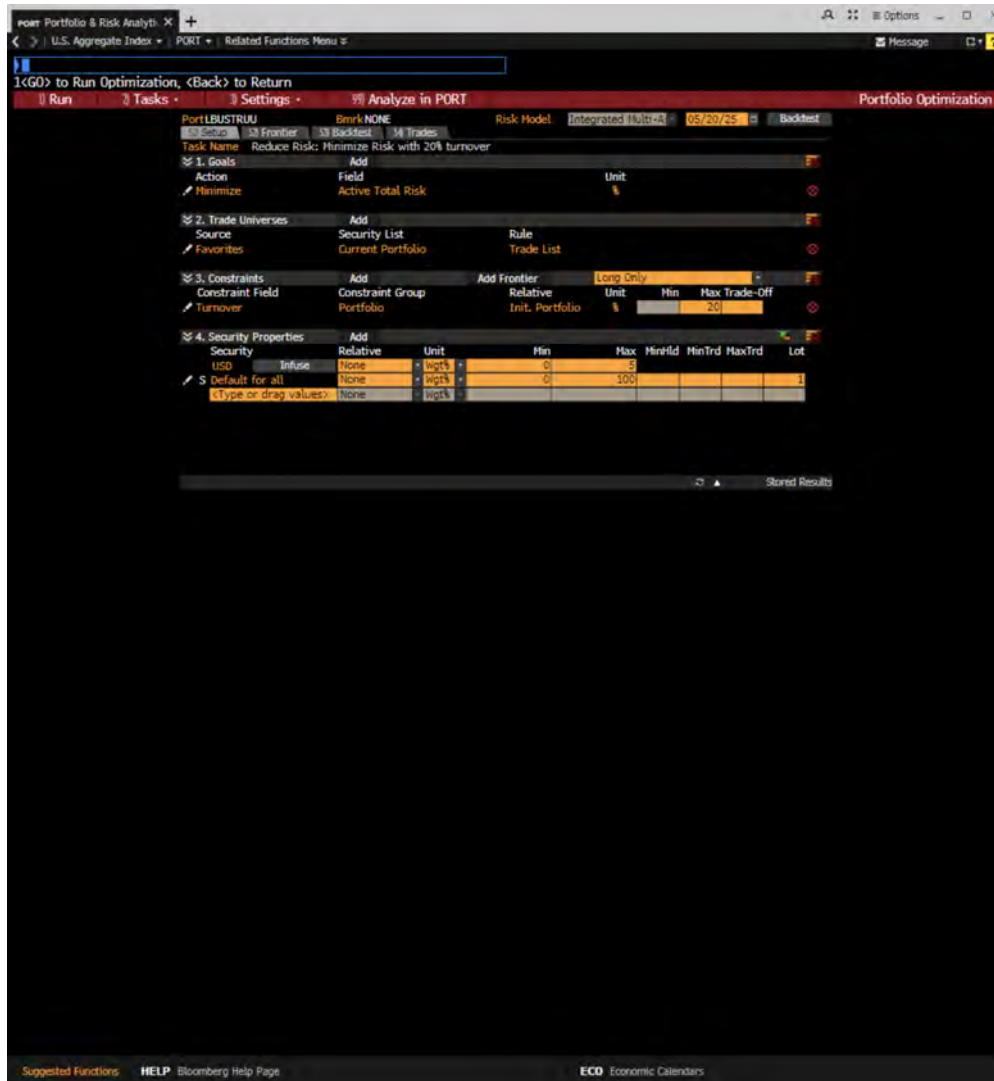
Figure A1: Ownership of U.S. Corporate Bonds by Domestic Investors



This figure illustrates the evolution of U.S. corporate bond holdings by domestic investor categories from 1990 to 2024. The data, sourced from the Federal Reserve Board's Financial Accounts of the U.S. (Z.1 release), exclude holdings by foreign investors and foreign monetary authorities. The depicted categories encompass households, mutual funds, pension funds, insurance companies, banks, and other domestic financial institutions. The data series can be accessed at: <https://www.federalreserve.gov/releases/z1/default.htm>.

A.2 Bloomberg portfolio allocation tool

Figure A2: Bloomberg portfolio allocation tool (PORT)



(a) Bloomberg PORT optimizer.

(Continued on the next page.)



(b) Bloomberg PORT with corporate bond classifications.

This figure presents the screenshots of Bloomberg portfolio allocation tool. Subfigure (a) demonstrates the tradeoff between tracking error and turnover which the tool defaults to, and subfigure (b) highlights the saliency of Bloomberg industry classifications.

A.3 Fund samples summary statistics

Table A1: Descriptive statistics for our Canadian sample of funds

Year	Counts			Portfolio size (billion CAD)			Avg. # Holdings		
	All	Passive	ETF	All	Passive	ETF	All	Passive	ETF
2003	61	4	0	35.8	9.7	0.0	85.0	391.3	–
2004	62	6	2	42.2	10.8	0.4	90.0	288.5	1.0
2005	64	5	1	48.3	11.2	0.4	97.1	345.4	1.0
2006	72	5	1	53.8	10.8	0.5	99.7	365.0	56.0
2007	81	9	5	65.1	11.9	1.6	108.4	250.7	96.4
2008	88	9	5	62.2	12.2	2.3	110.7	278.2	148.8
2009	94	9	5	80.3	15.4	4.3	129.2	332.3	209.0
2010	103	12	8	98.6	16.5	5.2	136.3	293.2	184.6
2011	113	15	11	106.8	16.5	5.8	136.9	273.5	178.3
2012	119	17	13	132.9	20.4	8.4	153.4	301.2	214.1
2013	133	18	14	140.7	19.2	7.3	150.5	304.2	214.4
2014	140	21	17	155.7	19.3	7.9	152.7	272.0	192.3
2015	149	23	19	172.4	21.9	10.0	165.0	272.7	199.0
2016	155	23	19	194.9	24.8	12.8	175.2	292.8	223.5
2017	158	26	22	198.9	26.6	15.0	177.1	283.2	220.0
2018	167	30	26	194.4	27.8	16.0	181.8	299.5	246.7
2019	177	35	31	216.6	34.7	22.1	204.8	299.7	254.9
2020	185	41	36	247.5	39.6	26.0	237.6	298.0	256.1
2021	192	45	40	264.6	41.6	28.5	258.5	309.2	270.6
2022	193	44	39	236.2	36.3	25.9	257.0	335.3	299.0
2023	198	46	41	246.1	39.3	29.9	252.3	335.3	299.8

Portfolio size includes bond holdings only. We report the number of fund portfolios, not unique share classes. Average number of holdings is bond count per portfolio at year-end. A dash indicates not applicable.

Table A2: Descriptive statistics for our U.S. sample of funds

Year	Counts			TNA (billion USD)			Avg. # Holdings		
	All	Passive	ETF	All	Passive	ETF	All	Passive	ETF
2012	322	41	26	2242.685	325.522	212.590	494.0	1072.8	925.7
2013	344	44	28	2139.409	339.701	219.573	522.0	1169.7	1054.2
2014	356	45	29	2156.750	415.081	272.908	540.8	1320.6	1251.0
2015	382	53	37	2120.953	452.882	310.035	563.0	1301.1	1246.5
2016	410	58	41	2366.264	543.013	373.542	563.9	1267.1	1189.7
2017	424	58	45	2741.912	652.338	446.594	583.3	1310.3	1184.3
2018	436	70	53	2732.357	680.011	458.905	593.2	1270.4	1180.2
2019	456	75	57	3232.458	813.755	532.020	601.2	1265.8	1146.3
2020	467	80	62	3735.388	997.250	673.113	682.1	1394.6	1252.8
2021	464	80	64	3952.417	1093.153	726.867	715.2	1397.6	1266.7
2022	447	79	64	3204.082	959.411	629.991	717.3	1413.1	1275.0
2023	439	80	65	3491.727	1099.141	719.353	742.4	1504.6	1360.1

TNA figures include matched separate accounts. We report the number of fund portfolios, not unique share classes. Average number of holdings is corporate bond count per portfolio at year-end.

