Optimal Index-Linked Rebalancing with Anticipatory Trading*

Stefano Pegoraro† Marco Sammon‡ John J. Shim§

August 1, 2026

Abstract

Using a model of index-linked rebalancing around reconstitution events, we show front-runners provide liquidity to index investors and benefit them. Index investors trade off execution costs against tracking-error concerns, while speculators maximize trading profit. In competitive markets, even loose index trackers optimally rebalance at index reconstitution, rationalizing concentrated reconstitution-day trading with little contemporaneous price impact. Empirically, index-linked investors bear small rebalancing costs relative to standard transaction-cost estimates and traders short sell shares to provide liquidity at reconstitution. We show fast-track additions increase IPO issue prices but impose costs on index investors when the newly-listed stocks are illiquid and hard to short.

Keywords: index rebalancing, index funds, passive investing, benchmark tracking, tracking error, front running, anticipatory trading, liquidity provision, price impact, fast-track additions, IPO.

---

*We thank Rob Battalio, Thomas Eisenbach, Miklos Farkas, Alessandro Melone, Taisiya Sikorskaya, Adi Sunderam, and the audience at the Washington University in St. Louis, Bristol Financial Markets Conference, and University of Washington Summer Finance Conference for helpful comments.

†Mendoza College of Business, University of Notre Dame, United States of America. Email: s.pegoraro@nd.edu

‡Harvard Business School, Harvard University, United States of America. Email: mcsammon@gmail.com

§Mendoza College of Business, University of Notre Dame, United States of America. Email: jshim2@nd.edu

1 Introduction

As of the end of 2025, investors delegated \$19 trillion to funds whose sole mandate is to track an index. In the U.S. equity market, index funds hold 19% of total market capitalization.[1] These figures likely understate the true scale of index-linked investing. Pension funds, active funds, and other institutional investors also benchmark their portfolios to indexes (Pavlova and Sikorskaya, 2023). Including these investors nearly doubles the estimated amount of index-tracking assets relative to strict index funds alone (Chinco and Sammon, 2024).

Because index providers regularly change the composition of their indexes, index-linked investors must trade to keep their portfolios aligned with the benchmark and minimize tracking error. These trades are both highly predictable and concentrated. When S&P and Russell index changes are implemented, index trackers often trade more than 20% of a stock's shares outstanding on a single day in the closing auction (Chinco and Sammon, 2024). Despite this enormous trading volume, there is little apparent price impact on the implementation day itself. This pattern has naturally raised the concern that the cost of predictable rebalancing may instead be effectively paid *before* implementation, as anticipatory traders move prices ahead of index-linked demand. As a specific example, Harvey et al. (2025) shows that institutional investors that mechanically rebalance between stocks and bonds face large costs from predictable cross-asset trading, which they estimate at \$16 billion per year.

These facts raise a puzzle. On the one hand, index-trackers trade large amounts in a single day with minimal price impact. On the other hand, they appear to trade suboptimally by executing trades at an already inflated price. To further add to the confusion, researchers have documented that even active investors who can flexibly track an index also tend to trade at the exact same time as "strict" index trackers, such as index funds (Chinco and Sammon, 2024). In this paper, we build a model to explain this puzzle, and reconcile these facts as outcomes of a unified economic framework.

Specifically, we argue these patterns emerge from the strategic interaction between index trackers and speculators who specialize in providing liquidity around index changes and who have become increasingly competitive and efficient over the last 15 years. Over this period, a sizable industry has emerged that tries to forecast and trade ahead of index additions and deletions.[2] These speculators build positions outside of the implementation

[1] Source: 2026 Investment Company Factbook, https://www.icifactbook.org/.

[2] For example, see articles by Bloomberg (https://www.bloomberg.com/news/features/2022-08-01/goldman-lost-2-traders-then-accused-them-of-accessing-secret-code) and Business Insider (https://www.businessinsider.com/millennium-hedge-funds-glen-schei

date and then supply shares to index trackers at the closing auction at implementation. Given that many additions to and deletions from indexes are highly predictable, these speculators can begin building positions weeks or months in advance of the event itself. Because "front running" is often viewed as predatory (Brunnermeier and Pedersen, 2005), it is tempting to interpret their behavior as costly for index fund investors. But that interpretation is incomplete. We claim that index investors, who are concerned with tracking error, gain from trading with speculators who do not share those concerns. We further claim that, to truly assess the cost speculators impose, one needs to assess the counterfactual costs that index investors would incur if speculators did not act as intermediaries by pre-positioning liquidity. After accounting for gains from trade and counterfactual costs, our framework highlights that index investors have been increasingly benefiting from the activity of intermediaries over the recent decades.

We start by expanding on and linking the facts originally documented in Chinco and Sammon (2024) and Greenwood and Sammon (2025). Namely, volume is concentrated on the so-called *implementation date*; that is, the trading day as of which the index change becomes effective. However, returns are relatively small on that day. As a dramatic example, between 2010 and 2024, average implementation-day returns for direct additions to the S&P 500 were -9 bps, despite trading volume that is 20-25 times these stocks' past average daily trading volume (ADV). This does not mean that prices are flat throughout the pre-event window. For S&P events, prices may move between announcement of the index change and implementation. For Russell events, index changes are predictable well in advance of implementation, because Russell adds or drops stocks to its indexes based on how they rank in terms of market capitalization. Therefore, pre-rank returns mechanically affect index inclusions and exclusions. The striking fact, however, is that the day on which index trackers appear to trade is not the day on which we observe the largest price movements.

We argue that, to rationalize these patterns, one needs to view speculators as intermediaries. Index trackers face a dual mandate of maintaining low tracking errors and minimizing trading costs. Speculators face no tracking-error concerns. By acquiring inventory ahead of the implementation date and selling at implementation, speculators intermediate shares over time and provide liquidity to index trackers when they value it the most. In deciding their optimal rebalancing schedule, the index tracker trades off tracking error concerns with the rents that speculators earn by prepositioning. In this paper, we present a model that formalizes this tradeoff and we establish conditions whereby index trackers optimally trade on the implementation day.

`nberg-index-rebalance-tesla-2022-1`).

In the model, an index-tracking fund must sell an outgoing constituent and buy an incoming constituent around the implementation date. The fund can split its order across periods, but incurs a tracking-error cost if it trades outside the implementation date. Strategic speculators trade to maximize profits. By trading, all agents generate transitory and persistent price impacts, which they internalize. We solve the model under two equilibrium concepts. Under no commitment, the fund chooses its schedule while taking speculators' positions as given. Under full commitment, the fund can credibly pre-announce its schedule and speculators respond to that announcement.

The fund faces a trade-off between avoiding tracking error and reducing transaction costs, including the rents that accrue to speculators. When tracking error concerns are sufficiently strong, the fund's unique optimum is to rebalance entirely on the implementation day. Furthermore, if a fund can credibly commit to a trading schedule, concentrated trading on the implementation day is optimal under even weaker tracking-error cost conditions. By committing to a concentrated trading schedule, the fund can coordinate the provision of liquidity on that date, thereby minimizing tracking-error while also reducing price impact. In contrast, if a fund cannot credibly commit, the temptation to rebalance outside the reconstitution date crowds out intermediary liquidity provision at the reconstitution, raising the costs for the fund. Hence, speculator activity is a form of intermediation or liquidity provision, rather than a purely predatory activity that harms index fund investors.

For any level of tracking-error concerns, the fund increasingly trades on the implementation date as competition between speculators intensifies. In particular, if a fund can credibly commit to a trading schedule, we show that, for any degree of tracking-error concerns, concentrated trading becomes optimal when competition between intermediaries is sufficiently intense. Intuitively, as speculators compete more fiercely, the rents they can extract fall to zero and they implement a socially optimal allocation. By committing to a trading schedule, a fund enjoys both zero tracking errors and cheap liquidity provided by competitive intermediaries.

The model also delivers a set of results that jointly rationalize the observed separation between the timing of prices and the timing of quantities. As speculators gradually preposition and accumulate inventory, prices increase gradually ahead of the implementation date. As they sell to the fund on the implementation date, prices change only modestly despite the large trading volume, consistent with empirical evidence in Chinco and Sammon (2024) and Greenwood and Sammon (2025).

We use our model to evaluate the welfare implications of fast-tracking the addition of a stock to an index shortly after the initial public offering (IPO). Fast-track additions are particularly costly for index investors when the stock is illiquid and short selling is constrained. Both conditions are likely to apply shortly after an IPO. With less time to build up inventory, speculators incur large transaction costs, intermediating a lower share of the volume and forcing index investors to suffer steeper price reversals. In contrast, fast-track additions are beneficial for issuers and initial stock owners, especially when speculators are highly competitive. To build inventory in a short time period, competitive speculators bid up the stock's price as soon as trading is possible, boosting IPO prices.

Finally, we also find complementarity between loose and strict index trackers in concentrating rebalancing at implementation if intermediation is sufficiently competitive. Specifically, we consider the optimal rebalancing schedule of an investor with weak tracking error concerns that trades alongside strict index trackers that rebalance on the implementation date. We show that if the weak tracker's trade is small enough relative to the passive trade and speculators are sufficiently competitive, the weak tracker also concentrates trading on implementation date. Intuitively, if the weak tracker is small compared to the strict tracker, by trading outside implementation it cannot move prices in its favor. Moreover, because speculators are competitive, the weak tracker cannot realize profitable trading gains given prices and tracking error concerns. This result may explain why trading volume on implementation day far exceeds the trading volume that strict index funds and shadow indexers alone generate.

To better isolate intermediation activity in the data, we examine short interest and securities lending data. The model predicts that speculators acquire shares over multiple periods to minimize their price impact. As a part of this strategy, speculators also acquire shares *after* the implementation day. That is, they short-sell on the implementation day and cover the short in the subsequent periods. Using short-selling data, we isolate genuine intermediation activity from generic asset sales (such as mutual funds attempting to "sell high" to generate alpha). We find that short interest rises on implementation dates and then falls afterwards as speculators cover. We also document the reverse pattern for negative index-induced demand shocks.

In an extension of the model, we examine how increasing the number of periods impacts trading volume around the implementation date. When speculators have fewer trading periods to pre-position ahead of implementation, trading volume becomes concentrated on days closer to implementation. We test this prediction using migrations between the Russell 1000 and 2000. Russell uses a transparent rank formula and a predetermined rank cutoff to determine which stocks will switch indexes and which will remain in their existing index. Some stocks are well beyond the cutoff and the market will correctly infer that the stock is certain to switch, even before the official rank date.

Other stocks are closer to the cutoff, leaving open the possibility that a return shock for that stock or other stocks close to the cutoff might still affect the index they will be assigned to on the rank date.

We use a measure of the pre-rank-date predictability as a way of capturing the number of periods available for speculators to pre-position shares for implementation day. That is, stocks far from the cutoff are available for pre-positioning well in advance of the rank date, while speculators are much more likely to wait until closer to the rank date or even the rank date itself to pre-position shares. We find that less predictable migrations have more trading between the rank and implementation date, as well as after the implementation date, consistent with speculators having less time to pre-position and therefore needing to intermediate over more constrained timeframes.

Finally, thanks to the model, we identify the post-implementation price reversal as the key measure of speculator competition and market efficiency. Intuitively, price reversals measure the *avoidable* execution cost borne by index investors. Empirically, we show that post-implementation reversals for S&P 500 direct additions have fallen sharply over time, and that the decline is even more pronounced when reversals are scaled by the size of the index demand shock, as the model prescribes. Strikingly, index funds trade very cheaply, as they transact extremely large quantities – often more than 20 times average daily volume – at a single point in time with costs that are small relative to standard price-impact benchmarks (Frazzini et al., 2012).

With our model and empirical results, we argue that, to assess the welfare effect of speculators around index reconstitution, one needs to consider a counterfactual market with no speculators. With no speculators pre-positioning liquidity around a predictable benchmark-driven trade, the fund would incur even higher costs, either in terms of tracking error or execution costs.

RELATED LITERATURE. This paper sits at the intersection of research on the index effect, benchmark-driven demand, and strategic liquidity provision.

A first related literature studies price pressure around index additions and deletions. Early work in this area documents significant abnormal returns around reconstitution dates (Harris and Gurel, 1986; Shleifer, 1986; Petajisto, 2011). More recent evidence shows that, in major indices, announcement-day and implementation-day effects have weakened while the price response has shifted earlier (Greenwood and Sammon, 2025). Our contribution is to emphasize why the timing of *prices* and *quantities* has diverged and to offer a mechanism for that divergence. In our model, benchmark-linked investors continue to concentrate their trading on the implementation date, but intermediaries spread their own trading across the surrounding window by building inventory in advance and unwinding short positions afterward. This allows the intermediary to effectively act as an agent for the index fund, smoothing the trades – and thus the returns – over time.

A second literature studies execution and anticipatory trading. Standard models predict gradual trading if the informed investor is expected to have price impact (Kyle, 1985; Almgren and Chriss, 2001), while theories of concentrated trading emphasize coordination during periods of deep liquidity (Admati and Pfleiderer, 1988). Related work highlights the benefits of commitment power in execution (Shim and Wang, 2025), and that predictable trading can invite strategic responses by rent-extracting intermediaries (Brunnermeier and Pedersen, 2005). We bring these ideas together in an index-reconstitution setting where benchmark-tracking concerns create a motive to trade on the implementation date, and anticipatory traders endogenously provide liquidity by pre-positioning for expected demand. In our model, synchronized execution at implementation is an equilibrium outcome because off-event fund trading crowds out intermediary liquidity provision at the close.

A closely related paper is Harvey et al. (2025), who study predictable rebalancing by institutional investors that aim to maintain target stock-bond allocations. They estimate that these predictable cross-asset trades impose costs of $16 billion per year on investors. Their setting, however, differs from ours in two important ways. First, their rebalancing occurs *across* asset classes, whereas ours occurs *within* equities. This distinction matters because the price impact of cross-asset rebalancing is likely governed by a macro-elasticity, while within-equity index rebalancing is likely governed by a micro-elasticity (Gabaix and Koijen, 2021). Because intermediating aggregate stock-market exposure is riskier than intermediating relative demand across individual stocks, macro elasticities are likely smaller than micro elasticities, making expected price impact larger in their setting. Further, their $16 billion estimate captures the total cost of predictable cross-asset rebalancing, including price pressure generated by the rebalancing investors' own trades. By contrast, in our setting, much of the (avoidable) price impact is generated by intermediaries who trade in anticipation of index demand. Our model therefore tells us to focus on the avoidable component of implementation shortfall: the part of price pressure that reverses after the implementation date.

A third literature studies the rise of passive and benchmarked investing and the extent of inelastic demand (Koijen and Yogo, 2019; Coles et al., 2022; Pavlova and Sikorskaya, 2023). Benchmark-driven demand has grown substantially, making it important to understand how markets have adapted to large, predictable rebalancing trades. Consistent with an ecosystem that has adapted to these predictable events (Lo, 2004), trading volume is extraordinarily concentrated on the reconstitution date itself (Chinco and Sammon, 2024). Our model helps rationalize this concentration by emphasizing a coordination channel that is specific to pre-announced and predictable index events. Specifically, index providers like Russell and S&P communicate rules, calendars, and implementation details in advance. This transparency supports a commitment-like environment in which intermediaries can build inventory ahead of the implementation date and supply liquidity when index funds trade to minimize tracking error.

Finally, our paper can inform index design, especially in settings where pre-positioning is constrained. Our mechanism suggests that concentrated rebalancing is less costly when index changes are predictable and intermediaries have time to build inventory before implementation. By contrast, short implementation windows can limit pre-positioning and concentrate price pressure on the implementation date. Consistent with this logic, Murray and Sammon (2026) show that fast-track IPO inclusions, which are costly to short and do not have many periods between IPO and index inclusion, amplify implementation-day price impact and the subsequent reversals relative to a world with fewer frictions for intermediaries.

2 Data and Variables

This subsection describes the data sources, sample construction, and the event-time datasets used in our analyses of index front-running and rebalancing around S&P and Russell index reconstitution. We construct the data separately for (i) S&P direct additions and deletions, (ii) S&P tier migrations (e.g., between large-, mid-, and small-cap indices), (iii) Russell migrations between the Russell 1000 and Russell 2000, and (iv) Russell direct additions and deletions from outside the Russell universe. Constructing the data separately reduces the risk of overlap, as some stocks may be added to or dropped from multiple indexes within the same event window.

We also clarify definitions and the language used in this paper. We use the term *reconstitution* to indicate the change in the composition of a certain index. In the context of the S&P 500 index, a reconstitution involves the *addition* of a stock to the index and the *deletion* of a stock. Many additions and deletions, particularly in the context of the Russell 1000 and 2000 indexes, are *migrations* between indexes.

We use the term *implementation date* to refer to the trading day when the index provider implements the reconstitution and, hence, the updated index becomes effective. We use the term *rebalancing* to specifically indicate the trading activity of index-tracking investors who trade on or near the implementation date in order to align their portfolio holdings with the updated index.