A.4 Bond samples summary statistics

Table A3: Descriptive statistics for the Canadian sample

Variable	N	Mean	SD	Min	p10	p50	p90	Max
Panel A: Prices and returns								
Yield	124,110	3.807	1.408	0.785	1.990	3.766	5.584	8.164
Yield spread	124,110	1.444	0.766	0.186	0.593	1.340	2.362	4.953
Yield change	124,110	-0.007	0.235	-0.705	-0.275	-0.011	0.264	0.757
Yield spread change	124,110	-0.004	0.144	-0.594	-0.123	-0.006	0.114	0.730
Total return	118,663	0.004	0.020	-0.068	-0.017	0.003	0.025	0.080
Panel B: BMI levels								
BMI	124,110	10.315	2.594	4.742	6.707	10.004	13.083	16.447
Sparse BMI	124,110	4.144	5.122	0.000	0.000	0.002	12.185	16.280
Sparse BMI (Rating)	124,110	4.764	5.291	0.000	0.000	1.167	12.497	16.285
Sparse BMI (Industry Rating)	124,110	4.787	5.315	0.000	0.000	1.484	12.523	16.332
Sparse BMI (Sector Rating)	124,110	4.858	5.316	0.000	0.000	2.174	12.578	16.285
Sparse BMI (Rating Maturity)	124,110	4.778	5.416	0.000	0.000	0.000	12.627	16.285
Sparse BMI (Industry Rating Maturity)	124,110	4.866	5.458	0.000	0.000	0.000	12.653	16.332
Sparse BMI (Sector Rating Maturity)	124,110	4.896	5.479	0.000	0.000	0.000	12.692	16.332
Panel C: BMI changes								
BMI	123,765	0.037	0.306	-0.656	-0.224	-0.007	0.309	1.838
Sparse BMI	123,765	0.008	0.436	-3.410	-0.115	0.000	0.142	3.401
Sparse BMI (Rating)	123,765	0.007	0.634	-5.324	-0.133	0.000	0.169	4.885
Sparse BMI (Industry Rating)	123,765	0.004	0.691	-5.998	-0.130	0.000	0.163	5.206
Sparse BMI (Sector Rating)	123,765	0.002	0.923	-7.318	-0.139	0.000	0.175	6.681
Sparse BMI (Rating Maturity)	123,765	0.002	0.862	-7.992	-0.130	0.000	0.165	6.863
Sparse BMI (Industry Rating Maturity)	123,765	0.001	0.977	-8.890	-0.130	0.000	0.162	8.043
Sparse BMI (Sector Rating Maturity)	123,765	0.005	1.327	-10.346	-0.132	0.000	0.166	9.994
Panel D: Ownership levels								
Total ownership	124,110	11.580	9.173	0.000	1.657	9.320	24.786	100.000
Active ownership	124,110	9.554	8.776	0.000	0.221	7.194	22.279	100.000
Passive ownership	124,110	2.027	1.378	0.000	0.500	1.870	3.642	17.396
Panel E: Ownership changes								
Total ownership	124,110	-0.058	0.880	-8.420	-0.428	0.000	0.367	10.105
Active ownership	124,110	-0.054	0.790	-7.213	-0.404	0.000	0.331	9.102
Passive ownership	124,110	-0.002	0.154	-1.370	-0.069	0.000	0.074	1.539
Panel F: Other bond characteristics								
log(Amount outstanding)	124,110	19.384	0.769	16.013	18.421	19.337	20.367	21.822
Rating (numeric)	124,110	1.719	0.697	1.000	1.000	2.000	3.000	4.000
Time to maturity (years)	124,110	11.124	8.782	1.500	2.201	7.181	25.142	29.999
Bid-ask spread (Refinitiv)	89,478	0.512	2.041	0.000	0.144	0.467	0.826	198.80

This table reports summary statistics of the main variables used in analysis for Canadian corporate bond monthly sample in 2004–2023. All variables in panels A to E are in percentage points. All sparse BMIs are above median.

Table A4: Descriptive statistics for the U.S. sample

Variable	N	Mean	SD	Min	p10	p50	p90	Max
Panel A: Prices and returns								
Yield	605,914	3.520	1.476	0.190	1.580	3.470	5.460	16.680
Yield spread	605,914	1.294	0.764	-0.223	0.528	1.157	2.193	15.209
Yield change	605,914	0.015	0.349	-5.420	-0.320	0.010	0.360	9.940
Yield spread change	605,914	-0.006	0.287	-5.341	-0.201	-0.008	0.169	10.536
Total return	605,914	0.204	2.688	-41.370	-2.520	0.190	2.760	31.630
Panel B: BMI levels								
BMI	605,914	17.059	3.008	12.531	13.795	16.611	21.752	28.369
Sparse BMI	605,914	9.435	8.126	0.000	0.000	11.096	17.916	28.329
Sparse BMI (Segment)	605,914	8.414	8.688	0.000	0.000	4.700	17.993	28.330
Sparse BMI (Industry)	605,914	8.454	8.628	0.000	0.000	5.658	17.957	28.330
Sparse BMI (Sector)	605,914	8.526	8.528	0.000	0.000	10.751	17.913	28.329
Sparse BMI (Rating)	605,914	8.861	8.451	0.000	0.000	11.217	17.956	28.329
Sparse BMI (Segment Rating)	605,914	8.413	8.599	0.000	0.000	5.416	17.970	28.329
Sparse BMI (Industry Rating)	605,914	8.479	8.576	0.000	0.000	9.768	17.934	28.331
Sparse BMI (Sector Rating)	605,914	8.465	8.455	0.000	0.000	10.226	17.868	28.329
Sparse BMI (Rating Maturity)	605,914	8.785	8.576	0.000	0.000	10.800	17.983	28.329
Sparse BMI (Segment Rating Maturity)	605,914	8.456	8.761	0.000	0.000	0.000	18.019	28.329
Sparse BMI (Industry Rating Maturity)	605,914	8.472	8.757	0.000	0.000	0.000	18.016	28.332
Sparse BMI (Sector Rating Maturity)	605,914	8.395	8.736	0.000	0.000	0.000	17.980	28.332
Panel C: BMI changes								
BMI	605,914	0.128	0.924	-2.488	-0.440	0.015	0.522	7.504
Sparse BMI	605,914	0.014	1.559	-12.749	-0.359	0.000	0.363	10.953
Sparse BMI (Segment)	605,914	0.011	1.453	-17.156	-0.284	0.000	0.311	16.251
Sparse BMI (Industry)	605,914	0.004	1.614	-17.151	-0.293	0.000	0.313	15.789
Sparse BMI (Sector)	605,914	0.008	1.808	-17.133	-0.319	0.000	0.334	16.304
Sparse BMI (Rating)	605,914	0.008	1.495	-16.515	-0.336	0.000	0.346	15.039
Sparse BMI (Segment Rating)	605,914	0.012	1.636	-17.136	-0.294	0.000	0.320	15.633
Sparse BMI (Industry Rating)	605,914	0.005	1.819	-17.151	-0.310	0.000	0.327	17.223
Sparse BMI (Sector Rating)	605,914	0.006	2.044	-17.134	-0.340	0.000	0.350	17.960
Sparse BMI (Rating Maturity)	605,914	0.004	1.661	-21.939	-0.305	0.000	0.326	16.118
Sparse BMI (Segment Rating Maturity)	605,914	-0.002	1.985	-21.978	-0.269	0.000	0.303	18.097
Sparse BMI (Industry Rating Maturity)	605,914	0.003	2.286	-20.723	-0.278	0.000	0.310	23.152
Sparse BMI (Sector Rating Maturity)	605,914	0.013	2.766	-20.501	-0.291	0.000	0.322	19.642
Panel D: Ownership levels								
Total ownership	605,914	11.806	11.212	0.000	2.597	8.635	24.324	100.000
Active ownership	605,914	7.437	10.726	0.000	0.002	3.757	18.882	100.000
Passive ownership	605,914	4.376	2.459	0.000	1.424	4.041	7.695	43.934
Panel E: Ownership changes								
Total ownership	605,914	-0.034	1.331	-21.184	-0.449	0.000	0.398	91.094
Active ownership	605,914	-0.062	1.300	-20.353	-0.374	0.000	0.212	93.520
Passive ownership	605,914	0.030	0.240	-5.403	-0.080	0.002	0.178	4.226
Panel F: Other bond characteristics								
log(Amount outstanding)	605,914	13.384	0.629	12.429	12.612	13.304	14.221	16.250
Rating (numeric)	605,914	7.388	1.898	1.000	5.000	8.000	10.000	10.000
Time to maturity (years)	605,914	11.288	8.993	1.510	2.500	7.290	26.240	29.920
Bid-ask spread (WRDS)	454,062	0.388	0.382	0.000	0.080	0.280	0.800	22.140

This table reports summary statistics of the main variables used in analysis for United States corporate bond monthly sample in 2012–2023. All variables in panels A to E are in percentage points. All sparse BMIs are above median.