2.1 Data

S&P 500 reconstitution events. S&P 500 additions and deletions are based on index-change lists provided by the index provider and include both announcement and effective dates. We also identify a set of unusual cases that are not typical index changes because they are driven by corporate events such as mergers and associated delistings. In the baseline analysis we remove these unusual cases. In some analyses we additionally separate S&P changes associated with regularly scheduled quarterly rebalances from ad hoc changes that occur when a slot opens up due to a merger or other corporate event.

The S&P event data also indicate whether an event reflects a movement across S&P size tiers. In our baseline analyses, we treat migrations as movements between the S&P 500 and the mid-cap 400 or small-cap 600. In many analyses, we separate these from stocks that enter from or exit to outside the S&P 1500 index family.

Russell 1000/2000 migrations. Russell migration events are based on index-change lists provided by the index provider and Russell reconstitution files that provide the annual rank date (when membership is determined) and the effective reconstitution date (when index constituent changes take effect). We classify each event as either a move into the Russell 2000 from the Russell 1000, or the reverse move into the Russell 1000 from the Russell 2000. We separately analyze Russell direct events. These are cases where firms enter or exit the Russell 1000 or Russell 2000 from outside the Russell universe.

Index Returns For return-based outcomes, we benchmark each event against the return of the relevant index. For S&P events, we use the S&P index return that matches the event (e.g., S&P 500 for S&P 500 additions/deletions and S&P tier returns for tier migrations). For Russell events, we use the Russell 1000 return for Russell 1000 additions and the Russell 2000 return for Russell 2000 additions, and analogously benchmark migration events against the destination/source index as appropriate. We obtain all daily total index return data from Datastream.

Stock-Level Data and Institutional Holdings Our daily stock-level outcomes come from CRSP and are organized at the security-trading-day level.

2.2 Main Data Set

We build a single main data set at the security-trading-day level that combines CRSP data with the reconstitution event-time datasets. Unless stated otherwise, the baseline event-time panels require a complete event window for each stock and event. Specifically, we require the stock to have trading-day data for all $\tau \in [-120, +20]$ relative to the implementation date. This mechanically excludes firms that list recently or otherwise do not have a full pre-reconstitution history, such as some spinoffs and corporate events associated with mergers and related index changes.

Trading Calendar and Date Alignment Index dates (announcement, rank, effective) can fall on weekends or market holidays. Our convention is to map any non-trading event date to the next trading day (e.g., Saturday to Monday). We implement this using a CRSP-based trading calendar and a next-trading-day map from calendar dates to valid CRSP trading dates. This alignment is applied before constructing event windows.

Event-Time Panel Construction For each reconstitution event group we construct a long event-time panel that indexes observations by event, by the implementation date, and by event time (days relative to the anchor). We use the same procedure for each group. First, we clean the raw event information and align announcement, rank, and implementation dates to trading days. Second, we assign a unique event identifier and create separate versions of each event anchored on the relevant dates (announcement versus implementation for S&P, and rank versus implementation for Russell). Third, we expand each anchored event into a fixed event window and map event time to trading dates using the trading calendar. Lastly, we merge in daily outcomes from the main data set.

Sample Selection and Target Event Buckets Our baseline figures and tables group events into four conceptually distinct categories. First, for the S&P 500 we study direct additions and direct deletions. Second, we study S&P tier migrations between the S&P 500 and adjacent size tiers. Third, for Russell we study direct additions and direct deletions into the Russell 1000 and into the Russell 2000, meaning entry from or exit to outside the Russell universe. Fourth, we study Russell migrations between the Russell 1000 and Russell 2000. For S&P analyses, we exclude unusual cases (e.g., those driven by listings, delistings, mergers, acquisitions, etc.) by default, and conduct robustness checks also excluding ad-hoc additions/deletions (i.e., additions/deletions which do not occur at the regular quarterly rebalancing events).

2.3 Event-Time Outcome Variables

This subsection describes the construction of the main event-time outcomes used in the analysis. These include scaled trading volume and cumulative abnormal returns.

Scaled Volume Scaled volume is defined as shares traded divided by shares outstanding, or

$$scaledvolume_{it} = \frac{vol_{it}}{shrout_{it} \cdot 1000}.$$

Cumulative Abnormal Returns (CAR) For a given event, define the stock cumulative return from the anchor date to τ as

$$R^{\text{stock}}(\tau) = \prod_{s=0}^{\tau}(1 + r_s^{\text{stock}}) - 1,$$

and define the cumulative return on the appropriate index benchmark over the same days as

$$R^{\text{bench}}(\tau) = \prod_{s=0}^{\tau}(1 + r_s^{\text{bench}}) - 1.$$

Cumulative abnormal return is then

$$\text{CAR}(\tau) = R^{\text{stock}}(\tau) - R^{\text{bench}}(\tau).$$

A key implementation detail is that we cumulate stock and benchmark returns separately and subtract *after* cumulation.

Benchmarks and anchors We use index-specific benchmarks that match the event. S&P 500 events are benchmarked to the S&P 500 return. Russell 1000 direct additions and deletions are benchmarked to the Russell 1000 return, and Russell 2000 direct additions and deletions are benchmarked to the Russell 2000 return. For migration events, we benchmark to the index most directly tied to the event. For S&P events, we anchor event time on the announcement date. For Russell events, we anchor event time on the effective reconstitution date.

Winsorization To reduce the influence of extreme outliers, we winsorize the outcome variables within each event time and decade. Our baseline approach uses one-sided winsorization at the 99th percentile.

3 MOTIVATING FACTS

In this section, we provide background and institutional details on indexes that are commonly and explicitly tracked by index funds and ETFs, and implicitly tracked by mutual funds and institutional investors. We then outline several facts on index investing that directly inform our model.

3.1 BACKGROUND AND INSTITUTIONAL DETAILS

Indexes are a common way to track segments of the market across and within asset classes, as well as serve as benchmarks for asset managers. Over the last thirty years, they have grown as a way to construct investment portfolios. Among the most common indexes used for investing are the S&P 500 index and the Russell 2000 index.

Indexes add and drop constituents from the indexes, usually at regular intervals. The ways that indexes reconstitute vary, ranging from complete transparency following a published formula (FTSE Russell indexes) to using subjective decision-making with limited information about selection criteria (S&P indexes). Depending on the policy, market participants find it simple or difficult to predict what changes will be made to an index in advance. Most indexes have regularly scheduled dates when the index is adjusted, where the scheduled reconstitution is announced at some point before the implementation date of the index reconstitution. There are some exceptions to this regular schedule, usually where index adjustments are made because of certain types of market events (e.g., a sudden merger, acquisition, or bankruptcy).

There are many types of investors or investment vehicles that follow indexes. We will broadly refer to these investors as *index trackers*. The most strict index trackers are index funds and ETFs. Most of these investment vehicles replicate the index with very little tracking error, and specifically cite an explicit goal to minimize tracking error. For index funds and ETFs that track indexes with relatively liquid underlying securities (e.g., U.S. equities), the tracking error is relatively small (Sammon and Shim, 2026), but can be larger due to selective replication when the underlying securities are illiquid (e.g., high-yield corporate bonds).

There are other index trackers that are not as strict. Many investors now engage in direct replication, where an investor or institution mimics an index but with some flexibility on several dimensions. These dimensions include holding only a representative subset of the index, strategically trading securities to engage in tax-loss harvesting, avoiding illiquid securities, and timing the purchase and sale of index securities. The reasons for market timing vary, from economizing on trading costs and index-provider fees, or because mandates, risk-systems, and compensation are tied to benchmarks (Pavlova and Sikorskaya, 2023). These index trackers may differ in how closely they follow the index, but are still considered index trackers in that, at any given point in time, their portfolios will largely be similar to the tracked index.

The last group of index trackers we highlight are so-called "closet indexers." These investors do not disclose that they are indexers and may even state that they are active investors, selecting securities they believe will outperform an index or benchmark. However, their portfolio holdings appear to closely replicate some index (and often the index they are benchmarked against). We think of closet indexers as also varying how closely they track the index, perhaps based on the degree of manager risk aversion or agency issues – if these factors are large, they may track the index as closely as index funds or ETFs.

More broadly, predictable rule-based rebalancing is not limited to stock market indexes. For example, many institutional investors maintain stock-bond portfolios with target weights, such as 60% invested in stocks and 40% invested in bonds, and rebalance at regular calendar dates or when portfolio weights deviate significantly from predetermined targets (Harvey et al., 2025). While these investors do indeed follow a predictable rule-based rebalancing policy, our main analysis focuses on index-linked rebalancing within equities, where investors substitute between individual stocks rather than between aggregate asset classes.

Each index typically specifies index-eligible shares – a type of adjusted shares outstanding – for each constituent stock. Index-eligible shares combined with prices are used to compute the index price level and index returns. Prices are almost always drawn from official closing prices when they are available. For U.S. equities, these prices come from the closing auction that marks the end of regular-hours trading for each stock. For more illiquid securities without official closing prices, indexes can use prices that come from independent pricing services or market snapshots.

The price that indexes use is relevant for index trackers that attempt to exactly replicate the index. Because these strict index trackers minimize tracking error, they attempt to buy and sell securities at the same price that the index uses. For example, these index trackers will often attempt to buy adds and sell drops for U.S. equity indexes at the closing auction in the exact units they need for rebalancing to incur as close to zero tracking error as possible.

3.2 Motivating Facts

In this subsection, we provide two facts on trading volume and prices around implementation days of index reconstitutions. The first fact is that there is a huge burst of trading volume on the implementation day itself, with much less trading before or after the event. The second fact is that, despite all this volume, prices move very little. Collectively, these two facts present a puzzle that motivates our model.

Fact 1: Index trackers trade on the index implementation day.

The first fact is that there is a huge amount of trading volume on the exact day index reconstitutions take place, which we refer to as the implementation day. When an index adds or drops a constituent, index trackers tend to trade in the closing auction of the implementation day. This allows index trackers to buy and sell securities at the same price that the index uses in constructing the index price level, and therefore minimizes tracking error.

Figure 1 shows volume and cumulative abnormal returns around implementation days for S&P 500 and Russell 1000/2000 index events. The S&P 500 events are separated into direct additions and deletions, which enter or exit the S&P 500 from outside the S&P index universe, and migration additions and deletions, which move between the S&P 500 and another S&P size index (the S&P 400 midcap or the S&P 600 smallcap). For Russell events, we focus on migrations between the Russell 1000 and Russell 2000. For both sets of events, we focus on reconstitutions taking place between 2010 and 2024.

The figure shows that, in every case, there is an extremely large spike in volume on the implementation day. For S&P 500 direct additions, nearly 20% of shares outstanding trade on the implementation day. For S&P 500 deletions and migrations, implementation-day volume is also extremely large, ranging from roughly 18% to 26% of shares outstanding. Russell migrations have smaller levels of implementation-day volume, but still show a clear and large spike relative to surrounding days.

Surprisingly, there is not a comparable run-up in volume before implementation, nor is there a comparable continuation of volume after implementation. This is not surprising for index funds and ETFs, i.e., strict index trackers. Most of these trackers have very small tracking error budgets. If the index uses the closing price on the implementation day, then these trackers naturally want to trade at that same closing price.

What is more surprising is that more flexible index trackers and closet indexers also seem to trade on implementation day. Chinco and Sammon (2024) document that the volume in the closing auction on implementation days is double what we would expect from strict index trackers alone. This suggests that implementation-day trading is not limited to mechanical index funds and ETFs. It also appears to include other benchmark-linked investors that have at least some flexibility in when to trade.

Fact 2: There is little to no return on index implementation day, despite enormous trading volume.

The second fact is that implementation-day returns are small despite this extraordinary volume. Given the significant demand for newly added index constituents and the large amount of trading on the implementation day, it seems sensible to believe that the price movement would have to be large to clear the market. In other words, if index trackers buy or sell up to 20% of all shares outstanding for a stock, and do so on a single day, we would expect significant price impact for such extreme demand. For context, average daily trading volume for S&P 500 stocks is on the order of 1% of shares outstanding.

Historically, there were large returns around reconstitution in the early years of index-linked investing. Shleifer (1986) and Harris and Gurel (1986) document that in the 1970s and 1980s, the index addition price effect for S&P 500 stocks was large, with an abnormal one-day return of around 3%. However, as the dollars devoted to indexing have grown, the price movements on announcement and implementation days themselves have become *smaller*, not larger (Greenwood and Sammon, 2025).

Figure 1 shows this fact in the modern, post-2010 sample. The bars show the implementation-day volume spike, while the black line shows cumulative abnormal returns around the same event. For S&P 500 direct additions, the average abnormal return on the implementation day is only -0.09%. For S&P 500 migration additions, it is -0.24%. For Russell migrations, the average implementation-day abnormal returns are similarly small: 0.10% for stocks moving from the Russell 1000 to the Russell 2000 and 0.11% for stocks moving from the Russell 2000 to the Russell 1000.

The S&P 500 deletion panels are somewhat different, but they point to the same tension. Direct deletions have a larger implementation-day abnormal return of -1.48%, and they also show a large negative pre-event return. These are unusual cases, however, as firms directly deleted from the S&P 500 are firms that historically were large, but became so small they are no longer eligible to be included even in the S&P 600 small cap. Migration deletions, by contrast, have an implementation-day abnormal return of only -0.45%. Thus, even in the deletion panels, where the events are less clean than additions, the implementation-day return is modest relative to the amount of trading.

The figure also shows that returns are not completely flat before implementation. For example, S&P 500 direct additions have some positive run-up before implementation, while S&P 500 direct deletions have a pronounced decline. These patterns are not surprising. For S&P events, there can be announcement effects before implementation. For Russell events, the ranking date is more predictable, and pre-rank returns mechanically affect index assignment. A stock that moves from the Russell 1000 to the Russell 2000 has typically performed poorly relative to the cutoff, while a stock that moves from the Russell 2000 to the Russell 1000 has typically performed well. Thus, some of the pre-event price movement likely reflects selection into the index event itself, rather than only anticipation of index-tracker demand.

The key fact is therefore not that prices don't move at *any* point before implementation. The key fact is that, on the implementation day itself, prices move very little relative to the amount of trading volume. The implementation day is when index trackers appear to trade, yet it is not when we see the kind of price movement that would normally be associated with trading 20% of shares outstanding.

This is especially striking when compared with estimates of price impact for institutional investors. For example, Frazzini et al. (2012) show that average price impact from trading 6% of average daily volume is around 30 basis points, and trading 11% of average daily volume is around 40 basis points. The index trades in our sample are orders of magnitude larger: implementation-day trading can be roughly 20 to 25 *times* average daily volume. The largest volume bin in Frazzini et al. (2012) is 11% of average daily volume, while these index events involve trading that is closer to 2,000% to 2,500% of average daily volume. Extrapolating from the estimates in Frazzini et al. (2012), we might expect to see price impact of over 6,800 basis points. Instead, we see little to no implementation-day return.

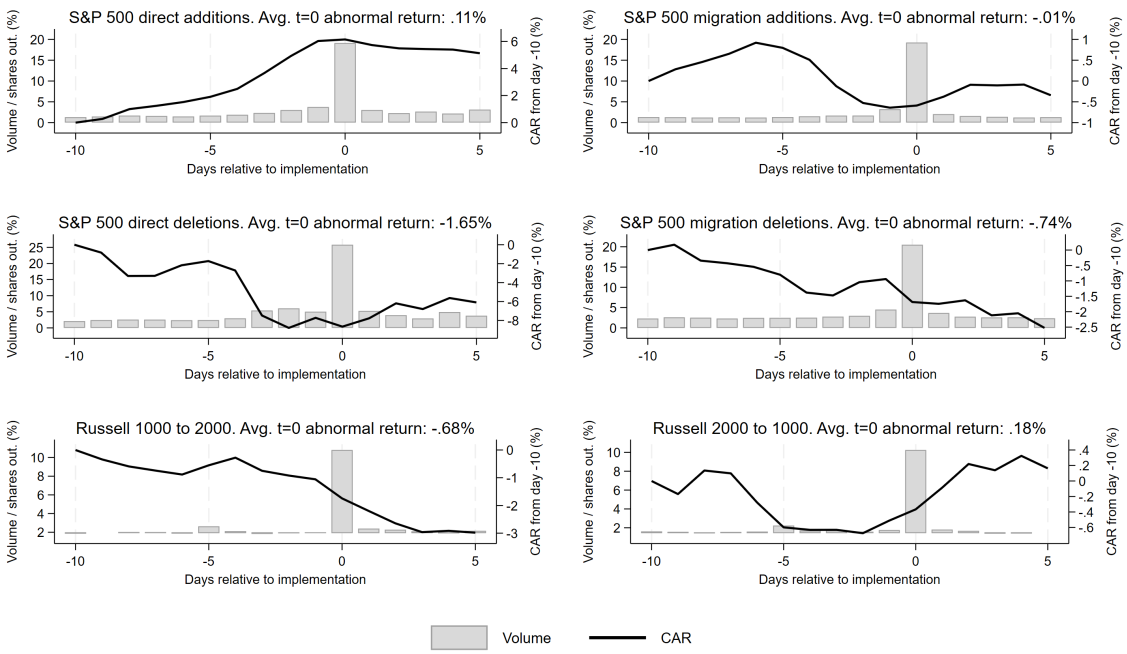
This does not mean that index trackers face zero trading costs. The implementation-day return is not the same thing as the total cost of trading. If speculators acquire shares in advance and sell them to index trackers at implementation, then the cost to index trackers may be embedded in the pre-event price rather than in the implementation-day return itself. Measuring the actual trading cost therefore requires a model of who trades before implementation, who supplies liquidity at implementation, and what the counterfactual price path would have been without anticipatory trading.

3.3 Discussion

The two indexing facts together present a puzzle. Index trackers – including loose trackers who have flexibility in when to trade – all seem to trade on the index implementation day. But prices do not jump on implementation day despite the significant volume.

So, the natural question is: why do prices not move? The amount of shares traded on

Figure 1: Volume and Returns Around Index Implementation Days



Notes. This figure shows average daily volume and cumulative abnormal returns around index implementation days. The gray bars show daily volume scaled by shares outstanding, in percent, and correspond to the left axis. The black line shows cumulative abnormal returns from day -10, in percent, and corresponds to the right axis. We separately show S&P 500 direct additions, S&P 500 migration additions, S&P 500 direct deletions, S&P 500 migration deletions, Russell 1000 to Russell 2000 migrations, and Russell 2000 to Russell 1000 migrations. S&P 500 direct additions and deletions enter or exit the S&P 500 from outside the S&P index universe. S&P 500 migration additions and deletions move between the S&P 500 and another S&P size index. Russell migrations move between the Russell 1000 and Russell 2000. The figure uses index events from 2010 to 2024.

the implementation day represents a demand shock that is, by any measure, extremely large and inelastic. Since the shock is widely anticipated, it is possible that the shares are provided with little additional price concession on the actual implementation day. If the price movement occurs before implementation, this implies that index trackers may buy at high prices and sell at low prices, which can be interpreted as a transaction cost or implementation shortfall. Still, it is striking to see 5% to 20% of a stock's shares outstanding changing hands with little contemporaneous price movement.

This in turn suggests a natural second question: why do index trackers trade on implementation day instead of in the days before? If prices rise leading up to an addition, it seems like they would benefit from trading before the implementation day itself. This seems especially valuable for loose index trackers and/or "enhanced" indexers that have less strict tracking error constraints and have flexibility to pay, what seems like, a slightly lower price by trading slightly early or late (Arnott et al., 2023).

Given the volume we observe on implementation dates, however, a significant proportion of index trackers do not appear to trade outside of the implementation date. If anyone trades in the days or months before the implementation day, it is likely speculators, who provide the shares to the index trackers on the implementation day itself. The final question, therefore, is whether speculator activity is sustainable for index trackers in equilibrium, and whether this activity represents an excessive cost to investors.

At a glance, it seems as though speculators are imposing a cost on index trackers because their behavior resembles front running (or predatory trading more broadly) (Brunnermeier and Pedersen, 2005). On the other hand, we do not know the counterfactual, which is what price impact and trading costs would be for index trackers if speculators did not acquire the shares in advance and provide shares on the implementation day, or if index trackers trading became less predictable.

The model that follows is designed to discipline this counterfactual. It asks why an index tracker may rationally concentrate trading at implementation even when prices appear to move beforehand, and how speculators who trade in advance can simultaneously earn profits, reduce implementation-day price impact, and reduce the total cost borne by index investors.

4 Model

We consider a model with two types of agent. An index fund that needs to rebalance their portfolios in response to an index change, and a set of speculators that can trade in anticipation of the fund's demand. We study the fund's optimal order execution when they have a price impact and speculators can trade in anticipation of the fund's demand. All proofs are in the Internet Appendix IA.1, which is available upon request.

4.1 Model Setup

The model spans five periods, $t \in \{0, 1, 2, 3, 4\}$. Period 0 is the announcement date: the index change is publicly announced, but no trading occurs and all positions are at their long-run fundamental values. Period 1 is the pre-implementation trading phase, during which the fund and speculators may establish positions in anticipation of the upcoming index change. Period 2 is the implementation date, on which the index change takes effect and the fund executes its mandatory rebalancing trade. Period 3 is the post-implementation trading phase, in which remaining positions may be adjusted. Period 4 is a terminal payoff period: no trading occurs, but each security pays its fundamental value

$V_{J,4}$ per share to its holders. Because the terminal payoff equals the fundamental value rather than the prevailing market price, any price impact incurred during periods $\{1,2,3\}$ represents a real cost that is not recovered at the terminal date.

Two assets trade in the market. The *original security* (O) is the incumbent index constituent, and the *new security* (N) is the incoming constituent, added to the index in period 2. In period 2, the index substitutes N for O: the target shifts from Q shares of O and zero shares of N to zero shares of O and Q shares of N. Accordingly, the fund must sell Q shares of O and purchase Q shares of N. We require that the fund's portfolio be fully rebalanced by the end of period 3: $\sum_{t=1}^{3} x_t = Q$.

The market contains a *fund* (F) and $L \geq 1$ *identical speculators*. The fund minimizes the sum of net execution cost and tracking-error cost. It may trade both securities in any period $t \in \{1,2,3\}$: let x_t^N denote N-shares purchased and x_t^O denote O-shares traded (negative for sales) in period t. We focus on the symmetric case $x_t \equiv x_t^N = -x_t^O$, so the fund buys x_t shares of N and sells x_t shares of O in each period. This *lockstep* assumption is natural when the fund finances new-security purchases with proceeds from original-security sales, as it minimizes cash-flow mismatch and keeps the fund's total portfolio value approximately constant throughout the rebalancing.

Each speculator $l \in \{1,\ldots,L\}$ maximizes trading profit net of execution and holding costs. Speculators may trade both securities in any period $t \in \{1,2,3\}$: let $s_{N,t,l}$ and $s_{O,t,l}$ denote speculator l's N-side and O-side trades in period t, respectively. Speculators are required to close all positions by the end of period 3, so that cumulative holdings satisfy $S_{N,3,l} = \sum_{\tau=1}^{3} s_{N,\tau,l} = 0$ and $S_{O,3,l} = \sum_{\tau=1}^{3} s_{O,\tau,l} = 0$ for every speculator l. Speculators compete in *Cournot* fashion, taking the quantities of other speculators as given. We characterize the symmetric equilibrium in which all speculators adopt identical strategies, $s_{N,t,l} = s_{N,t}$ and $s_{O,t,l} = s_{O,t}$ for all l, so that aggregate speculator order flows are $S_{N,t} = L\, s_{N,t}$ and $S_{O,t} = L\, s_{O,t}$.

We analyze two equilibrium concepts that differ in the degree of commitment available to the fund. Under *no commitment* (NC), the fund selects its trading schedule taking speculator positions as given, and equilibrium is determined by the mutual best responses of all players simultaneously. Under *full commitment* (FC), the fund publicly commits to its complete trading schedule (x_1, x_2, x_3), satisfying $x_1 + x_2 + x_3 = Q$, before speculators act. Each speculator then observes the committed schedule and chooses trades $(s_{N,t,l},\, s_{O,t,l})$ for $t \in \{1,2,3\}$ to maximize profit, taking the fund's schedule and other speculators' quantities as given. The fund selects the cost-minimizing schedule in anticipation of this best-response behavior, and the commitment is credible at the execution stage. The FC equilibrium thus has a Stackelberg structure in which the fund moves first.

4.1.1 Fundamental Values and Prices

Each security has a fundamental (long-run) value that evolves as a martingale with IID innovations,

$$V_{J,t} = V_{J,t-1} + \varepsilon_{J,t}\,, \qquad J \in \{N, O\}\,, \tag{1}$$

where $\{\varepsilon_{J,t}\}$ are IID across time with $\mathbb{E}[\varepsilon_{J,t}] = 0$ and finite variance; innovations need not be independent across securities. We write $M_{J,t} \equiv \sum_{\tau=1}^{t} \varepsilon_{J,\tau}$ for the cumulative martingale innovation, so that $V_{J,t} = \bar{v} + M_{J,t}$. Let $P_{J,t}$ denote the transaction price of security $J \in \{N, O\}$ at time t. All transaction prices and fundamental values are initialized at the common level $\bar{v}$ at $t = 0$:

$$V_{N,0} = V_{O,0} = P_{N,0} = P_{O,0} = \bar{v}\,. \tag{2}$$

A key parameter is the expected absolute difference in fundamental innovations between N and O,

$$\nu \;\equiv\; \mathbb{E}\big[|\varepsilon_{N,t} - \varepsilon_{O,t}|\big] > 0\,, \tag{3}$$

which is strictly positive whenever $\varepsilon_{N,t} \neq \varepsilon_{O,t}$ with positive probability—the natural case for distinct securities. In the terminal period 4, each security pays its fundamental value $V_{J,4} = \bar{v} + M_{J,4}$ to its holders.

Prices are governed by a reduced-form linear price-impact function that captures how net order flow affects transaction prices at each trading date as in Almgren and Chriss (2001). Let $z_{J,t}$ denote the net order flow in security J in period t,

$$z_{J,t} = x_t^J + \sum_{l=1}^{L} s_{J,t,l},$$

and let $Z_{J,t} = \sum_{\tau=1}^{t} z_{J,\tau}$ denote the cumulative net order flow up to period t, with $Z_{J,0} = 0$. The transaction price is

$$P_{J,t} = \bar{v} + \lambda_J\, Z_{J,t} + \rho_J\, z_{J,t} + M_{J,t}\,, \tag{4}$$

where λ_J is the persistent price-impact coefficient and ρ_J is the temporary price-impact coefficient. For algebraic convenience, we also define

$$\mu_J \equiv \lambda_J + \rho_J$$

which measure the total within-period price impact of flows. We allow $\lambda_O \leq \lambda_N$ and

$\rho_O \leq \rho_N$, so that $\mu_O \leq \mu_N$, reflecting the possibility that the original security is a more liquid position with lower total per-unit price impact than the new constituent. When order flow is negative (net selling), the transaction price falls accordingly.

When the fund trades both securities in lockstep, $x_t^N = x_t$ and $x_t^O = -x_t$, the effective within-period impact is $\mu_N + \mu_O$. No trading occurs in period 4. For later use, define

$$\bar{P}_N \equiv \bar{v} + \lambda_N Q, \tag{5}$$

the N-security price after all trading has taken place and cumulative N-side net order flow equals Q.

4.1.2 Costs and Objectives

The fund's total cost has two components: an implementation shortfall that captures price-impact costs, and a tracking-error penalty that penalizes deviations from the index between trading dates.

Tracking error. The fund's tracking error in period t measures the expected absolute return gap between the fund's portfolio and the underlying index. Because both securities pay their fundamental values at the terminal date, the relevant source of divergence is the martingale innovation component of price changes rather than any price-impact term. Let $h_{N,t}$ and $h_{O,t}$ denote the fund's holdings of N and O at the end of period t, with initial positions $h_{N,0} = 0$ and $h_{O,0} = Q$. The index holds Q shares of O for $t < 2$ and Q shares of N from period 2 onward. The per-period tracking error is

$$\mathrm{TE}_t = \nu \cdot |h_{N,t-1} - h_{N,t-1}^{\mathrm{idx}}|, \tag{6}$$

where $\nu = \mathbb{E}[|\varepsilon_{N,t} - \varepsilon_{O,t}|] > 0$ is the expected absolute innovation gap defined in (3). Because the fund and the index hold identical portfolios at $t = 0$, no tracking error arises in period 1. From period 2 onward, each share of N held instead of O (or vice versa) generates an expected absolute return gap of ν per period. Equivalently, $\kappa\,\mathrm{TE}_t$ can be interpreted as a Lagrange multiplier representation of a hard constraint on portfolio deviations from the index: any fund required to hold no more than a given number of shares away from the index weight will face a shadow cost proportional to that deviation, regardless of the return volatility or cross-security correlation of the two assets. The parameter ν scales the expected return consequence of a one-share deviation, but the linearity of the penalty in share deviations is a property of the position constraint itself, not of any particular assumption on ν.

Lemma 1 (Tracking-error formulas). *Under the lockstep schedule,* $\mathrm{TE}_1 = 0$ *because the fund and the index hold identical portfolios at* $t = 0$*;* $\mathrm{TE}_2 = \nu\, x_1$ *because the fund holds* x_1 *excess shares of N entering period* 2*; and* $\mathrm{TE}_3 = \nu\, x_3$ *because the fund holds* $Q - x_1 - x_2 = x_3$ *excess shares of O entering period* 3*. Tracking error is linear in the fund's off-implementation trades, and price impact does not enter it. The total tracking-error cost is therefore*

$$\kappa \sum_{t=1}^{3} \mathrm{TE}_t = \kappa\, \nu\, (x_1 + x_3), \tag{7}$$

where $\kappa > 0$ *is the tracking-error penalty parameter.*

Implementation shortfall. The fund's *implementation shortfall* IS^F is the total execution cost of buying Q shares of N and selling Q shares of O across periods $\{1, 2, 3\}$, measured relative to the initial fair value $\bar{v}$. Because the terminal payoff equals the fundamental value, any price impact incurred during trading represents a real cost not recovered at the terminal date. For period t trading of x_t shares, the cost is the price-impact component of the current transaction price: $\lambda_N\, x_t\, Z_{N,t} + \rho_N\, x_t\, z_{N,t}$ on the buy side and $-\lambda_O\, x_t\, Z_{O,t} - \rho_O\, x_t\, z_{O,t}$ on the sell side. The compact coefficients $\mu_N = \lambda_N + \rho_N$ and $\mu_O = \lambda_O + \rho_O$ are used below when they shorten formulas. The total implementation shortfall is

$$\mathrm{IS}^F = \sum_{t=1}^{3} \Big[\lambda_N\, x_t\, Z_{N,t} + \rho_N\, x_t\, z_{N,t} - \lambda_O\, x_t\, Z_{O,t} - \rho_O\, x_t\, z_{O,t}\Big], \tag{8}$$

with $Z_{N,0} = Z_{O,0} = 0$. The cross-period terms depend on cumulative order flows that include speculator activity, which is the channel through which the fund's cost is affected by equilibrium speculator strategies (Huberman and Stanzl, 2004).

Fund's objective. The fund chooses the trading schedule (x_1, x_2, x_3) to minimize the sum of implementation shortfall and tracking-error cost:

$$C^F = \mathrm{IS}^F + \kappa \sum_{t=1}^{3} \mathrm{TE}_t = \mathrm{IS}^F + \kappa\, \nu\, (x_1 + x_3), \tag{9}$$

where the second equality uses (7).

Speculator's objective. Each speculator l maximizes trading profit net of execution and holding costs. The execution cost from N-side trades of $s_{N,t,l}$ shares in period t is $\lambda_N\, s_{N,t,l}\, Z_{N,t} + \rho_N\, s_{N,t,l}\, z_{N,t}$, with the O-side defined analogously. The compact coefficient $\mu_N = \lambda_N + \rho_N$ is used in the algebra below where convenient. In addition, each speculator bears a holding cost proportional to the square of cumulative positions,

$$\mathrm{HC}_l^S = c \sum_{t=1}^{3} \big[(S_{N,t,l})^2 + (S_{O,t,l})^2\big], \tag{10}$$

where $S_{N,t,l} = \sum_{\tau=1}^{t} s_{N,\tau,l}$ and $S_{O,t,l} = \sum_{\tau=1}^{t} s_{O,\tau,l}$ are cumulative positions, and $c \geq 0$ captures financing, margin, and inventory risk. Because speculators must close all positions by the end of period 3, the terminal payoff terms that would otherwise appear in the profit expression vanish identically. The per-speculator profit across both securities therefore reduces to the negative of total execution cost minus holding cost:

$$\Pi_l^S = -\sum_{t=1}^{3} \big[P_{N,t}\, s_{N,t,l} + P_{O,t}\, s_{O,t,l}\big] - \mathrm{HC}_l^S, \tag{11}$$

where $P_{J,t}$ is the execution price from (4). The speculator profits when the period-2 execution gain—from selling into the fund's demand at temporarily elevated prices—exceeds the combined execution and holding costs from periods 1 and 3.