A.5 Fund tracking errors

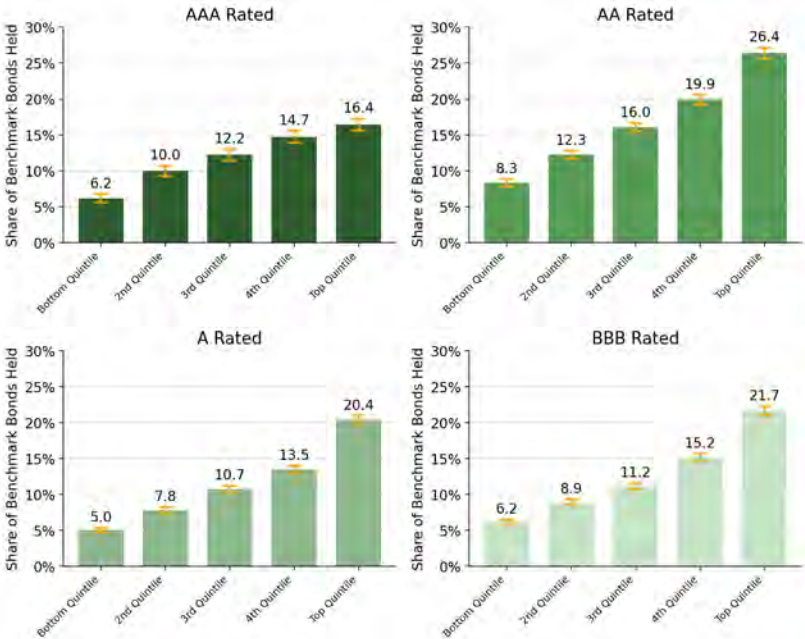
Table A5: Tracking errors of bond funds

	Rolling 3-year TE			Rolling 5-year TE			N
	Mean	Median	St. dev.	Mean	Median	St. dev.	
Panel A: US							
Active Fund	2.070	1.377	2.087	2.064	1.445	1.845	57,393
Passive Fund	0.922	0.326	1.251	0.798	0.302	1.075	10,257
Panel B: Canada							
Active Fund	1.164	0.686	1.290	1.195	0.773	1.240	25,887
Passive Fund	0.590	0.175	0.930	0.452	0.153	0.678	4,221

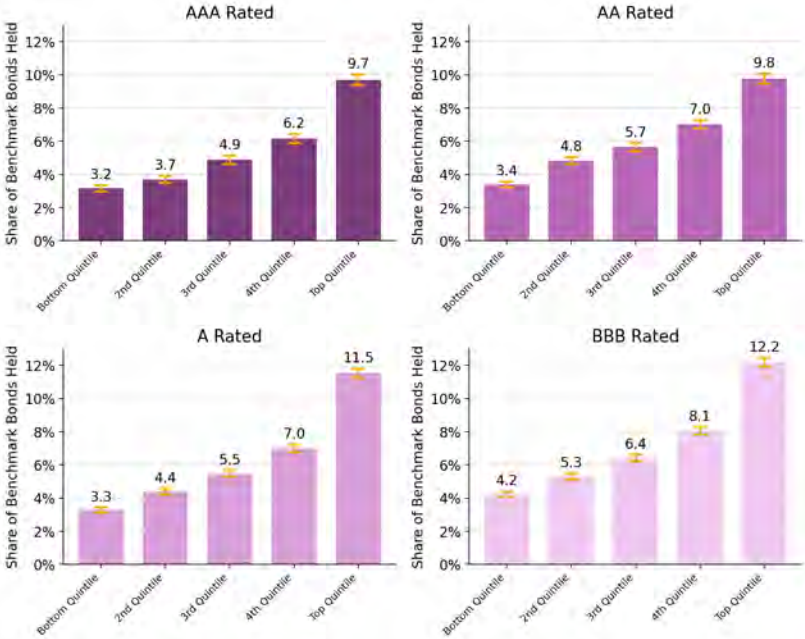
This table reports tracking errors (TE) of funds in our US (panel A) and Canadian (panel B) samples, all in percent. TEs are computed from monthly fund return deviations from its benchmark return using rolling windows: “3-year” is a 36-month window and “5-year” is a 60-month window. N reports fund/portfolio-month observations over which statistics are computed.

A.6 Share of benchmark bonds held across benchmark weight quintiles by credit ratings

Figure A3: Share of benchmark bonds held across benchmark weight quintiles by credit ratings



(a) Canada



(b) US

This figure illustrates the share of benchmark bonds held across benchmark weight quintiles by credit ratings.

A.7 Panel regressions of fund holdings on benchmark index membership and weights

Table A6: Regressions of active and passive fund holdings on benchmark index membership and weights for Canada and US

	Fund weight fund weight \neq 0			Dummy(fund weight > 0)		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A1: Canada – Passive funds						
Dummy(within benchmark)		0.002** (0.001)	0.002** (0.001)		0.589*** (0.013)	0.450*** (0.016)
Benchmark weight			0.944*** (0.058)			48.70*** (3.351)
Observations	167,944	167,944	167,944	577,325	577,325	577,325
R^2	–	72.9%	73.6%	–	61.8%	62.4%
Within R^2	–	0.22%	2.64%	–	29.33%	30.42%
Panel A2: Canada – Active funds						
Dummy(within benchmark)		0.001 (0.001)	0.001 (0.001)		0.152*** (0.007)	0.080*** (0.006)
Benchmark weight			0.634*** (0.124)			30.70*** (2.629)
Observations	268,227	268,227	268,227	3,046,415	3,046,415	3,046,415
R^2	–	75.4%	75.4%	–	15.1%	15.4%
Within R^2	–	0.06%	0.14%	–	1.65%	1.92%
Panel B1: US – Passive funds						
Dummy(within benchmark)		0.051*** (0.009)	-0.008 (0.009)		0.535*** (0.007)	0.481*** (0.008)
Benchmark weight			0.755*** (0.033)			1.051*** (0.079)
Observations	2,639,677	2,639,677	2,639,677	18,084,928	18,084,928	18,084,928
R^2	–	83.9%	84.1%	–	35.1%	35.3%
Within R^2	–	0.04%	0.76%	–	19.79%	19.99%
Panel B2: US – Active funds						
Dummy(within benchmark)		0.106*** (0.007)	0.053*** (0.011)		0.053*** (0.001)	0.022*** (0.001)
Benchmark weight			0.502*** (0.088)			0.568*** (0.016)
Observations	2,430,714	2,430,714	2,430,714	96,720,178	96,720,178	96,720,178
R^2	–	97.8%	97.8%	–	6.76%	6.98%
Within R^2	–	0.06%	0.08%	–	0.57%	0.80%

This table reports estimates of the conditional correlation between fund weight or propensity to hold and benchmark membership dummy variable and benchmark index weight separately for passive and active funds. Panels A1 and A2 report the results for Canada, and Panels B1 and B2 for the US. All regressions are saturated by including fund-by-date and bond-by-date fixed effects. Standard errors clustered at the bond and year-month levels are presented in parentheses. Significance levels are marked as: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A.8 Construction of historical benchmark data for US funds

We assemble a dataset of historical benchmarks of US mutual funds and ETFs from the database of historical fund prospectus available on the website of the U.S. Securities and Exchange Commission (SEC)⁸. We further validate them using two additional sources: (1) snapshot of benchmarks in Morningstar as of March 2025, and (2) SEC Mutual Fund Prospectus Risk/Return Summary data sets (MFRR)⁹. Benchmarks are mentioned in the annual return summary published in prospectuses.

A.8.1 Extract Benchmarks

For the `crsp_cik_map` table in WRDS (containing a total of 2586 company CIKs), all the 485APOS and 485BPOS files for 2504 company CIKs were downloaded using the `sec_edgar_downloader` Python package. By comparing several samples, it was confirmed that all required file types were successfully downloaded. Through a manual check, it was verified that the remaining company CIKs without downloads do not have data available on the EDGAR website.

From the 485APOS and 485BPOS files, extract the table(s) containing or immediately following the string “(average) annual (total) return(s)” (with parentheses indicating optional terms), marked by the `<table></table>` tags. If the table contains any of the following terms, retain the table:

- s&p, russell, crsp, msci, dj, dow jones, nasdaq, ftse, topix, tse, schwab, barclays, wilshire, bridgeway, guggenheim, calvert, kaizen, lipper, redwood, w.e. donoghue, essential treuters, barra, ice bofaml, bbgbarc, cboe
- benchmark, index (note: does not match *index fund* or *index series*)

Use the large language model (Gemini API) to determine if the extracted table is an average annual return table. If it is, extract the corresponding benchmark for each fund and label the benchmark in sequence.