4.1.3 Strategies and Payoffs

In the symmetric equilibrium, each speculator employs a *three-period strategy* that closes all positions by the end of period 3. On the N-side, the strategy is parameterized by $s_1 \equiv s_{N,1}$, the pre-positioning purchase in period 1, and $r \equiv s_{N,3}$, the covering trade in period 3; the implied period-2 trade $-(s_1 + r)$ sells the accumulated position into the fund's implementation demand. The O-side is symmetric, with $s_{O,1}$ denoting the period-1 short and $r_O \equiv s_{O,3}$ the period-3 re-short. Aggregate positions across L speculators are

$$S_1 \equiv L\, s_1, \quad R \equiv L\, r \quad \text{(N-side);} \qquad S_{O,1} \equiv L\, s_{O,1}, \quad R_O \equiv L\, r_O \quad \text{(O-side).} \tag{12}$$

By construction, cumulative N-side holdings per speculator are s_1 at the end of period 1, $-r$ at the end of period 2, and 0 at the end of period 3; the O-side is symmetric. All positions close by design, confirming that the terminal-payoff terms in (11) vanish and that the holding cost per speculator is $c\,(s_1^2 + r^2)$ on the N-side plus $c\,(s_{O,1}^2 + r_O^2)$ on the O-side.

Suppressing zero-mean fundamental innovations, the net order flows in each period are

$$\text{Period 1:} \quad z_{N,1} = x_1 + S_1, \qquad z_{O,1} = -x_1 + S_{O,1};$$

$$\text{Period 2:} \quad z_{N,2} = x_2 - S_1 - R, \qquad z_{O,2} = -x_2 - S_{O,1} - R_O;$$
$$\text{Period 3:} \quad z_{N,3} = x_3 + R, \qquad z_{O,3} = -x_3 + R_O.$$

Execution prices follow from the pricing function (4). Period 4 involves no trading; holders receive the fundamental value $V_{J,4}$.

4.2 Equilibrium Analysis

We now characterize equilibrium. We begin by deriving the symmetric Cournot best-response functions for $L \geq 1$ speculators on each side, taking the fund's schedule as given. Then, we characterize the fund's optimal trading schedule under no commitment and full commitment, respectively. In both cases, if the tracking error penalty is sufficiently high, the fund optimally concentrates all trading at implementation.The full-commitment threshold is below its no-commitment counterpart: by publicly announcing its schedule, the fund can secure the implementation-day outcome under a strictly broader set of parameters.

4.2.1 Speculator's General Three-Period Strategy

Next, we characterize the speculators' symmetric best responses to an arbitrary trading schedule implemented by the fund.

Proposition 1 (L-speculator symmetric Cournot equilibrium)**.** *Suppose $\rho_N > 0$ (and correspondingly $\rho_O > 0$ on the O-side). With $L \geq 1$ identical speculators in symmetric Cournot competition and an arbitrary fund schedule (x_1, x_3) (with $x_2 = Q - x_1 - x_3$):*

1. Equilibrium positions at the implementation-date rebalancing schedule. *At $(x_1, x_3) = (0, 0)$, the aggregate equilibrium positions are*

$$S_1(L) = L\, s_1(L), \qquad R(L) = L\, r(L), \tag{13}$$

   with $s_1(L) > 0$ for all $L \geq 1$. Closed-form expressions for $s_1(L)$ and $r(L)$ are given in (IA.7) *in the proof.*

2. Affine best-response structure. *For a general schedule (x_1, x_3), the per-speculator best responses are affine corrections to the implementation-date rebalancing positions:*

$$s_1^*(x_1, x_3) = s_1(L) + \frac{\alpha_{s,1}\, x_1 + \alpha_{s,3}\, x_3}{\Delta(L)}\,, \tag{14}$$

$$r^*(x_1, x_3) = r(L) + \frac{\alpha_{r,1}\, x_1 + \alpha_{r,3}\, x_3}{\Delta(L)}\,.,\tag{15}$$

where $\Delta(L) > 0$ is defined in (IA.6) *in the proof. All correction coefficients are negative except $\alpha_{r,1}$, which has ambiguous sign (see the proof for explicit expressions). Hence fund trades strictly reduce speculator equilibrium positions on both the accumulation and close-out margins.*

3. Crowding out: aggregate liquidity provision decreases when the fund trades outside of the implementation date. *The aggregate N-side liquidity provision on the implementation date $\Sigma_N^*(x_1, x_3) \equiv S_1^*(x_1, x_3) + R^*(x_1, x_3)$ is strictly decreasing in both x_1 and x_3:*

$$\frac{\partial \Sigma_N^*}{\partial x_j} = \frac{L\,(\alpha_{s,j} + \alpha_{r,j})}{\Delta(L)} < 0, \qquad j \in \{1, 3\}.\tag{16}$$

That is, the more the fund trades outside of the implementation date, the less total speculative liquidity is available at implementation.

4. Fund trades increase net order flow. *Each aggregate derivative is bounded below by -1:*

$$\frac{\partial S_1^*}{\partial x_j},\ \frac{\partial R^*}{\partial x_j},\ \frac{\partial \Sigma_N^*}{\partial x_j}\ >\ -1, \qquad j \in \{1, 3\}.\tag{17}$$

Hence, the net order flow $z_{N,t}$ is strictly increasing in fund trades outside of the implementation date.

The O-side positions and best-response functions are given by the same expressions with $(\rho_O, \lambda_O, \mu_O)$ replacing $(\rho_N, \lambda_N, \mu_N)$.

The crowding-out result established in Part 3 is the central economic mechanism linking the two equilibrium concepts. When the fund trades before or after implementation ($x_1 > 0$ or $x_3 > 0$), it compresses speculator margins and induces speculators to scale back their positions. Because speculators supply the liquidity that absorbs the fund's period-2 demand x_2, reduced speculative participation raises the fund's implementation-date execution cost. Committing to the implementation-date rebalancing schedule $(x_1, x_3) = (0, 0)$ therefore maximizes speculative liquidity at implementation, thus reducing implementation shortfall.

The condition $\rho_N > 0$ (and correspondingly $\rho_O > 0$ on the O-side) is necessary for the crowding-out derivatives in Parts 2–4 to be strictly negative and is maintained throughout the analysis.

4.2.2 The Role of Speculators

Before characterizing the equilibrium, we first establish that speculators can be beneficial for the fund by reducing implementation shortfall.

Proposition 2 (Speculators reduce fund costs)**.** *Consider a fund that wants to implement an implementation-date rebalancing strategy ($x_1^* = x_3^* = 0$).*

1. *Without speculators, the fund's implementation shortfall is*

$$\mathrm{IS}^F_{\text{no-spec}} = (\mu_N + \mu_O)\, Q^2.$$

2. *With speculators, the fund's implementation shortfall is*

$$\begin{aligned}\mathrm{IS}^F_{\text{spec}} &= (\mu_N + \mu_O)\, Q^2 \\ &\quad - \big[\rho_N\, S_1(L) + \mu_N\, R(L)\big]\, Q + \big[\rho_O\, S_{O,1}(L) + \mu_O\, R_O(L)\big]\, Q.\end{aligned}$$

3. *In both equilibria with full or no commitment, speculators reduce the fund's cost:*

$$\begin{aligned}\mathrm{IS}^F_{\text{spec}} - \mathrm{IS}^F_{\text{no-spec}} &= -\big[\rho_N\, S_1(L) + \mu_N\, R(L)\big]\, Q \\ &\quad + \big[\rho_O\, S_{O,1}(L) + \mu_O\, R_O(L)\big]\, Q \\ &< 0\end{aligned}$$

whenever $\rho_N > 0$ and $\rho_O > 0$.

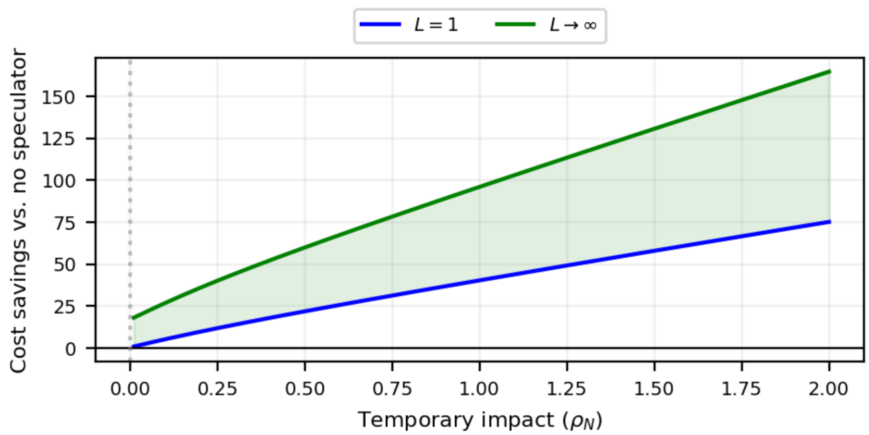
Figure 2 illustrates these cost comparisons as a function of the temporary-impact parameter ρ_N, plotting savings relative to the no-speculator benchmark for $L = 1$ and the competitive limit ($L \to \infty$), along with the incremental benefit of competition.

This result distinguishes our setting from predatory trading as analyzed by Brunnermeier and Pedersen (2005), where a strategic trader exploits forced liquidation. Here, the fund's demand is predictable rather than forced, speculators trade in the opposite direction in period 2, and their participation reduces rather than increases the fund's execution cost. Anticipatory trading in the index-reconstitution context is therefore a form of liquidity provision, not predation.

4.2.3 Equilibrium with No Commitment

We formally define an equilibrium with no commitment as follows.

Figure 2: Cost savings relative to the no-speculator benchmark as a function of ρ_N.



Notes. The gap (shaded area) is the additional benefit from competition. The horizontal axis varies ρ_N; ρ_O is held fixed. Parameters: $Q = 10$, $\lambda_N = 0.5$, $\lambda_O = 0.2$, $\rho_O = 0.2$, $c = 0.3$.

Definition 1 (Equilibrium with no commitment)**.** *An* equilibrium with no commitment *is a strategy profile* $(x_1^*, x_2^*, x_3^*; (s_{1,l}^*, r_l^*)_{l=1}^L)$ *such that: (i) each speculator's strategy* $(s_{1,l}^*, r_l^*)$ *is a best response to the fund's strategy and the other speculators' strategies; and (ii) the fund's strategy* (x_1^*, x_2^*, x_3^*) *is a best response to the aggregate speculator positions, taking speculator behavior as given. Neither party commits ex ante; strategies are chosen simultaneously.*

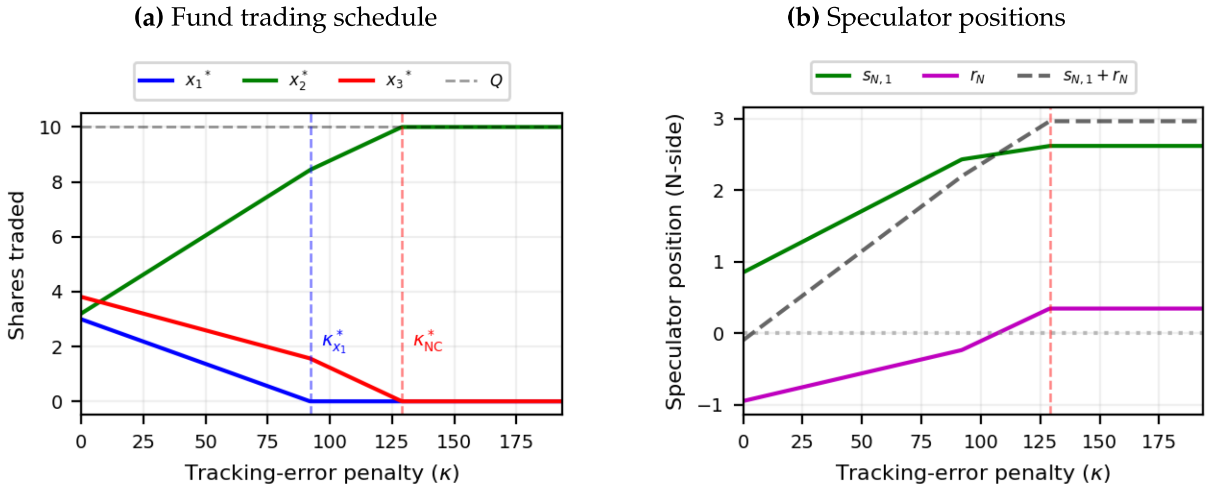
That is, with no commitment, the fund takes speculator positions as given when choosing its trading schedule. The equilibrium is determined by the mutual best responses of all players simultaneously. The fund's optimal schedule is therefore a fixed point of the best-response functions, and the speculators' equilibrium positions are determined by the fund's schedule through the best-response functions derived in Proposition 1. In this case, the equilibrium is characterized by the following result.

Proposition 3 (Fund's optimal rebalancing schedule)**.** *There exist a no-commitment threshold* $\kappa_{\mathrm{NC}}^* > 0$ *(see* (IA.14)*) and a second threshold* $\kappa_{x_1}^* \in (0, \kappa_{\mathrm{NC}}^*)$*, both derived explicitly in the proof, such that the equilibrium has three regimes:*

1. $\kappa \geq \kappa_{\mathrm{NC}}^*$ ***(implementation-date rebalancing):*** *The fund's unique optimal strategy is* $(x_1^*, x_3^*) = (0, 0)$*, and hence,* $x_2^* = Q$*. That is, the fund trades its entire demand in period* 2 *and incurs exactly zero tracking error.*

2. $\kappa_{x_1}^* \leq \kappa < \kappa_{\mathrm{NC}}^*$ ***(intermediate regime):*** $x_1^* = 0$ *and* $x_3^* > 0$.

Figure 3: No-commitment equilibrium as a function of the tracking-error penalty κ.



Notes. Panel 3(a) shows fund trade sizes x_1^*, x_2^*, and x_3^*. For $\kappa \geq \kappa^*_{\mathrm{NC}}$ (red dashed line) implementation-date rebalancing obtains ($x_2^* = Q$). In the intermediate regime $\kappa^*_{x_1} \leq \kappa < \kappa^*_{\mathrm{NC}}$ (blue dashed line), $x_1^* = 0$ and $x_3^* > 0$; below $\kappa^*_{x_1}$ the solution is fully interior. Panel 3(b) shows N-side speculator positions at the no-commitment equilibrium— accumulation $s_{N,1}$ and close-out trade r_N—together with the total period-2 sell $s_{N,1} + r_N$. As κ rises, the fund concentrates trading in period 2, increasing speculators' front-running incentives. Parameters: $Q = 10$, $\lambda_N = 0.5$, $\rho_N = 0.55$, $\lambda_O = 0.2$, $\rho_O = 0.2$, $c = 0.3$, $\nu = 0.15$, $L = 1$.

3. $\kappa < \kappa^*_{x_1}$ ***(interior regime):*** *$x_1^* > 0$ and $x_3^* > 0$. That is, the fund spreads trades across all three periods, with $x_3^* > x_1^*$.*

Figure 3 illustrates the three-regime structure as a function of κ: the left panel traces equilibrium fund trade sizes (x_1^*, x_2^*, x_3^*) across the three regimes, and the right panel shows the corresponding equilibrium speculator positions on the N-side.

The optimality of implementation-date rebalancing admits a clean marginal-cost interpretation. Starting from $(x_1, x_3) = (0, 0)$, the fund's marginal cost of shifting one share from implementation to period 3 is

$$G_3 \;\equiv\; \frac{\partial C^F}{\partial x_3}\bigg|_{(0,0)} \;=\; \underbrace{\kappa\,\frac{\partial\,\mathrm{TE}}{\partial x_3}\bigg|_{(0,0)}}_{\kappa\,\nu} \;+\; \underbrace{\frac{\partial\,\mathrm{IS}^F}{\partial x_3}\bigg|_{(0,0)}}_{-\iota(L)} \;=\; \kappa\,\nu - \iota(L), \tag{18}$$

where $\kappa\,\nu$ is the per-share tracking-error penalty and $\iota(L)$ is the marginal reduction in implementation shortfall from thinning the period-2 block by one share, evaluated at the implementation-date rebalancing positions. Implementation-date rebalancing is optimal when $G_3 \geq 0$, that is, when the tracking-error penalty outweighs the implementation-shortfall saving: $\kappa \geq \iota(L)/\nu = \kappa^*_{\mathrm{NC}}$. Speculators affect the threshold by compressing $\iota(L)$: their period-2 sell flow lowers the fund's implementation-date execution cost, reducing the implementation-shortfall benefit of thinning the block further. Because the fund takes speculator positions as given, it cannot internalize the liquidity it would destroy by pre- or post-trading; the threshold therefore remains strictly positive.