⁸<https://www.sec.gov/edgar/searchedgar/mutualsearch.html>

⁹<https://www.sec.gov/dera/data/mutual-fund-prospectus-risk-return-summary-data-sets>

If a file lists all series names (fund names) at the beginning using the <series-name> tag, identify the closest preceding series name for each table.

For tables where the closest preceding series name was not found in the previous step, identify the most similar series name and its corresponding series ID from the SEC-provided mapping using the fund name(s) in the table.

If neither of the above methods succeeds, use the Gemini API to analyze the text preceding each table and identify the string closest to the table that is most likely to be a fund name.

For each table, prioritize the selection of the extracted series name as follows:

1. First, use the closest preceding series name as the extracted series name.
2. If no series name is found, use the closest preceding string most likely to be a fund name, as identified from the text preceding the table.
3. If neither of the above is available, use the fund name within the table as the extracted series name.

The SEC provides the company CIK, company name, series ID, and series name for each fund for every year from 2010 to 2024. For each “average annual return” table in year t , extract the SEC mapping for the corresponding company CIK for year $t-1$, t , and $t+1$. If the table is from a year earlier than the earliest available year for the company CIK, retrieve the minimum available year and the next year’s SEC mapping. For each extracted series name, identify the totally-matched series name and its corresponding series ID from the SEC-provided mapping.

A.8.2 Fund Categories

We classify US mutual funds and ETFs into active and passive following [Pavlova and Sikorskaya \(2023\)](#).

Active v.s Inactive

1. If `et_flag` is empty and `index_fund_flag` is empty, it is an “Active Fund”.

2. If `et_flag` is empty and `index_fund_flag` is “B” (Index Based Fund), it is an “Active Fund”
3. If `index_fund_flag` is D or E, it is an “Index Fund”
4. If the fund name contains key words: ‘index’, ‘indx’, ‘ idx ’, ‘s&p’, ‘ sp’, ‘nasdaq’, ‘msci’, ‘crsp’, ‘ftse’, ‘barclays’, ‘ dj ’, ‘ dow ’, ‘jones’, ‘russell’, ‘ nyse ’, ‘wilshire’, ‘400’, ‘500’, ‘600’, ‘1000’, ‘1500’, ‘2000’, ‘2500’, ‘3000’, ‘5000’, ‘ dfa ’, ‘program’, it is an “Index Fund”.
5. If `et_flag` is not empty, it is an “ETF”
6. If the fund name contains key words: ‘spdr’, ‘trackers’, ‘holders’, ‘powershares’, ‘street-tracks’, ‘etf’, ‘exchange traded’, ‘exchange-traded’, it is an ETF

Assets Types

1. If `crsp_obj_cd` starts with ‘ED’, it is “Equity Domestic”
2. If `crsp_obj_cd` starts with ‘EF’, it is “Equity Foreign”
3. If `crsp_obj_cd` starts with ‘IF’, it is “Fixed Income Foreign”
4. If `crsp_obj_cd` starts with ‘IM’, it is Fixed Income Domestic Money Market”
5. If `crsp_obj_cd` starts with ‘I’ but doesn’t start with ‘IF’ or ‘IM’, it is “Fixed Income Domestic”
6. If `crsp_obj_cd` starts with ‘M’, it is “Mixed FI & Equity”
7. All the other categories are “Other”

We only used fixed income domestic funds with their primary benchmarks.

A.8.3 Clean Benchmarks

We let Gemini API to create a mapping, manually check it, and then use it to unify different names of the same benchmark into a single name, such as unifying ‘lehman aggregate bond index’, ‘bloomberg us agg bond tr usd’, ‘bloomberg capital u.s. aggregate tr usd’ as ‘bloomberg u.s aggregate bond tr usd’.

A.8.4 Correct Mismatched Funds and Benchmarks

If a SEC file contains multiple funds and their corresponding benchmarks in succession, for example, “fund 1, benchmark 1, fund 2, benchmark 2, fund 3, benchmark 3”, but the SEC does not provide the seriesId for fund 2, or the name of fund 2 in the file does not match the name of fund 2 in the SEC-provided mapping, then both benchmark 1 and benchmark 2 might be assigned to fund 1, leading to incorrect matching. Given the low probability of this situation occurring, I have adopted the following method to correct it: if a fund consistently uses benchmark A but suddenly shows benchmark B midway (like A A A A B A A A A), benchmark B is clearly a mismatched entry, and I replace B with A.

A.8.5 Validation

Validation Against MFRR We validate primary benchmarks extracted from SEC with MFRR time series benchmark data. Due to the incompleteness of MFRR data, 4.4 trillion are only in our sample, 0.3 trillion are only in MFRR, so the validation is for 2.1 trillion co-exist funds for the same quarter. The matching rate is 81% (based on the latest total net assets), that is, these funds have the same benchmark for the same quarter between MFRR and our extracted data.

Validation Against Morningstar We only include the primary benchmarks for both SEC and MS benchmarks.

We deleted 205 seriesId (SEC fund id) which have more than one primary prospectus benchmarks in Morningstar. It’s usually because the mapping file wrongly map different funds’ secid (morningstar id) to the same seriesId.

Since MS benchmarks are the latest snapshot, for each seriesId (fund), I only compare the latest benchmark extracted from SEC with MS.

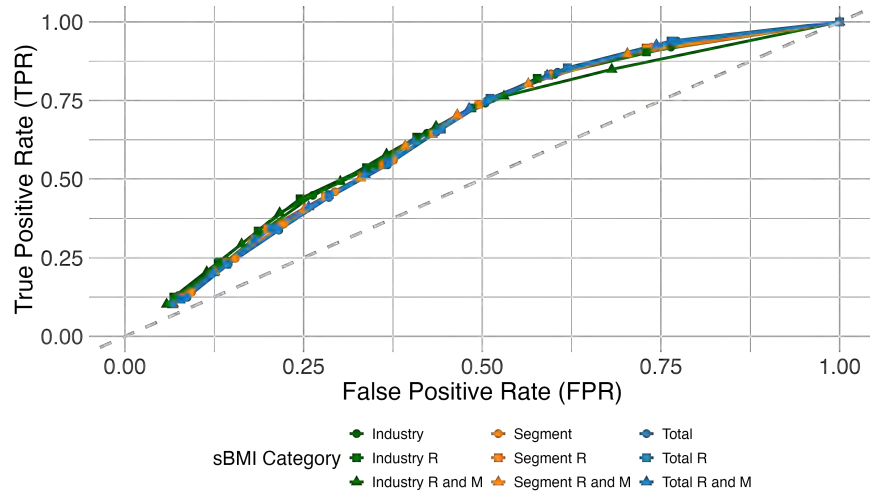
There are 1982 series (2.82 trillion) which only have benchmarks from SEC, and 248 series (0.28 trillion) which only have benchmarks from MS. We can only compare 1360 series (3.63 trillion) which have benchmarks both in SEC and MS.

86% (by the latest total net assets) have the same benchmark.

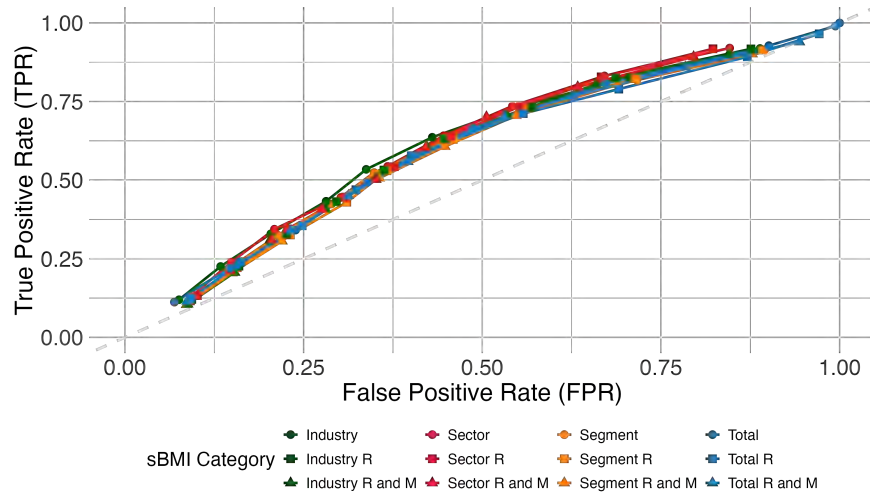
B Appendix: Empirical Analysis

B.1 What is the right level of sparsity?