The three-regime structure and back-loading pattern $x_3^* \geq x_1^*$ follow from a simple asymmetry: the implementation-shortfall saving from shifting one share to period 3 is strictly larger than from shifting to period 1. A period-1 purchase permanently displaces the N-security price upward *before* the implementation block, raising the fund's execution cost on all x_2 shares still to be bought at period 2. A period-3 purchase arrives *after* the block, leaving the permanent price at implementation unaffected. The tracking-error penalty $\kappa\,\nu$ is identical for both deviations, so the net marginal cost of a period-1 shift always exceeds that of a period-3 shift:

$$G_1 \;\equiv\; \frac{\partial C^F}{\partial x_1}\bigg|_{(0,0)} \;=\; G_3 + A, \qquad A \;\equiv\; \mu_N\, S_1 - \rho_N R - \mu_O\, S_{O,1} + \rho_O R_O \;>\; 0. \tag{19}$$

The quantity A reflects precisely how speculator positions scale the displacement cost. The dominant term $\mu_N\, S_1 > 0$ captures the fact that the fund's marginal period-1 share lands alongside the S_1 pre-positioning shares already in the market: the fund pays an execution price proportional to the entire flow $(x_1 + S_1)$, so even an infinitesimal x_1 incurs an additional cost of $\mu_N S_1$ relative to a period-3 deviation, where no speculator flow is co-directional with the fund at that margin. The term $-\rho_N R$ provides a partial offset: shifting to period 1 instead of period 3 means the fund avoids co-trading with the speculator covering flow R at period 3, slightly reducing that period's execution cost; however, this saving is smaller than the displacement cost $\mu_N S_1$, so $A > 0$ throughout. Consequently, as κ falls below κ^*_{NC}, G_3 turns negative first: the fund shifts trades to period 3 before period 1. Only when κ falls further below $\kappa^*_{x_1}$ does G_1 also turn negative and an interior solution with $x_1^* > 0$ emerge; throughout that region, $x_3^* > x_1^*$.

4.2.4 Equilibrium with Full Commitment

Full commitment is the more natural assumption for large fund managers, and a repeated-game perspective shows why it can be sustained endogenously. Index reconstitutions recur on a predictable calendar, so the fund and speculators interact repeatedly: speculators observe the fund's realized execution pattern after each event and update their pre-positioning strategies accordingly. In such a setting, the implementation-date rebalancing schedule $(x_1, x_3) = (0, 0)$ can be self-enforcing as a reputation equilibrium: if the fund deviates by trading outside the implementation window, speculators interpret this as a signal of opportunistic behavior, scale back their pre-positioning in subsequent reconstitutions, and thereby raise the fund's future execution costs. Under standard folk-theorem arguments, a sufficiently patient fund therefore finds it optimal to maintain the committed schedule in every period, rendering commitment credible without any external enforcement mechanism. We therefore define the full-commitment equilibrium as follows.

Definition 2 (Equilibrium with full commitment)**.** *An* equilibrium with full commitment *is a strategy profile* $(x_1^*, x_2^*, x_3^*; (s_{1,l}^*, r_l^*)_{l=1}^L)$ *such that: (i) the fund publicly commits to its trading schedule* (x_1^*, x_2^*, x_3^*) *to minimize total cost, anticipating the speculators' best responses; and (ii) each speculator's strategy* $(s_{1,l}^*, r_l^*)$ *is a best response to the committed schedule and the other speculators' strategies. The fund is bound by its stated schedule.*

That is, with full commitment, the fund moves first: it selects and publicly commits to a cost-minimizing schedule (x_1^*, x_2^*, x_3^*), taking into account how speculators will respond to that commitment. Each speculator then observes the complete schedule and chooses trades to maximize profit, competing in Cournot fashion with the remaining $L-1$ speculators. This notion of equilibrium differs from no commitment in that the fund internalizes the effect of its committed schedule on speculator behavior— namely, the crowding-out effect described in Proposition 1—rather than treating speculator positions as exogenous. The fund's optimal commitment schedule is characterized by the following result.

Proposition 4 (Fund's optimal rebalancing schedule under full commitment)**.** *There exists a full-commitment threshold* κ_{FC}^* *satisfying* $0 < \kappa_{\mathrm{FC}}^* \leq \kappa_{\mathrm{NC}}^*$*, with strict inequality under the parameter conditions of Remark 4.2.1. The following hold.*

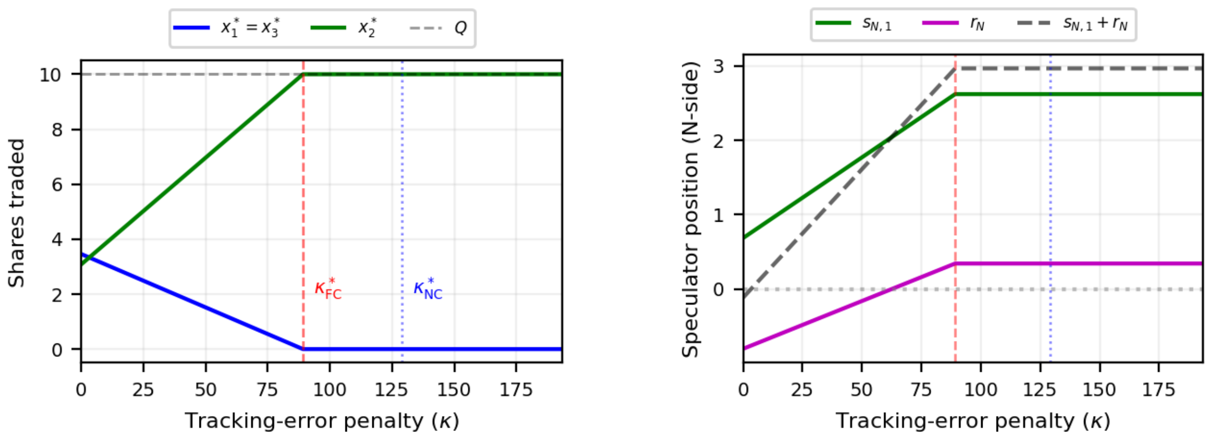
1. $\kappa \geq \kappa_{\mathrm{FC}}^*$ ***(implementation-date rebalancing):*** *The fund's unique optimal commitment is* $(x_1^*, x_3^*) = (0, 0)$*, and hence* $x_2^* = Q$*. The fund trades its entire demand in period* 2 *and incurs exactly zero tracking error. The implementation-day schedule also maximizes aggregate speculator liquidity provision:* $\Sigma_N^* = S_1^* + R^*$ *and* $\Sigma_O^* = |S_{O,1}^*| + |R_O^*|$ *are both strictly decreasing in* x_1 *and* x_3*.*

2. $\kappa < \kappa_{\mathrm{FC}}^*$***:*** *The fund's unique optimal commitment satisfies* $x_1^* = x_3^* > 0$*: trades are spread* equally *across periods* 1 *and* 3*, with the remainder concentrated in period* 2*. Speculator positions on both sides are strictly reduced relative to the implementation-day benchmark.*

Figure 4 illustrates these features: the left panel shows the fund's optimal committed schedule, with the interior solution satisfying $x_1^* = x_3^*$ and the switch to implementation-date rebalancing at κ_{FC}^* (strictly below the no-commitment threshold κ_{NC}^*); the right panel

Figure 4: Full-commitment equilibrium as a function of the tracking-error penalty κ.

(a) Fund trading schedule **(b)** Speculator positions



Notes. Panel 4(a) shows the fund's optimal trading schedule. By Proposition 4, the interior solution satisfies $x_1^* = x_3^*$ (symmetric around implementation); implementation-date rebalancing ($x_1^* = x_3^* = 0$) obtains at the full-commitment threshold κ^*_{FC} (red dashed), strictly below the no-commitment threshold κ^*_{NC} (blue dotted). Panel 4(b) shows N-side speculator positions at the full-commitment equilibrium. When the fund spreads trading ($x_1^* = x_3^* > 0$), speculators reduce their accumulation $s_{N,1}$ and covering r_N, reflecting the strategic-effect terms SE_N and SE_O in eq. (IA.53). Parameters: $Q = 10$, $\lambda_N = 0.5$, $\rho_N = 0.55$, $\lambda_O = 0.2$, $\rho_O = 0.2$, $c = 0.3$, $\nu = 0.15$, $L = 1$.

shows the corresponding reduction in speculator positions relative to the no-commitment benchmark.

The threshold κ^*_{FC} admits a marginal-cost decomposition analogous to (18). Starting from $(x_1, x_3) = (0, 0)$, the committed fund's total marginal cost of shifting one share to period 3 is

$$\tilde{G}_3 \equiv \frac{d\tilde{C}^F}{dx_3}\bigg|_{(0,0)} = \underbrace{\kappa \frac{d\,\mathrm{TE}}{dx_3}\bigg|_{(0,0)}}_{\kappa\,\nu} + \underbrace{\frac{\partial\,\mathrm{IS}^F}{\partial x_3}\bigg|_{(0,0)}}_{-\iota(L)} + \underbrace{\frac{\partial\,\mathrm{IS}^F}{\partial \mathbf{S}^*} \cdot \frac{d\mathbf{S}^*}{dx_3}\bigg|_{(0,0)}}_{\mathrm{SE}(L)\,>\,0} \tag{20}$$

where the first two terms are identical to the no-commitment gradient (18) and the third term is the indirect effect on IS^F through the speculator best responses $\mathbf{S}^*(x_1, x_3) = (S_1^*, R^*, S^*_{O,1}, R^*_O)$, which equals $\mathrm{SE}(L) > 0$ (defined in the proof). Implementation-date rebalancing is optimal when $\tilde{G}_3 \geq 0$, i.e. when $\kappa \geq (\iota - \mathrm{SE}(L))/\nu = \kappa^*_{\mathrm{FC}}$.

Two features of the full-commitment solution deserve comment. First, the threshold κ^*_{FC} is strictly below its no-commitment counterpart κ^*_{NC}. The gap arises because the committed fund internalizes the strategic effect $\mathrm{SE}(L)$: by announcing $(x_1, x_3) = (0, 0)$, it preserves speculator margins and thereby secures the maximum speculative liquidity at implementation. An uncommitted fund ignores this externality (it treats speculator positions as fixed) and therefore perceives a smaller benefit from concentrating all trades at period 2. The result is that the no-commitment fund requires a higher tracking-error penalty to be willing to bear the full period-2 block, whereas the committed fund internalizes the liquidity it would destroy by deviating, and so switches to implementation-date rebalancing under a strictly broader parameter range.

Second, below the full-commitment threshold, the optimal schedule satisfies $x_1^* = x_3^*$: pre- and post-implementation trades are equal in size. This symmetry reflects the structure of the implementation shortfall under commitment. When the fund accounts for the speculator best-response functions (14)–(15), the crowding-out effect is symmetric between period 1 and period 3: announcing one share of pre-implementation trading compresses speculator accumulation by the same amount that announcing one share of post-implementation trading compresses speculator covering. Because the crowding-out derivatives enter the fund's effective cost symmetrically, and the tracking-error penalty $\kappa\,\nu$ is identical for deviations in either direction, the cost-minimizing committed schedule equates x_1 and x_3. This contrasts sharply with the no-commitment case, where the asymmetry in permanent price impact (period-1 purchases displace prices ahead of the implementation block, whereas period-3 purchases do not) generates back-loading ($x_3^* > x_1^*$) at every interior solution.

4.3 Equilibrium in the Competitive Limit

As the number of speculators grows, a sharp distinction emerges between the two equilibrium concepts. Under no commitment, the threshold $\kappa^*_{\mathrm{NC}}(L)$ remains bounded away from zero even as $L \to \infty$: even with a fully competitive speculator sector, the fund requires a strictly positive tracking-error penalty to prefer implementation-date rebalancing. Under full commitment, by contrast, $\kappa^*_{\mathrm{FC}}(L) \to 0$: once speculators are sufficiently numerous and the fund can credibly commit to its schedule, *any* fund—however small its tracking-error penalty—will optimally concentrate all trading at implementation. Commitment therefore does not merely lower the threshold; in the competitive limit it eliminates it entirely.

Proposition 5 (Competitive limit: $L \to \infty$)**.** *As the number of identical speculators $L \to \infty$:*

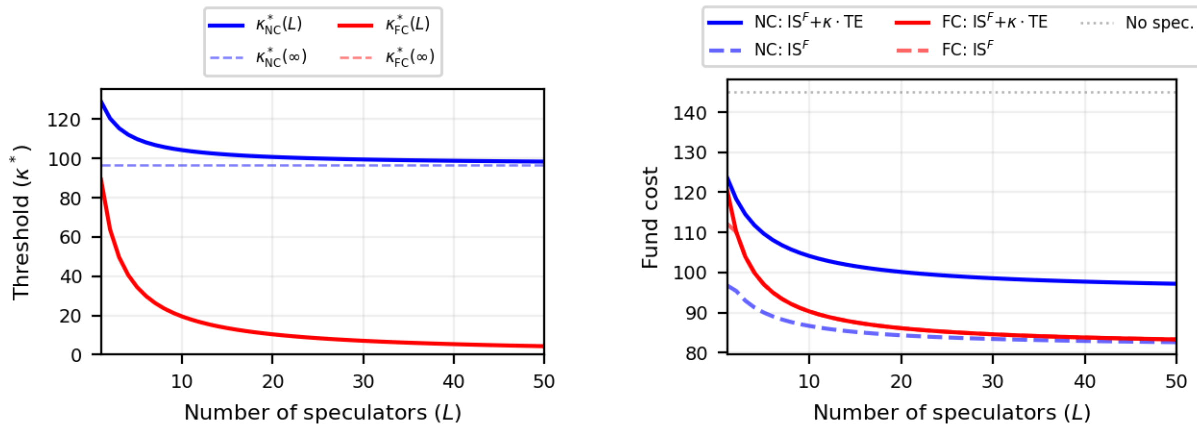
1. ***No-commitment threshold:*** *$\kappa^*_{\mathrm{NC}}(L)$ converges to a strictly positive limit:*

$$\kappa^*_{\mathrm{NC}}(\infty) = \frac{(\mu_N + \mu_O)\,Q}{\nu} > 0. \tag{21}$$

   When $c = 0$, $\kappa^*_{\mathrm{NC}}(L)$ is strictly decreasing in L.

Figure 5: Effect of the number of speculators L on equilibrium outcomes.

(a) Implementation-date rebalancing thresholds **(b)** Equilibrium fund cost



Notes. Panel 5(a) shows implementation-date rebalancing thresholds $\kappa^*_{\mathrm{NC}}(L)$ (blue) and $\kappa^*_{\mathrm{FC}}(L)$ (red) as functions of L, with dashed lines showing their limits as $L \to \infty$. Both thresholds decrease in L; the no-commitment limit is strictly positive ($\kappa^*_{\mathrm{NC}}(\infty) = (\mu_N + \mu_O)Q/\nu$) while the full-commitment threshold converges to zero (Proposition 5). Panel 5(b) shows fund equilibrium cost at fixed $\kappa = 80$. Solid lines show total cost IS$^F + \kappa \cdot \mathrm{TE}$; dashed lines show IS^F alone. The FC total cost (solid red) is always below the NC total cost (solid blue), confirming that commitment benefits the fund. When $\kappa^*_{\mathrm{FC}}(L)$ drops below 80 for large L, the FC fund switches to implementation-date rebalancing (zero TE, higher IS^F), whereas the NC fund continues to split trades. Parameters: $Q = 10$, $\lambda_N = 0.5$, $\rho_N = 0.55$, $\lambda_O = 0.2$, $\rho_O = 0.2$, $c = 0.3$, $\nu = 0.15$.

2. ***Full-commitment threshold:*** *$\kappa^*_{\mathrm{FC}}(L)$ is strictly decreasing in L and $\kappa^*_{\mathrm{FC}}(L) \to 0$ as $L \to \infty$. Consequently, for any $\kappa > 0$ there exists $L^*(\kappa)$ such that for all $L \geq L^*(\kappa)$, implementation-date rebalancing ($x^*_1 {=} x^*_3 {=} 0$) is the unique full-commitment equilibrium.*

Figure 5 illustrates both results: the left panel plots $\kappa^*_{\mathrm{NC}}(L)$ (blue) and $\kappa^*_{\mathrm{FC}}(L)$ (red) as functions of L, with dashed horizontal asymptotes; the right panel shows equilibrium fund cost at a fixed κ, confirming that the committed fund always achieves lower total cost and switches to implementation-date rebalancing at a smaller L.

A competitive speculator who pre-positions s shares reduces the fund's period-2 net flow by s and earns a temporary-impact premium of $\rho_N\, s$—but this exactly offsets the $\rho_N\, s$ impact cost incurred on the pre-positioning trade itself. The fund, by contrast, internalizes the impact saving on its *entire* residual order, a benefit proportional to Q, not to the marginal share moved. Competition drives speculator rents to zero but cannot eliminate this infra-marginal component of the fund's cost, so the NC threshold remains strictly positive regardless of how many speculators enter.

The full-commitment result has a sharper implication. Because the full-commitment threshold $\kappa^*_{\mathrm{FC}}(L) \to 0$, any fund operating in a sufficiently competitive market and able to credibly pre-commit to its schedule will, in equilibrium, concentrate *all* rebalancing at the implementation date—even if its tracking-error penalty is arbitrarily small. This provides a direct explanation for the empirical regularity that passive funds almost universally execute at implementation rather than spreading trades across the announcement window: credible commitment, backed by reputation and disclosure requirements, aligns the fund's incentives with the strategy that minimizes total cost, and competition among liquidity providers supports the optimality of the strategy. Crucially, this outcome is sustained by speculator participation, not undermined by it. At implementation, speculators act as endogenous liquidity providers, absorbing the fund's concentrated demand and thereby reducing the execution cost below what the fund would face trading alone. The fund's commitment to implementation-date rebalancing maximizes speculator margins, encouraging entry and competition; more competition in turn drives the threshold toward zero, making implementation-date rebalancing optimal for a wider set of funds. Commitment and speculator competition are therefore complementary forces: each reinforces the other, and together they sustain the concentrated-rebalancing equilibrium observed in practice.