Figure B1: Area-under-the-curve plots for sparse BMIs



(a) Canada

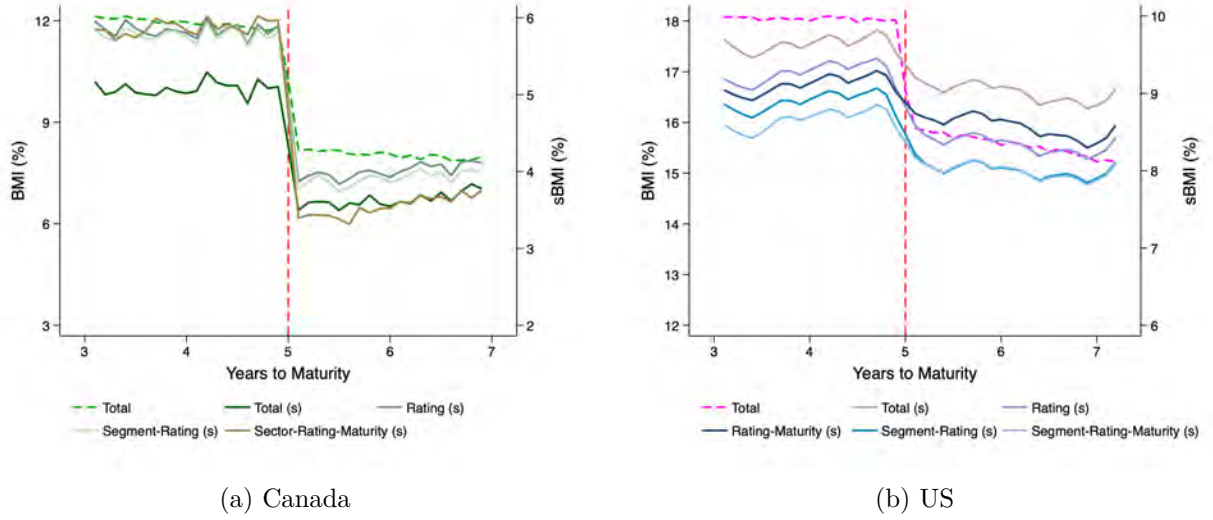


(b) US

This figure illustrates how the true and false positive rates of sparse BMIs change as we vary the level of sparsity. The rates are computed with respect to the sparse BMI predicting total fund ownership. In sparse BMI categories, 'R' stands for sparsity within rating and 'M' stands for sparsity within maturity.

B.2 (Sparse) BMIs around the 5-year cutoff

Figure B2: Total and sparse BMIs near the 5-year maturity cutoff



This figure shows how total and sparse BMIs change around the 5-year cutoff in Canada (a) and the United States (b).

B.3 Sparse BMIs around maturity cutoffs

Table B1: Sparse BMI changes around maturity cutoffs (United States and Canada)

	Total BMI		Sparse BMI			
		Rating	Segment– Rating	Segment– Rating– Maturity	Sector– Rating	Sector– Rating– Maturity
Panel A: United States						
<i>A1. Full sample (with recently issued bonds)</i>						
D(Crossed 3y)	3.921*** (0.060)	1.965*** (0.045)	1.900*** (0.047)	1.868*** (0.050)	1.787*** (0.048)	1.658*** (0.064)
D(Crossed 5y)	1.205*** (0.050)	0.472*** (0.043)	0.476*** (0.045)	0.335*** (0.075)	0.599*** (0.053)	1.010*** (0.095)
D(Crossed 10y)	0.095 (0.076)	−0.083 (0.069)	−0.176** (0.080)	−0.357*** (0.136)	−0.137 (0.087)	−0.126 (0.184)
Observations	722,238	722,238	722,238	722,238	722,238	722,238
R-squared	0.399	0.399	0.399	0.399	0.399	0.399
<i>A2. Baseline sample</i>						
D(Crossed 3y)	3.917*** (0.060)	1.958*** (0.045)	1.892*** (0.047)	1.861*** (0.051)	1.795*** (0.047)	1.661*** (0.065)
D(Crossed 5y)	1.183*** (0.049)	0.457*** (0.043)	0.468*** (0.044)	0.314*** (0.075)	0.585*** (0.055)	0.983*** (0.095)
D(Crossed 10y)	0.220*** (0.069)	−0.142 (0.103)	−0.227** (0.107)	−0.434* (0.223)	−0.059 (0.127)	−0.008 (0.264)
Observations	605,661	605,661	605,661	605,661	605,661	605,661
R-squared	0.393	0.393	0.393	0.393	0.393	0.393
Panel B: Canada						
<i>B1. Full sample (with recently issued bonds)</i>						
D(Crossed 5y)	1.791*** (0.022)	1.574*** (0.069)	1.577*** (0.074)	1.394*** (0.099)		
D(Crossed 10y)	−0.681*** (0.018)	−0.904*** (0.071)	−0.980*** (0.076)	−1.757*** (0.173)		
Observations	175,309	175,309	175,309	175,309		
R-squared	0.932	0.932	0.932	0.932		
<i>B2. Baseline sample</i>						
D(Crossed 5y)	1.801*** (0.019)	1.435*** (0.071)	1.435*** (0.074)	1.211*** (0.105)		
D(Crossed 10y)	−0.671*** (0.018)	−0.582*** (0.078)	−0.663*** (0.084)	−1.378*** (0.219)		
Observations	123,916	123,916	123,916	123,916		
R-squared	0.934	0.934	0.934	0.934		

This table reports coefficients on maturity-switch indicator variables (3/5/10-year for the US and 5/10-year for Canada) using changes in total BMI and sparse BMIs as dependent variables. Panels A1 and B1 use full samples, while panels A2 and B2 remove recently issued bonds (our baseline). All models include controls for log amount outstanding, numeric rating, bid–ask spread quintile, and time to maturity, and have bond and year–month fixed effects. Standard errors (clustered by year–month) in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results in panel B omit sector-based sparse BMIs due to the absence of this classification level in Canada (blank cells).

B.4 Sparse BMIs and bond prices

Table B2: Changes in BMIs and bond yields (United States and Canada)

	Total BMI		Sparse BMI			
		Rating	Segment– Rating	Segment– Rating– Maturity	Sector– Rating	Sector– Rating– Maturity
Panel A: United States						
$t-1$ to $t+1$	–0.005** (0.002)	–0.019*** (0.004)	–0.016*** (0.003)	–0.011*** (0.003)	–0.013*** (0.003)	–0.008*** (0.002)
Observations	605,726	605,726	605,726	605,726	605,726	605,726
R-squared	0.664	0.667	0.667	0.666	0.667	0.666
$t+3$	–0.002 (0.003)	–0.013*** (0.002)	–0.011*** (0.002)	–0.008*** (0.002)	–0.010*** (0.002)	–0.006*** (0.001)
Observations	589,024	589,024	589,024	589,024	589,024	589,024
R-squared	0.707	0.708	0.707	0.707	0.707	0.707
$t+6$	–0.005 (0.004)	–0.011*** (0.002)	–0.009*** (0.002)	–0.006*** (0.001)	–0.009*** (0.002)	–0.006*** (0.001)
Observations	564,148	564,148	564,148	564,148	564,148	564,148
R-squared	0.773	0.773	0.773	0.773	0.773	0.773
Panel B: Canada						
$t-1$ to $t+1$	–0.013 (0.015)	–0.010*** (0.002)	–0.008*** (0.001)	–0.003*** (0.001)	–0.005*** (0.001)	–0.002*** (0.000)
Observations	123,748	123,748	123,748	123,748	123,748	123,748
R-squared	0.733	0.733	0.733	0.733	0.733	0.733
$t+3$	–0.015 (0.022)	–0.008*** (0.002)	–0.006*** (0.002)	–0.003*** (0.001)	–0.005*** (0.001)	–0.003*** (0.001)
Observations	123,402	123,402	123,402	123,402	123,402	123,402
R-squared	0.752	0.752	0.752	0.752	0.752	0.752
$t+6$	–0.017 (0.028)	–0.007*** (0.003)	–0.004** (0.002)	–0.003** (0.001)	–0.005*** (0.001)	–0.003*** (0.000)
Observations	122,781	122,781	122,781	122,781	122,781	122,781
R-squared	0.777	0.777	0.777	0.777	0.777	0.777

This table reports the $\Delta sBMI$ coefficient estimates from regression (23) for the U.S. (Panel A) and Canadian (Panel B) samples. The dependent variable is the change in bond's yield to maturity. All regressions control for the logarithm of the issue amount outstanding, numeric credit rating, bid–ask spread quintile, and time to maturity, and include bond and year–month fixed effects. Standard errors, clustered by issuer and year–month, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B3: Changes in BMIs and bond returns (United States and Canada)

	Total BMI		Sparse BMI			
		Rating	Segment– Rating	Segment– Rating– Maturity	Sector– Rating	Sector– Rating– Maturity
Panel A: United States (full sample, accum)						
$t-1$ to $t+1$	-0.028 (0.000)	0.199 (0.000)	0.165 (0.000)	0.072 (0.000)	0.104 (0.000)	0.032 (0.000)
Observations	605,726	605,726	605,726	605,726	605,726	605,726
R-squared	0.613	0.618	0.617	0.614	0.616	0.613
$t+3$	-0.084 (0.060)	0.146*** (0.023)	0.119*** (0.020)	0.049*** (0.012)	0.077*** (0.011)	0.019*** (0.004)
Observations	589,018	589,018	589,018	589,018	589,018	589,018
R-squared	0.630	0.632	0.631	0.630	0.631	0.630
$t+6$	-0.083 (0.063)	0.141*** (0.024)	0.115*** (0.019)	0.034*** (0.012)	0.074*** (0.012)	0.013*** (0.005)
Observations	564,023	564,023	564,023	564,023	564,023	564,023
R-squared	0.671	0.672	0.672	0.671	0.672	0.671
Panel B: Canada (full sample, accum)						
$t-1$ to $t+1$	-0.001 (0.002)	0.001*** (0.000)	0.001*** (0.000)	0.000*** (0.000)		
Observations	118,339	118,339	118,339	118,339		
R-squared	0.597	0.597	0.597	0.597		
$t+3$	0.000 (0.003)	0.001*** (0.000)	0.001*** (0.000)	0.000*** (0.000)		
Observations	118,071	118,071	118,071	118,071		
R-squared	0.625	0.625	0.625	0.625		
$t+6$	-0.005* (0.003)	0.001*** (0.000)	0.001*** (0.000)	0.000*** (0.000)		
Observations	117,576	117,576	117,576	117,576		
R-squared	0.656	0.656	0.656	0.656		

This table reports the $\Delta sBMI$ coefficient estimates from regression (23) for the U.S. (Panel A) and Canadian (Panel B) samples. The dependent variable is the bond's total return (local currency return from Bloomberg for the U.S. or Refinitiv's total return for Canada). All regressions control for the logarithm of the issue amount outstanding, numeric credit rating, bid-ask spread quintile, and time to maturity, and include bond and year-month fixed effects. Standard errors, clustered by issuer and year-month, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.5 Ownership changes around maturity cutoffs

Table B4: Ownership changes around maturity cutoffs (United States and Canada)

	Total ownership	Active ownership	Passive ownership
Panel A: United States			
D(Crossed 3y)	0.385*** (0.067)	0.094 (0.061)	0.288*** (0.022)
D(Crossed 5y)	-0.412*** (0.093)	-0.101 (0.085)	-0.312*** (0.032)
D(Crossed 10y)	0.513** (0.233)	-0.106 (0.227)	0.614*** (0.075)
Observations	520,572	520,572	520,572
R-squared	0.204	0.204	0.287
Panel B: Canada			
D(Crossed 5y)	1.021*** (0.156)	0.815*** (0.157)	0.206*** (0.033)
D(Crossed 10y)	0.425 (0.282)	0.302 (0.263)	0.131** (0.066)
Observations	122,441	122,441	122,441
R-squared	0.534	0.509	0.348

This table reports coefficients on maturity-switch indicator variables (3/5/10-year for the US and 5/10-year for Canada) using (t-6; t+6) changes in fund ownership (total, active, and passive) as dependent variables. All regressions include controls for log amount outstanding, numeric rating, bid-ask spread quintile, and time to maturity, and have bond and year-month fixed effects. Standard errors (clustered by year-month) in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.6 Bonds around the 5-year cutoff

Table B5: Changes in BMIs and bond yield spreads (United States and Canada)

	Total BMI		Sparse BMI			
		Rating	Segment– Rating	Segment– Rating– Maturity	Sector– Rating	Sector– Rating– Maturity
Panel A: United States						
$t-1$ to $t+1$	–0.006*** (0.002)	–0.021*** (0.006)	–0.021*** (0.006)	–0.011*** (0.003)	–0.010*** (0.002)	–0.004*** (0.001)
Observations	46,175	46,175	46,175	46,175	46,175	46,175
R-squared	0.575	0.575	0.575	0.575	0.575	0.575
$t-1$ to $t+3$	–0.004* (0.003)	–0.014*** (0.004)	–0.014*** (0.004)	–0.007*** (0.003)	–0.004*** (0.001)	–0.002*** (0.001)
Observations	45,005	45,005	45,005	45,005	45,005	45,005
R-squared	0.620	0.620	0.620	0.620	0.620	0.620
$t-1$ to $t+6$	–0.001 (0.002)	–0.013*** (0.004)	–0.012*** (0.004)	–0.008*** (0.002)	–0.004*** (0.001)	–0.002** (0.001)
Observations	43,318	43,318	43,318	43,318	43,318	43,318
R-squared	0.654	0.654	0.654	0.654	0.654	0.654
Panel B: Canada						
$t-1$ to $t+1$	0.005 (0.003)	–0.006*** (0.002)	–0.004** (0.002)	–0.001 (0.001)		
Observations	33,972	33,972	33,972	33,972		
R-squared	0.654	0.654	0.654	0.654		
$t-1$ to $t+3$	0.009 (0.006)	–0.004* (0.002)	–0.003 (0.002)	–0.000 (0.001)		
Observations	33,865	33,865	33,865	33,865		
R-squared	0.712	0.712	0.712	0.712		
$t-1$ to $t+6$	0.005 (0.006)	–0.002 (0.003)	–0.000 (0.002)	–0.000 (0.001)		
Observations	33,700	33,700	33,700	33,700		
R-squared	0.751	0.751	0.751	0.751		

This table reports the $\Delta sBMI$ coefficient estimates from regression (23) for the U.S. (Panel A) and Canadian (Panel B) samples only for bonds around the 5-year cutoff (within 6 months for the U.S. and 2 years for Canada to maintain comparable sample size). The dependent variable is the change in yield spread (relative to a maturity-matched Treasury bond) in percentage points. All regressions control for the logarithm of the issue amount outstanding, numeric credit rating, bid–ask spread quintile, and time to maturity, and include bond and year–month fixed effects. Standard errors, clustered by bond and year–month, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results for Canada omit sector-based sparse BMIs (blank cells).

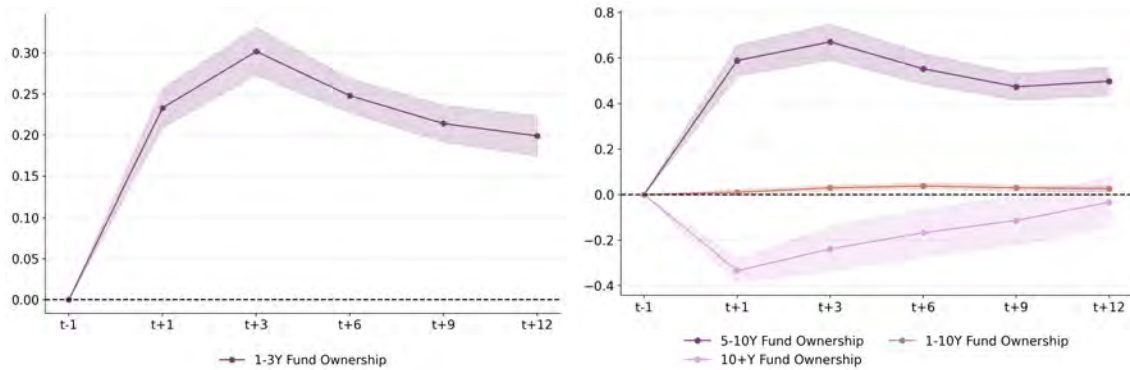
Table B6: Changes in ownership and sparse BMIs (United States and Canada)