4.4 PRICE AND QUANTITY-BASED MEASURE OF SPECULATIVE COMPETITION

All structural quantities derived above depend on parameters that are not directly observable. We now show that speculative competition leaves a clean signature in transaction prices alone. Define the *post-implementation reversal* as

$$\Gamma(L) \equiv \frac{\mathbb{E}\left[\bar{P}_N - P_{N,2}\right]}{P_{N,0}\, Q}, \tag{22}$$

the gap between the post-trading price $\bar{P}_N = \bar{v} + \lambda_N Q$ and the expected implementation-date transaction price, scaled by the initial N-security transaction price $P_{N,0}$ and the rebalancing quantity Q; taking expectations removes the mean-zero martingale innovations, isolating price-impact components.

Because all positions are closed by period 3, cumulative N-side net order flow satisfies $Z_{N,3} = Q$ regardless of L, so the post-trading price $\bar{P}_N = \bar{v} + \lambda_N Q$ is pinned by fundamentals and the rebalancing size, independent of speculator competition. Using the execution-price formula (4) with $Z_{N,1} = S_1(L)$ under implementation-date rebalancing:

$$\Gamma(L) = \frac{\mu_N\, R(L) - \rho_N(Q - S_1(L))}{P_{N,0}\, Q}, \tag{23}$$

where $\alpha \equiv \rho_N > 0$. The second term at the numerator captures the temporary concession (α) on the residual block not pre-positioned by speculators, which depresses the implementation-date price below $\bar{P}_N$; the first term reflects the price impact of period-3 covering flow that has not yet been priced in at period 2 and partially offsets the concession.

Proposition 6 (Post-implementation reversal)**.** *The post-implementation reversal $\Gamma(L)$ defined in* (22) *satisfies:*

1. ***Strict monotonicity.*** *$\Gamma(L) < 0$ and Γ is strictly increasing in L for all $L \geq 1$ and all parameters satisfying $\rho_N > 0$ and $c \geq 0$.*

2. ***Competitive limit.*** *$\Gamma(L)$ converges to a strictly negative limit,*

$$\Gamma(\infty) \;\equiv\; \lim_{L\to\infty} \Gamma(L) \;=\; -\frac{\rho_N^3}{P_{N,0}\big(3(\rho_N + \frac{\lambda_N}{2})^2 + \frac{\lambda_N^2}{4}\big)} \;<\; 0, \tag{24}$$

where $\Gamma(\infty) \to 0$ as $\rho_N \to 0$.

We complement the price-based measure Γ with a *volume-based* measure of speculative competition. Define the *intermediated volume*

$$\mathcal{V}(L) \;\equiv\; \frac{S_1(L) + R(L)}{Q}\,, \tag{25}$$

the fraction of the fund's rebalancing demand that is absorbed by speculators rather than by the market maker at implementation. Under implementation-date rebalancing, the period-2 net order flow is $z_{N,2} = Q - S_1 - R$, so $\mathcal{V}$ measures how much of Q is intermediated.

Proposition 7 (Intermediated volume)**.** *Under $\rho_N > 0$ and $c \geq 0$, the intermediated volume $\mathcal{V}(L)$ satisfies:*

1. ***Strict monotonicity.*** *$\mathcal{V}(L)$ is strictly increasing in L for all $L \geq 1$.*

2. ***Competitive limit.*** *$\mathcal{V}(L)$ converges to*

$$\mathcal{V}(\infty) \;\equiv\; \lim_{L\to\infty} \mathcal{V}(L) \;=\; \frac{2(\rho_N + \frac{\lambda_N}{2})^2 + \frac{\lambda_N^2}{2}}{3(\rho_N + \frac{\lambda_N}{2})^2 + \frac{\lambda_N^2}{4}} \;\in\; (0,1), \tag{26}$$

where the displayed ratio converges to 1 as $\rho_N \to 0$.

While Γ captures the price signature of competition, $\mathcal{V}$ captures its quantity signature: as more speculators enter, a larger share of the fund's demand is pre-positioned and covered by intermediaries, reducing the residual order flow that hits the market maker at implementation. The strict upper bound $\mathcal{V}(\infty) < 1$ reflects the same economic force that sustains $\Gamma(\infty) < 0$: temporary price impact creates a wedge that prevents full intermediation even under perfect competition.

5 Implications: Fast-Track Inclusions and Liquidity Complementarity

In this section, we apply the theoretical framework of Section 4 to evaluate the welfare implications of fast-track index additions after an initial public offering (IPO). We also study the strategic behavior of investors with weak tracking-error constraints.

The analytical results in Section 4 are derived for a three-trading-period model. In practice, the interval between the index-reconstitution announcement and the effective date can span many trading sessions, allowing both the fund and speculators to spread their activity over more periodsj, both before and after implementation. Therefore, we start this section by introducing a generalized model with an arbitrary number of trading periods. We solve the model numerically and derive the empirical implications. The full model specification and solution method are described in Internet Appendix IA.6. We then use the generalized model to study fast-track additions and weak index trackers.

5.1 Generalized Setup

In the generalized model, we denote by T the implementation date. There are $T-1$ pre-implementation periods and H post-implementation periods, for a total of $T+H$ trading periods, all indexed by $t \in \{1, \ldots, T+H\}$. As in the base model, period 0 represents the initial state and a terminal period at $T+H+1$ represents the realization of the fundamental value. Setting $T = 2$ and $H = 1$ recovers the base model exactly.

The fund chooses a schedule $\mathbf{x} = (x_1, \ldots, x_{T+H})$ with $\sum_{t=1}^{T+H} x_t = Q$ and $x_t \geq 0$. Each speculator l chooses trades $(s_{N,1,l}, \ldots, s_{N,T+H,l})$ subject to $\sum_t s_{N,t,l} = 0$ (zero-net-position constraint), and analogously on the O-side. Because speculator profit remains a quadratic function of trades, the symmetric Cournot best response is obtained by solving a $(2(T+H)) \times (2(T+H))$ KKT system; the resulting mapping $\mathbf{s}^*(\mathbf{x})$ is affine in the fund schedule. Substituting into the fund's objective yields a convex quadratic program, solved exactly by an active-set method for both the full-commitment and no-commitment equilibria (see

Appendix IA.6 for details).

5.1.1 Robustness of the Main Results

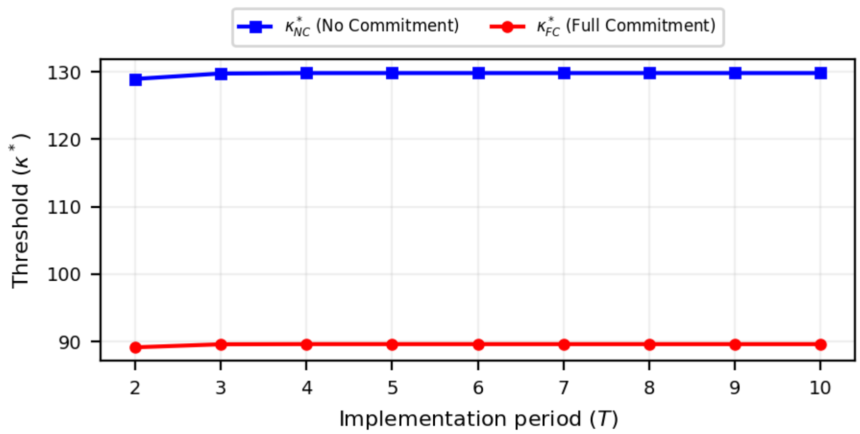
A central regularity in the index reconstitution literature is that abnormal price pressure clusters at or very near the implementation date, even when funds have an extended pre-event window in which to trade. The model replicates and rationalizes this pattern. For any tracking-error penalty above κ^*_{NC}, the fund concentrates its entire demand at period T irrespective of how many trading periods are available: implementation-date rebalancing is optimal even where gradual execution is feasible.

The key is that the marginal benefit from spreading execution diminishes with distance from implementation. Shifting a unit of demand from period T to period $T-1$ reduces the price spike at implementation; extending that shift to period $T-2$ yields a smaller additional saving, because the permanent impact from pre-T speculator accumulation is already partially impounded into prices. Meanwhile, the tracking-error penalty accrues linearly for every period the fund deviates from the index. The binding trade-off therefore always falls at the period immediately adjacent to implementation ($t = T \pm 1$); periods farther from the event contribute negligibly to either cost component. As a consequence, both thresholds κ^*_{NC} and κ^*_{FC} converge as T grows (Figure 6). Figure 7 illustrates for $T = 4$: as κ rises through the threshold, off-implementation trades collapse to zero and the full demand Q shifts discretely to period T, producing a concentrated price spike.

Moreover, the results of Proposition 4 carry directly to the generalized model. A fund that publicly commits to its trading schedule concentrates demand at period T under a wider set of parameters than one that does not. That is, $\kappa^*_{\mathrm{FC}} < \kappa^*_{\mathrm{NC}}$ for all T. Both thresholds converge as T grows, and the gap between them remains visible in Figure 6, confirming that neither result is an artifact of the three-period structure.

Finally, the distribution of trades across periods is informative about whether the fund does or does not have commitment. The uncommitted fund's schedule is back-loaded: execution is skewed toward post-implementation periods for all intermediate values of κ (Figure 7, left panel). This asymmetry reflects the fund's inability to deter pre-T speculator accumulation: knowing that front-running will drive up the execution price at implementation, the uncommitted fund shifts residual demand to post-T periods, when speculator position-covering partially reverses the pre-T price run-up and execution costs are correspondingly lower.

Figure 6: Implementation-date rebalancing thresholds as a function of time to implementation T.



Notes. The no-commitment threshold κ^*_{NC} (blue) and the full-commitment threshold κ^*_{FC} (red) both converge as T grows; the gap between them represents the commitment gain. This convergence reflects the fact that the binding deviation is always at the adjacent period $T \pm 1$. Parameters: $Q = 10$, $L = 1$, $\lambda_N = 0.5$, $\rho_N = 0.55$, $\lambda_O = 0.2$, $\rho_O = 0.2$, $\nu = 0.15$, $c = 0.3$.

5.1.2 Price Run-Up, Trading Volume, and the Role of Speculation

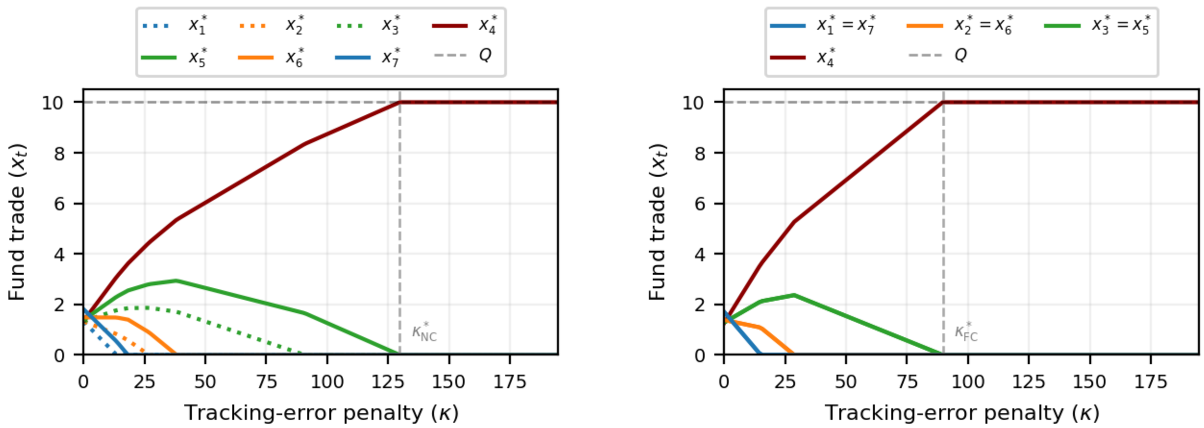
The model's most direct empirical implications concern the shape of the price path and trading volume around implementation. A set of predictions emerges that corresponds to empirical regularities documented in the literature: the secular decline in implementation-day abnormal returns, the migration of price discovery into the pre-event window, and the concentration of index-driven volume in the closing auction on the effective date.

Figure 8 traces the predicted execution-price path for the added security when the fund rebalances entirely at period $T = 30$.

With $L = 1$, speculators accumulate few shares pre-T, keeping pre-implementation prices low. As speculators sell their limited inventory at implementation against the large demand of the index tracker, prices spike at implementation, far above the post-trading benchmark $\bar{P}_N - \bar{v} = \lambda_N Q$, thus creating a large reversal. At $L = 10$, greater competition drives pre-T accumulation higher, thus smoothing the price spike at implementation and subsequent reversal. At $L = 100$, speculators are highly competitive, bid aggressively for stocks before implementation, and, as a result, push prices before and at implementation near their post-trading benchmark $\bar{P}_N - \bar{v} = \lambda_N Q$. Therefore, the model predicts a cross-sectional pattern: securities added to a competitive market should exhibit immediate price changes at the announcement and a smaller implementation-date spike than securities added to uncompetitive markets.

Figure 7: Fund trading schedule as a function of the tracking-error penalty κ in the generalized model.

(a) No commitment **(b)** Full commitment



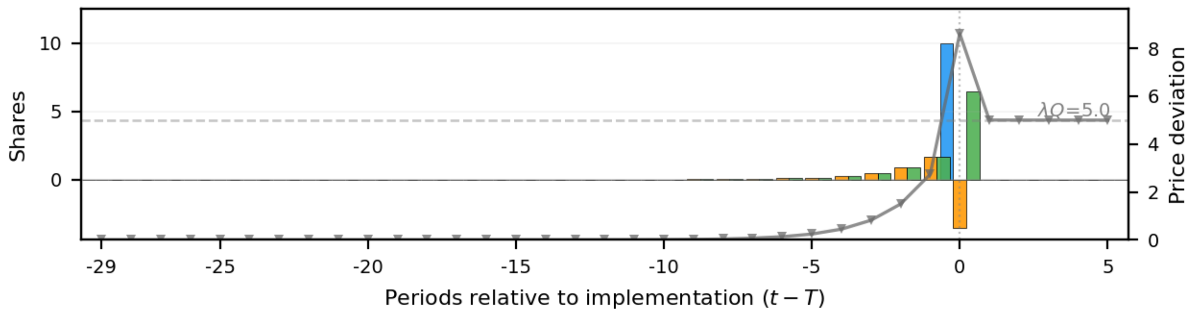
Notes. Each line shows the trade size for one period. Panel 7(a) shows the no-commitment equilibrium, with all periods plotted separately. At the threshold κ^*_{NC} (dashed line) the fund switches to implementation-date rebalancing. Panel 7(b) shows the full-commitment equilibrium. Because the FC solution is symmetric around the implementation date, periods equidistant from period T always trade identical quantities and share a legend entry. The commitment threshold κ^*_{FC} is strictly below κ^*_{NC}. Parameters: $T = 4$, $H = 3$ ($T + H = 7$ trading periods: periods 1–3 pre-implementation, period 4 implementation, periods 5–7 post-implementation), $Q = 10$, $L = 1$, $\lambda_N = 0.5$, $\rho_N = 0.55$, $\lambda_O = 0.2$, $\rho_O = 0.2$, $\nu = 0.15$, $c = 0.3$.

Speculators provide a form of price discovery by impounding the known demand shock into transaction prices as soon as information is available and they can trade on it. The same competition mechanism offers a time-series rationalization for the disappearing index effect documented by Greenwood and Sammon (2025): competition among speculators intensified and price discovery migrated progressively earlier in event time. The model predicts that this process leaves total cumulative price adjustment unchanged, because it is determined by the permanent impact parameter $\lambda_N Q$ alone. However, competitive speculators redistribute this price impact from the implementation date to the pre-event period, precisely the pattern documented by Greenwood and Sammon (2025).

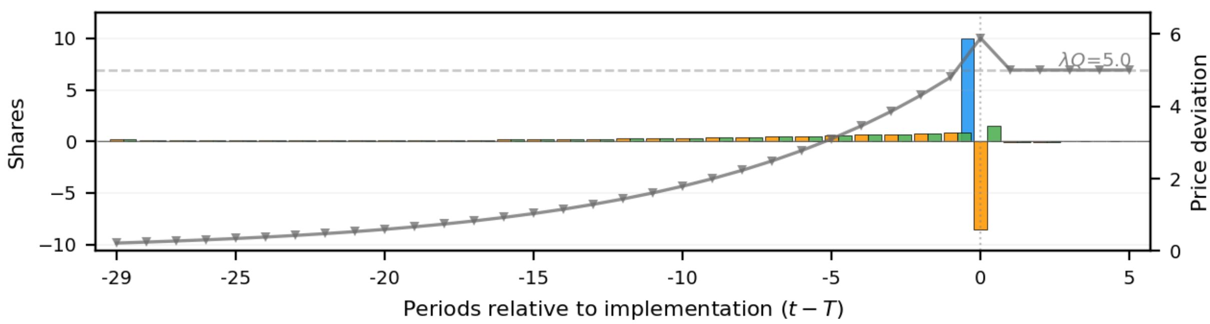
Moreover, these results provide a demand-side rationalization for the disproportionate concentration of index-driven trading in the closing auction on effective dates documented by Chinco and Sammon (2024): the closing auction is the economically optimal execution venue for any benchmarked investor whose tracking-error penalty exceeds the threshold, not merely a focal convention. As the intermediation sector has grown and speculator competition has intensified, κ^*_{NC} converges toward its competitive-limit value (Proposition 5), expanding the set of fund types for which implementation-date rebalancing is individually optimal.

Figure 8: Order flows and transaction price under implementation-date rebalancing.

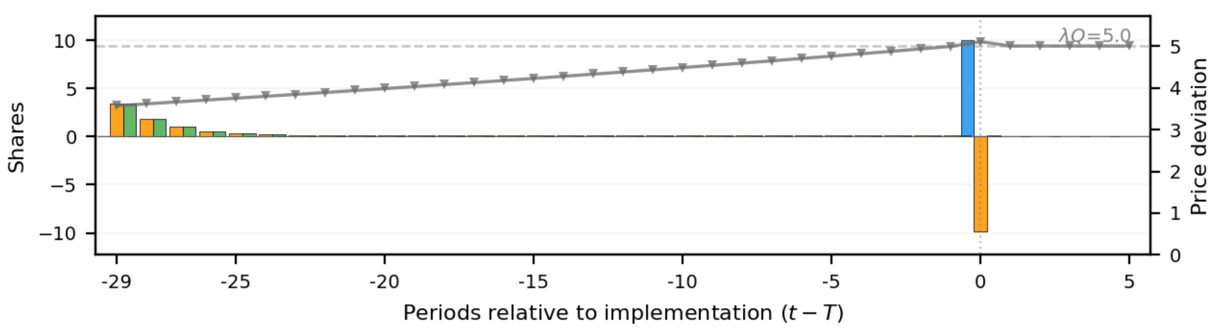
(a) $L = 1$ speculator.



(b) $L = 10$ speculators.



(c) $L = 100$ speculators.



Notes. $x_T^* = Q$, $x_t^* = 0$ for $t \neq T$; periods 1–35 shown (activity negligible thereafter). Bars: fund trade (blue), Speculators (orange), net order flow (green). Gray line: transaction price (right axis). Dashed line: $\lambda_N Q = 5$. Greater speculator competition shifts price pressure from the implementation spike to a pre-T run-up. Parameters: $T = 30$, $Q = 10$, $\lambda_N = 0.5$, $\rho_N = 0.55$, $\lambda_O = 0.2$, $\rho_O = 0.2$, $c = 0.3$.

5.2 FAST-TRACK INCLUSIONS

We now evaluate the cost index trackers incur when a stock is included in the index shortly after its IPO. Newly listed stocks tend to be illiquid (high ρ) and, by fast-tracking their inclusion (lowering T), index providers force speculators to concentrate their inventory accumulation over a short period of time, exacerbating costs for index trackers. Speculators can make up for the short pre-implementation window by shorting the stock at implementation and closing their short positions afterward. However, newly issued stocks are hard to short, as the supply of lendable shares is severely limited.

FAST-TRACK ADDITIONS ARE COSTLY FOR INDEX INVESTORS. We evaluate the cost for index trackers using the model-implied measure of avoidable transaction costs: the post-implementation price reversal Γ (Proposition 6). The left column of Figure 9 shows that, in general, index trackers incur larger price reversals when the stock is illiquid (high ρ) and speculators are not competitive (low L). However, as long as short selling is possible and speculation is sufficiently competitive, speculators can intermediate the demand shock by shorting at implementation and covering afterward, so index trackers incur only a marginally worse reversal when the stock is fast-tracked after its IPO (low T).

In contrast, the right column of Figure 9 shows that when short selling is prohibited, the reversal is much higher for fast-track inclusions, especially when the stock is illiquid. Even in highly competitive markets (high L), index trackers incur substantial reversal costs in illiquid stocks because speculators cannot intermediate the demand shock by shorting at implementation and covering afterward.

FAST-TRACK ADDITIONS INCREASE IPO PRICES. Whereas index trackers suffer from fast-track inclusions of illiquid, hard-to-short stocks, issuers greatly benefit from them. The left column of Figure 10 shows that when short selling is unconstrained, the first-trading-day price is only marginally higher for fast-track inclusions, especially when the stock is illiquid. In contrast, the right column of Figure 10 shows that when short selling is prohibited, the first-trading-day price is much higher for fast-track inclusions, especially when the stock is illiquid. Because speculators are forced to build inventory before implementation with no option to short at implementation, they bid up the price prior to implementation, leading to substantially higher prices on the first trading day for fast-track inclusions of illiquid, hard-to-short stocks.

Interestingly, issuers and initial stock owners benefit more when markets are highly competitive. As discussed earlier, competitive speculators trade aggressively and immediately bid up prices. In the context of an IPO, competitive speculators will therefore bid fiercely at the IPO stage or when the stock begins trading, ultimately transferring value to the issuer and initial stock owners.

Overall, these predictions are consistent with the findings in Murray and Sammon (2026). Our model provides a framework to rationalize and qualify their empirical findings. First, the model highlights that index investors suffer costs from fast-tracked additions only as long as markets are illiquid (high ρ) and short selling is constrained. Both conditions are likely after an IPO. In contrast, the surprise addition of a liquid and established stock to an index could be performed at a low cost to investors even with a short pre-implementation trading period, as trades are spread over the post-implementation period by short selling. Second, issuers and original owners especially benefit from fast-tracking in highly competitive markets. Conversely, in less competitive markets like those in the past, speculators would be able to retain a larger share of total profits.

Figure 11 illustrates the inventory and price dynamics behind these patterns. When short selling is possible, speculators can intermediate the demand shock by shorting at implementation and covering afterward, so index trackers incur only a marginally larger reversal when the stock is fast-tracked after an IPO (low T). When short selling is prohibited, speculators cannot intermediate the demand shock in this way, so index trackers incur substantial reversal costs in illiquid stocks, and issuers can expect high stock prices on their first trading day.

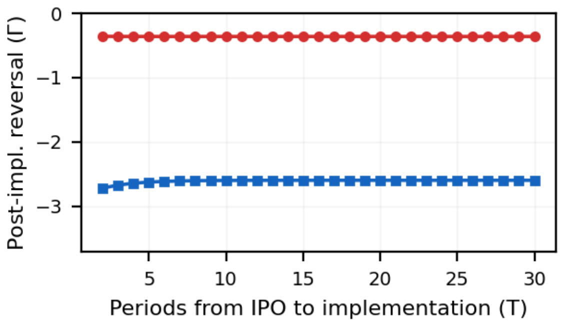
5.3 Trading with Weak and Strict Index Trackers

Many real-world funds track a benchmark but retain more flexibility on their tracking error: active funds with index benchmarks, enhanced index funds, and multi-asset managers that shadow a reconstitution-linked mandate. These *weak trackers* share the same benchmark as the index fund but face a lower tracking-error penalty κ. Chinco and Sammon (2024) show that benchmark-motivated trading, inferred from volume patterns, is substantially larger than what one would estimate from labeled index-fund AUM alone, reflecting the participation of weak-indexers and benchmarked institutions in the same closing-auction equilibrium as strict trackers (Pavlova and Sikorskaya, 2023).

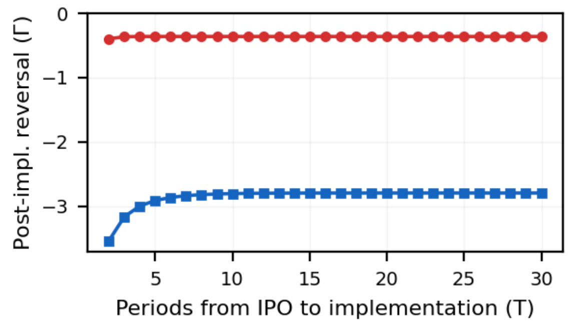
We study two scenarios involving weak trackers. First, as a benchmark, Figure 12 shows the weak tracker managing its schedule in isolation under NC (left column) and FC (right column). Second, Figures 13 and 14 place the same weak tracker alongside a passive fund committed unconditionally to period-T execution.

Figure 9: Post-implementation reversal Γ as a function of the pre-implementation horizon T: baseline versus no short-selling.

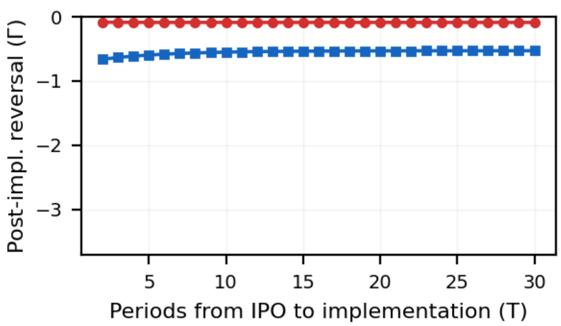
(a) Baseline, $L = 1$.



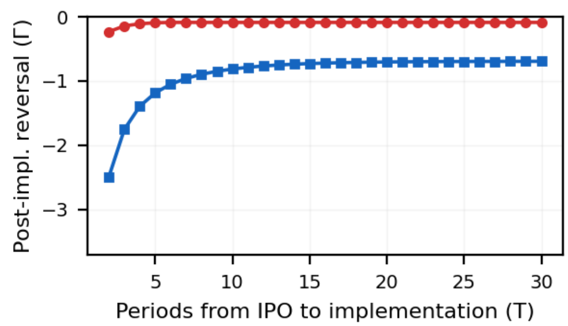
(b) No short-selling, $L = 1$.



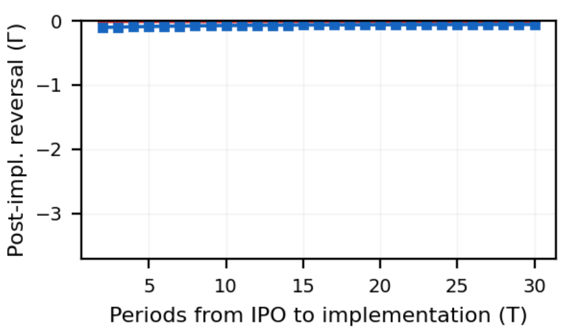
(c) Baseline, $L = 10$.



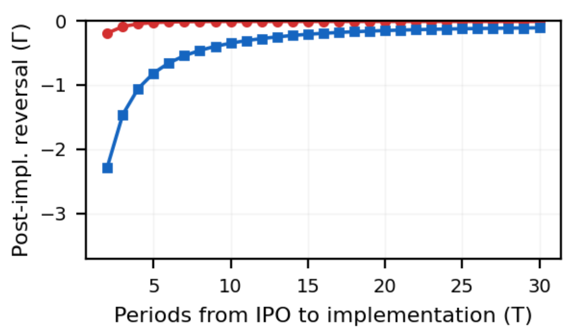
(d) No short-selling, $L = 10$.



(e) Baseline, $L = 100$.

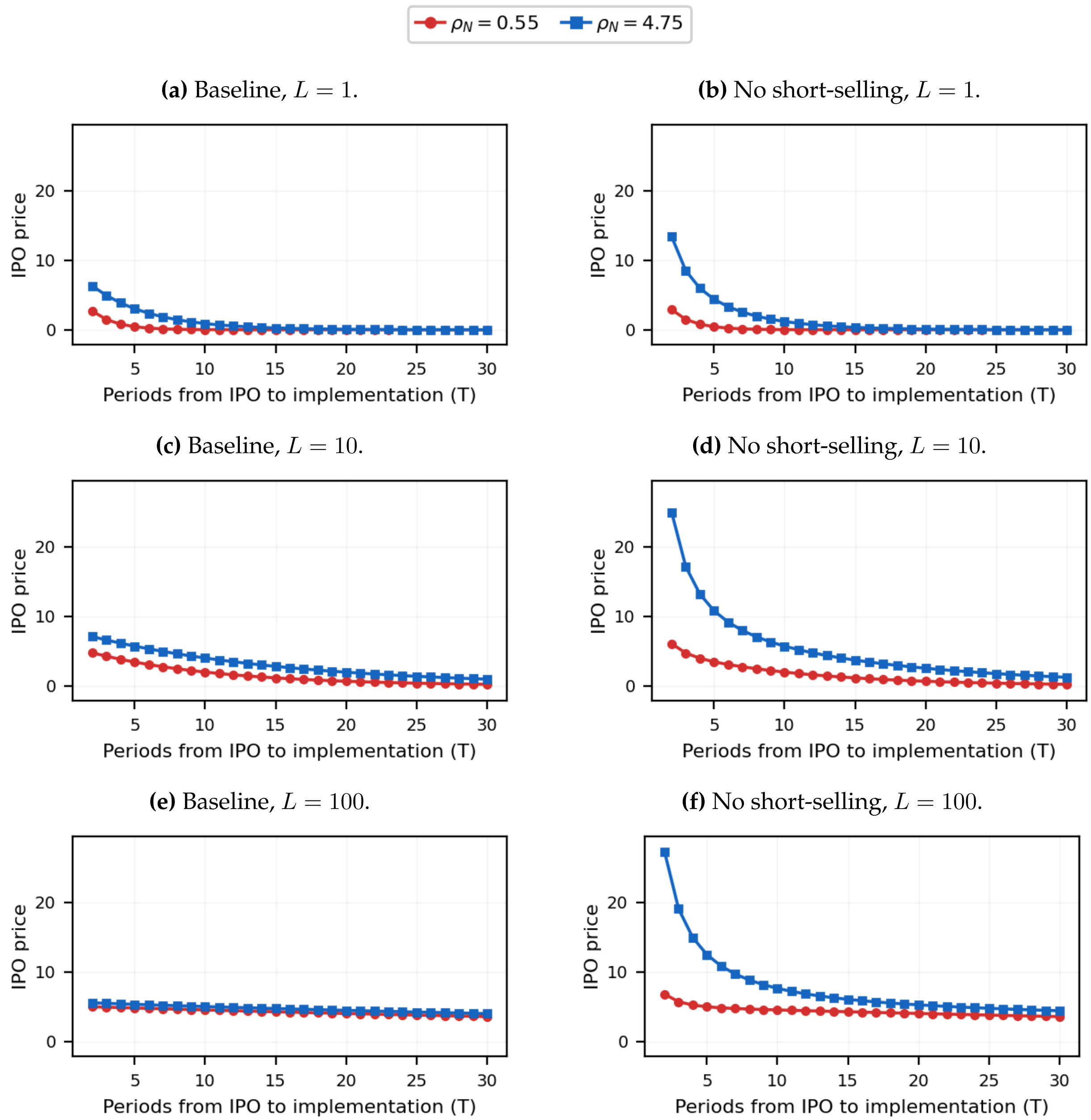


(f) No short-selling, $L = 100$.



Notes. Red and blue lines denote $\rho_N = 0.55$ and $\rho_N = 4.75$, respectively. In the left column, short selling is unrestricted. In the right column, short selling is not allowed. The first row shows outcomes with $L = 1$ speculators. The second row shows outcomes with $L = 10$ speculators. The third row shows outcomes with $L = 100$ speculators. Parameters: $Q = 10$, $\lambda_N = 0.5$, $c = 0.3$, $H = 29$.