	<u>Total BMI</u>		<u>Sparse BMI</u>			
		Rating	Segment– Rating	Segment– Rating– Maturity	Sector– Rating	Sector– Rating– Maturity
Panel A: United States (3-month horizon)						
<i>A1. Total ownership</i>						
$t-3$ to $t\pm 3$	–0.028 (0.031)	0.045*** (0.014)	0.034*** (0.013)	0.023*** (0.009)	0.019 (0.015)	0.001 (0.009)
Observations	44,433	44,433	44,433	44,433	44,433	44,433
R-squared	0.436	0.437	0.437	0.436	0.436	0.436
<i>A2. Active ownership</i>						
$t-3$ to $t\pm 3$	–0.003 (0.030)	0.044*** (0.013)	0.034*** (0.012)	0.021** (0.009)	–0.018 (0.021)	0.007 (0.009)
Observations	44,433	44,433	44,433	44,433	44,433	44,433
R-squared	0.433	0.434	0.434	0.433	0.433	0.433
<i>A3. Passive ownership</i>						
$t-3$ to $t\pm 3$	–0.024*** (0.009)	–0.000 (0.003)	–0.000 (0.003)	0.003 (0.002)	0.006 (0.003)	0.003 (0.002)
Observations	44,433	44,433	44,433	44,433	44,433	44,433
R-squared	0.423	0.423	0.423	0.423	0.423	0.423
Panel B: Canada (6-month horizon)						
<i>B1. Total ownership</i>						
$t-6$ to $t\pm 6$	0.230*** (0.076)	0.077*** (0.029)	0.068*** (0.025)	0.027 (0.017)		
Observations	33,560	33,560	33,560	33,560		
R-squared	0.680	0.680	0.680	0.680		
<i>B2. Active ownership</i>						
$t-6$ to $t\pm 6$	0.169** (0.077)	0.060** (0.028)	0.053** (0.025)	0.019 (0.019)		
Observations	33,560	33,560	33,560	33,560		
R-squared	0.650	0.650	0.650	0.650		
<i>B3. Passive ownership</i>						
$t-6$ to $t\pm 6$	0.062*** (0.017)	0.019*** (0.006)	0.017*** (0.005)	0.009*** (0.004)		
Observations	33,560	33,560	33,560	33,560		
R-squared	0.479	0.479	0.479	0.479		

This table reports the $\Delta sBMI$ coefficient estimates for the U.S. (Panel A, 3-month horizon) and Canadian (Panel B, 6-month horizon) samples only for bonds around the 5-year cutoff (within 6 months for the U.S. and 2 years for Canada to maintain comparable sample size). The dependent variables are changes in fund ownership (total, active, and passive) in percentage points. All regressions control for the logarithm of the issue amount outstanding, numeric credit rating, bid–ask spread quintile, and time to maturity, and include bond and year–month fixed effects. Standard errors, clustered by bond and year–month, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results for Canada omit sector-based sparse BMIs (blank cells).

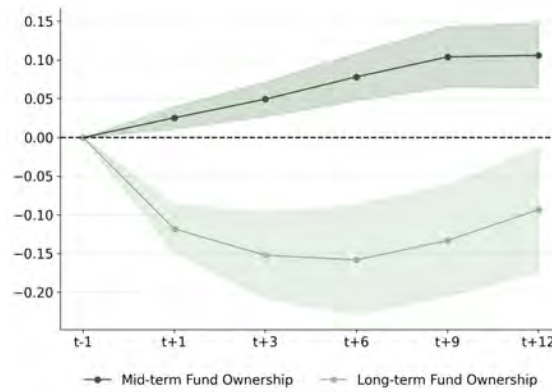
B.7 Maturity-based benchmarked ownership changes around maturity cutoffs

Figure B3: Changes in benchmarked ownership around the 3-year and 10-year maturity cutoffs



(a) US, 3-year cutoff

(b) US, 10-year cutoff

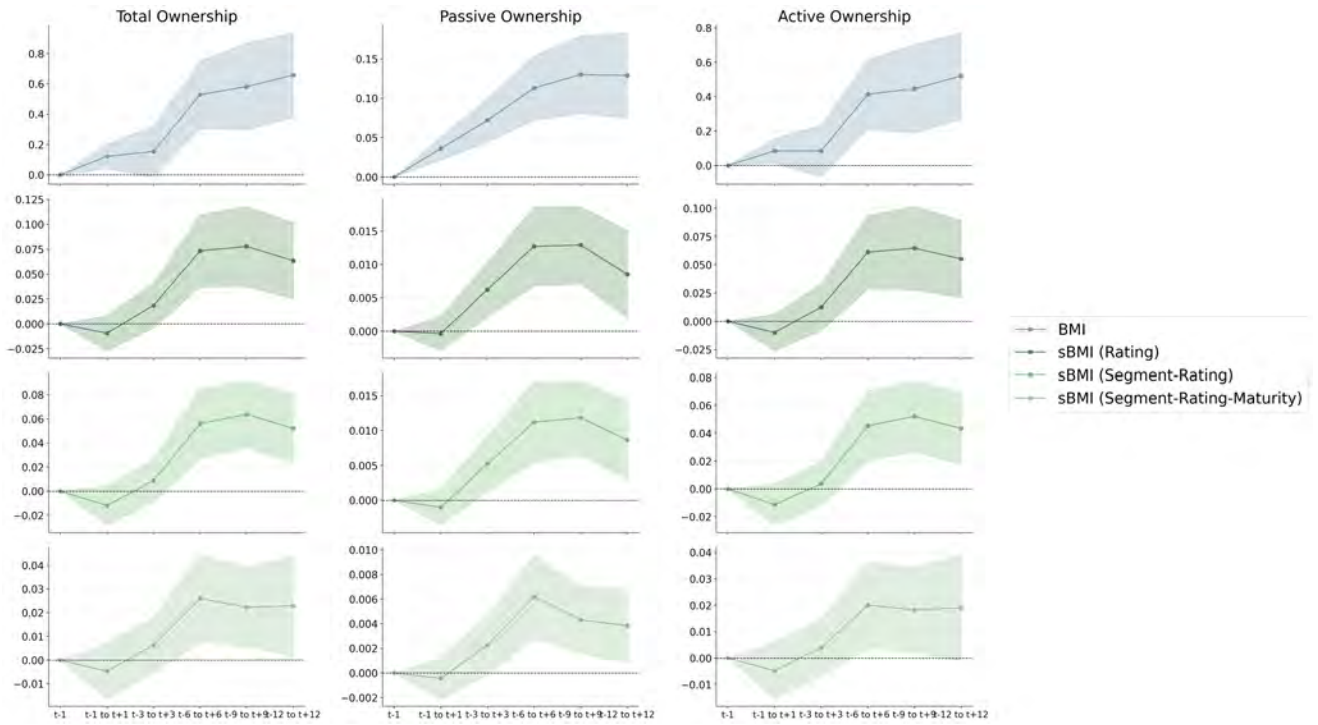


(c) Canada, 10-year cutoff

This figure illustrates how benchmarked fund ownership of corporate bonds changes around the 3/10-year maturity cutoffs in the United States and the 10-year cutoff in Canada. We report total fund ownership changes of funds with maturity-based benchmark indexes, where sufficient ownership data is available.^a All estimates and confidence intervals are from estimating specification (23) around the respective 3/10-year cutoffs. All y-axes correspond to p.p. changes in ownership in response to the move across the 3/10-year cutoff.

^aIn particular, 3-5 year funds have insufficient data for this exercise.

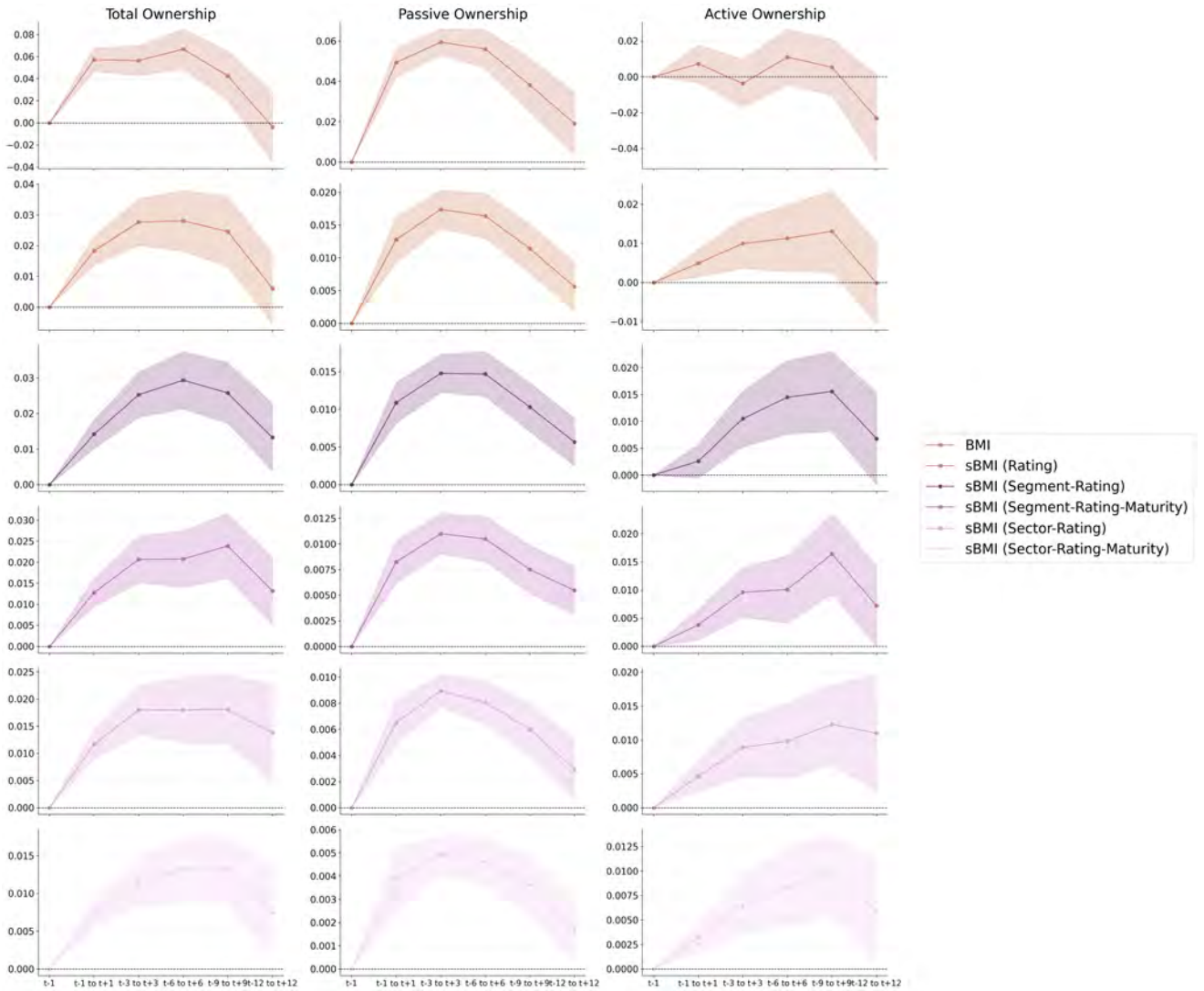
B.8 Changes in ownership at longer horizons



(a) Canada

(Continued on the next page.)

Figure B4: Changes in fund ownership at longer horizons

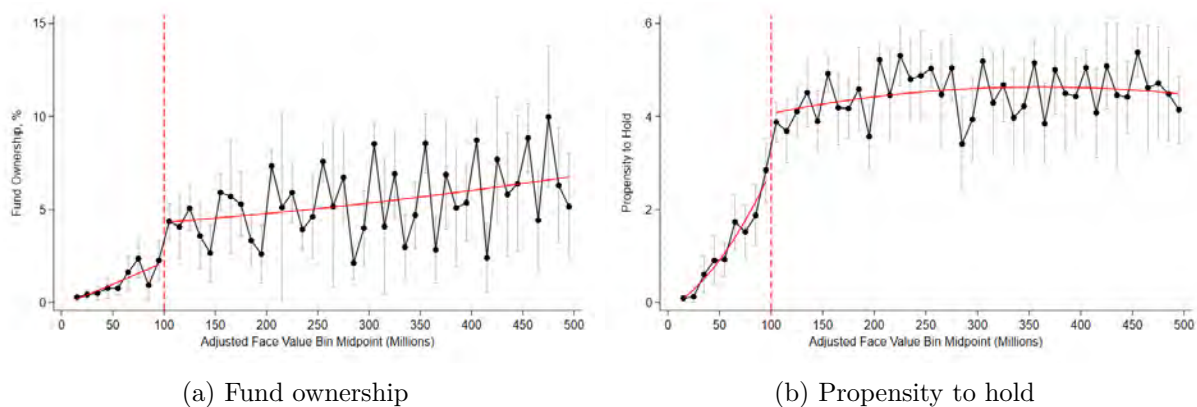


(b) US

This figure plots the $\Delta sBMI$ coefficient estimates for the Canadian (a) and U.S. (b) samples at more horizons than in the main text. The dependent variables are changes in fund ownership (total, active, and passive) in percentage points. Top row is for the total BMI, and the other rows are for the sparse BMIs (see the legend). All regressions control for the logarithm of the issue amount outstanding, numeric credit rating, bid-ask spread quintile, and time to maturity, and include bond and year-month fixed effects. We report 90% confidence intervals based on standard errors clustered by issuer and year-month.

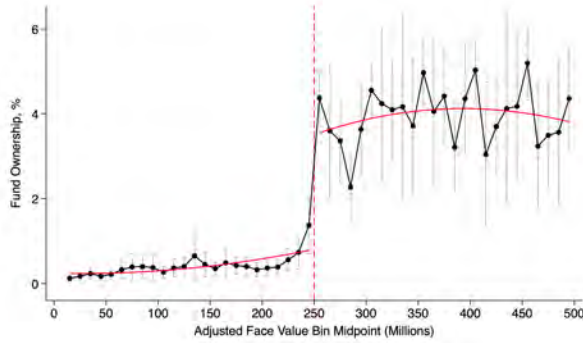
B.9 Benchmark Issue Size Cutoff and Fund Ownership

Figure B5: Ownership around the FTSE size cutoff in Canada

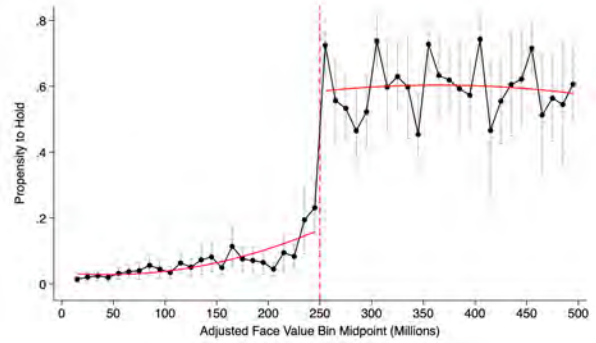


This figure illustrates fund ownership and propensity to hold corporate bonds around FTSE index issue size eligibility cutoff of 100 million CAD. The sample includes a union of bonds in FTSE Canada Overall Universe bond index with bonds in Bloomberg Canada universe.

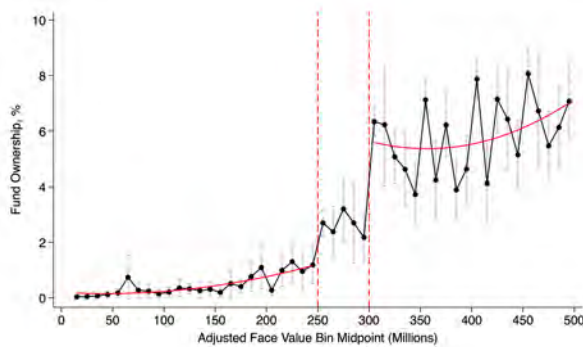
Figure B6: Ownership around the Bloomberg size cutoff in United States



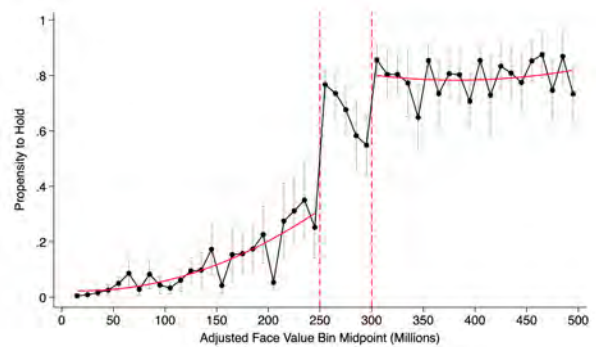
(a) Fund ownership before Q2-2017



(b) Propensity to hold before Q2-2017



(c) Fund ownership after Q2-2017



(d) Propensity to hold after Q2-2017

This figure illustrates fund ownership and propensity to hold corporate bonds around Bloomberg index issue size eligibility cutoff of 250 million USD before the second quarter of 2017 and around the new size cutoff of 300 million USD after. The sample includes a union of bonds in Bloomberg Aggregate bond index with bonds in WRDS sample.