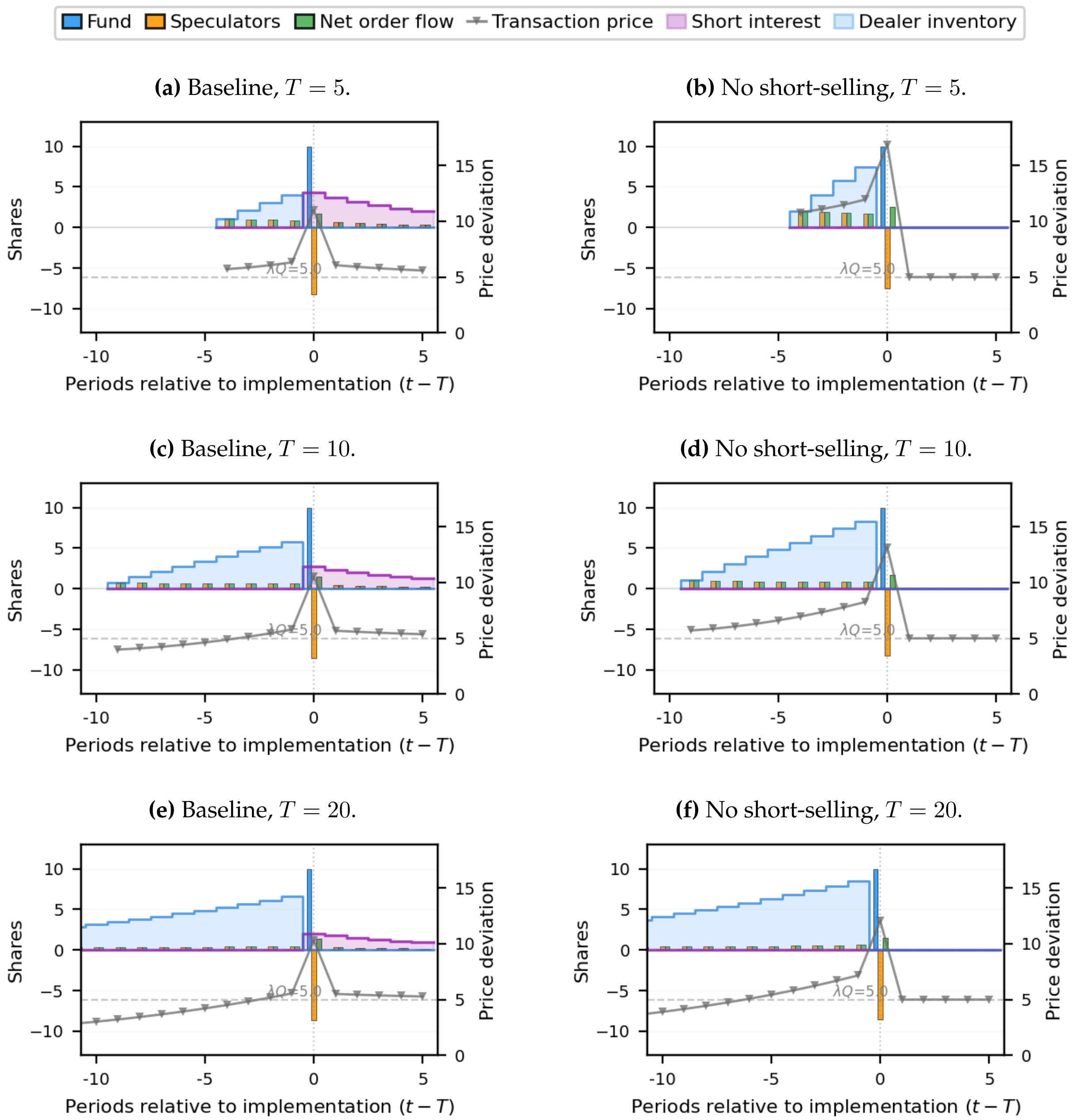
Figure 10: IPO price (first-trading-day transaction price at $t = 1$) as a function of the pre-implementation horizon T: baseline versus no short-selling.



(a) Baseline, $L = 1$. **(b)** No short-selling, $L = 1$. **(c)** Baseline, $L = 10$. **(d)** No short-selling, $L = 10$. **(e)** Baseline, $L = 100$. **(f)** No short-selling, $L = 100$.

Notes. Red and blue lines denote $\rho_N = 0.55$ and $\rho_N = 4.75$, respectively. In the left column, short selling is unrestricted. In the right column, short selling is not allowed. The first row shows outcomes with $L = 1$ speculators. The second row shows outcomes with $L = 10$ speculators. The third row shows outcomes with $L = 100$ speculators. Parameters: $Q = 10$, $\lambda_N = 0.5$, $c = 0.3$, $H = 29$.

Figure 11: Order flows, short interest, and speculators' inventory under high temporary impact ($\rho_N = 4.75$): baseline versus no short-selling.



Notes. Rows set $T = 5$ (first row), $T = 10$ (second row), and $T = 20$ (third row); columns compare unrestricted short selling (left) and constrained short selling (right). Bars: fund trade (blue), aggregate speculator flow (orange), net order flow (green). Purple shading: speculator short interest $\max(-LS_{N,t}, 0)$; light blue shading: speculators' inventory $\max(LS_{N,t}, 0)$. Gray line: transaction price (right axis); dashed line: $\lambda_N Q = 5$. Parameters: $Q = 10$, $\rho_N = 4.75$, $\lambda_N = 0.5$, $c = 0.3$, $L = 10$, $H = 29$.

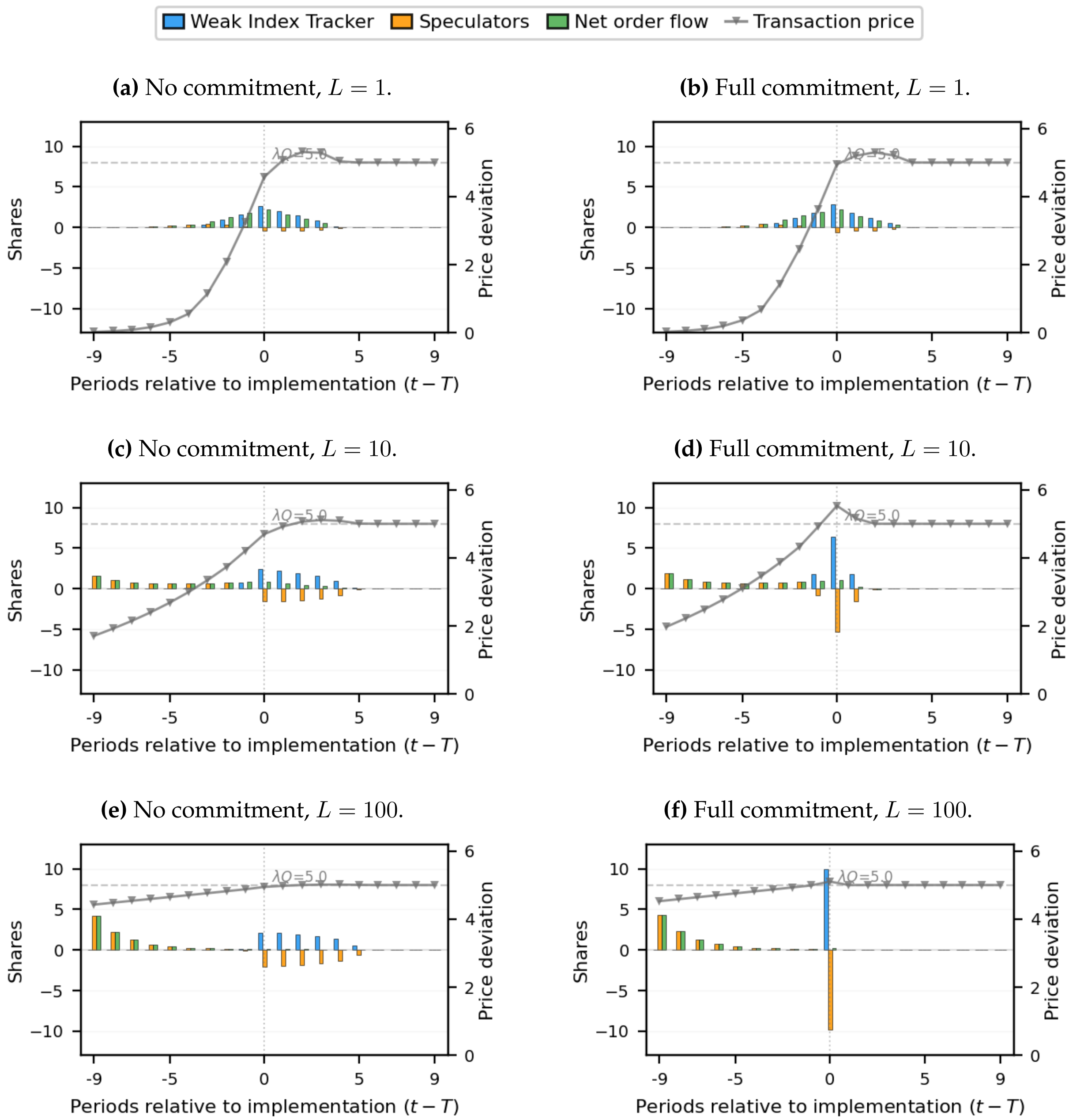
WEAK TRACKERS' OPTIMAL REBALANCING. Figure 12 illustrates how the equilibrium trading schedule of a weak tracker responds to speculator competition under no commitment (left column) and full commitment (right column). Under no commitment, limited competition leaves the fund trading in a relatively balanced way around implementation, but stronger competition makes the schedule increasingly back-loaded. The mechanism is that competitive front-running permanently displaces the N-security price before implementation; the uncommitted fund therefore retreats from the pre-T window and shifts demand to post-T periods, when speculator position-reversal partially unwinds the earlier run-up. Commitment changes this pattern.

As competition increases, speculators intermediate more of the predictable demand before and at T, producing a smoother price path and reducing the need for the fund to chase the post-implementation reversal. The figure therefore distinguishes two channels: without commitment, weak trackers avoid the implementation-date price pressure by back-loading; with commitment, the fund uses the announced schedule to reshape speculative inventory and keep execution closer to the benchmark date. Sammon and Shim (2026) provides direct evidence for the same cost channel: funds that execute rebalancing trades earlier pay lower average implementation shortfall, consistent with the model's prediction that avoiding the most crowded implementation-date pressure can reduce costs. The long-horizon pre-event volume accumulation documented by Chinco and Sammon (2024) corroborates the demand side: predictable reconstitutions attract intermediary positioning over months, implying that the relevant trading window for cost-minimizing funds is far longer than the narrow pre-implementation period.

WEAK TRACKERS AVOID PASSIVE BLOCKS IN THIN MARKETS BUT JOIN THEM IN DEEP MARKETS. Figures 13 and 14 introduce a fully passive counterpart that commits unconditionally to purchasing $\bar{Q}$ shares at period T, fixing total demand at $Q + \bar{Q} = 10$, under a passive-dominant split ($Q = 1$, $\bar{Q} = 9$) and an equal split ($Q = 5$, $\bar{Q} = 5$). The passive block raises implementation-date order flow mechanically, but its effect on the weak tracker's schedule is not monotone in the size of the block alone. A large passive order is also a large predictable intermediation opportunity. Whether the weak tracker trades around it or joins it depends on how much of that predictable demand speculators have already absorbed before T.

Under no commitment (Figure 13), thin speculative competition leaves the passive block only partly intermediated. The implementation-date price spike remains large, especially in the passive-dominant case, so the weak tracker avoids period T and shifts its own trades into nearby off-implementation periods. As L rises, this incentive is weakened

Figure 12: Order flows of the weak index tracker under no commitment and full commitment.



(a) No commitment, $L = 1$.

(b) Full commitment, $L = 1$.

(c) No commitment, $L = 10$.

(d) Full commitment, $L = 10$.

(e) No commitment, $L = 100$.

(f) Full commitment, $L = 100$.

Notes. Bars: weak index tracker trade (blue), speculators (orange), net order flow (green). Gray line: transaction price (right axis). Dashed line: $\lambda_N Q = 5$. Panels 12(a), 12(c), and 12(e) show order flows under no commitment; panels 12(b), 12(d), and 12(f) show order flows under full commitment. Parameters: $T = 10$, $H = 9$, $\kappa = 10$, $Q = 10$, $\lambda_N = 0.5$, $\rho_N = 0.55$, $\lambda_O = 0.2$, $\rho_O = 0.2$, $\nu = 0.15$, $c = 0.3$.

rather than strengthened. Competitive speculators pre-position against the predictable passive flow, moving more of the price pressure into the pre-implementation path and compressing the incremental spike at T. With a large $\bar{Q}$, this pre-intermediation is strong enough that the weak tracker is pulled back toward implementation-date execution. In the equal-split case, the passive block is smaller and the weak tracker itself accounts for a larger share of total demand; stronger competition still raises the attractiveness of period T, but the schedule remains more dispersed because the fund continues to benefit from trading against the post-implementation reversal.

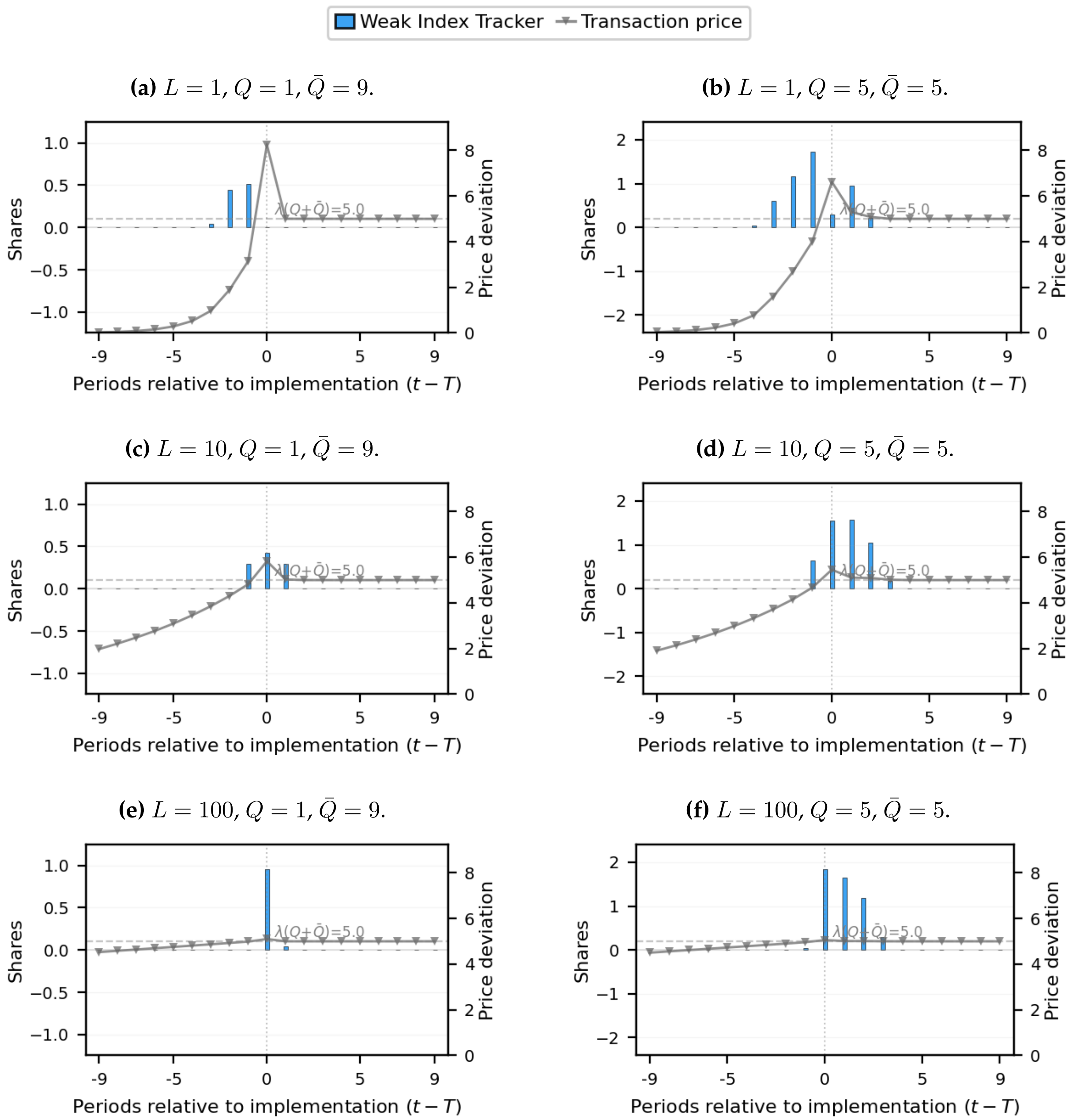
Full commitment (Figure 14) reinforces the same qualitative lesson. The committed weak tracker internalizes how its announced schedule changes speculator positioning, so it reacts less aggressively to the passive block than under no commitment. When L is low, even commitment does not eliminate the incentive to lean away from period T, because there is too little speculative capital to absorb the passive shock. When L is high, however, the large implementation flow has already attracted enough intermediation that trading around it is no longer as valuable. The committed weak tracker can therefore keep most of its own demand at the implementation date, and in highly competitive markets this concentration arises even when the passive block is large.

The general implication is that large implementation-date order flow does not necessarily push weak trackers away from the auction: in thin markets it does, but in competitive markets the same predictable flow attracts intermediation that diminishes the gains from avoiding it. Therefore, competitive speculators introduce a liquidity complementarity in the order flow of strict and weak index trackers: as the relative demand by fully passive investors increases and such demand is well intermediate, other benchmarked investors are increasingly incentivized to trade together with the large passive block. This mechanism is consistent with the evidence in Chinco and Sammon (2024), who show that nearly half of the excess demand for stocks at index additions can be attributed to institutions that are not strict index trackers.

6 Interpreting Index Events Through the Model

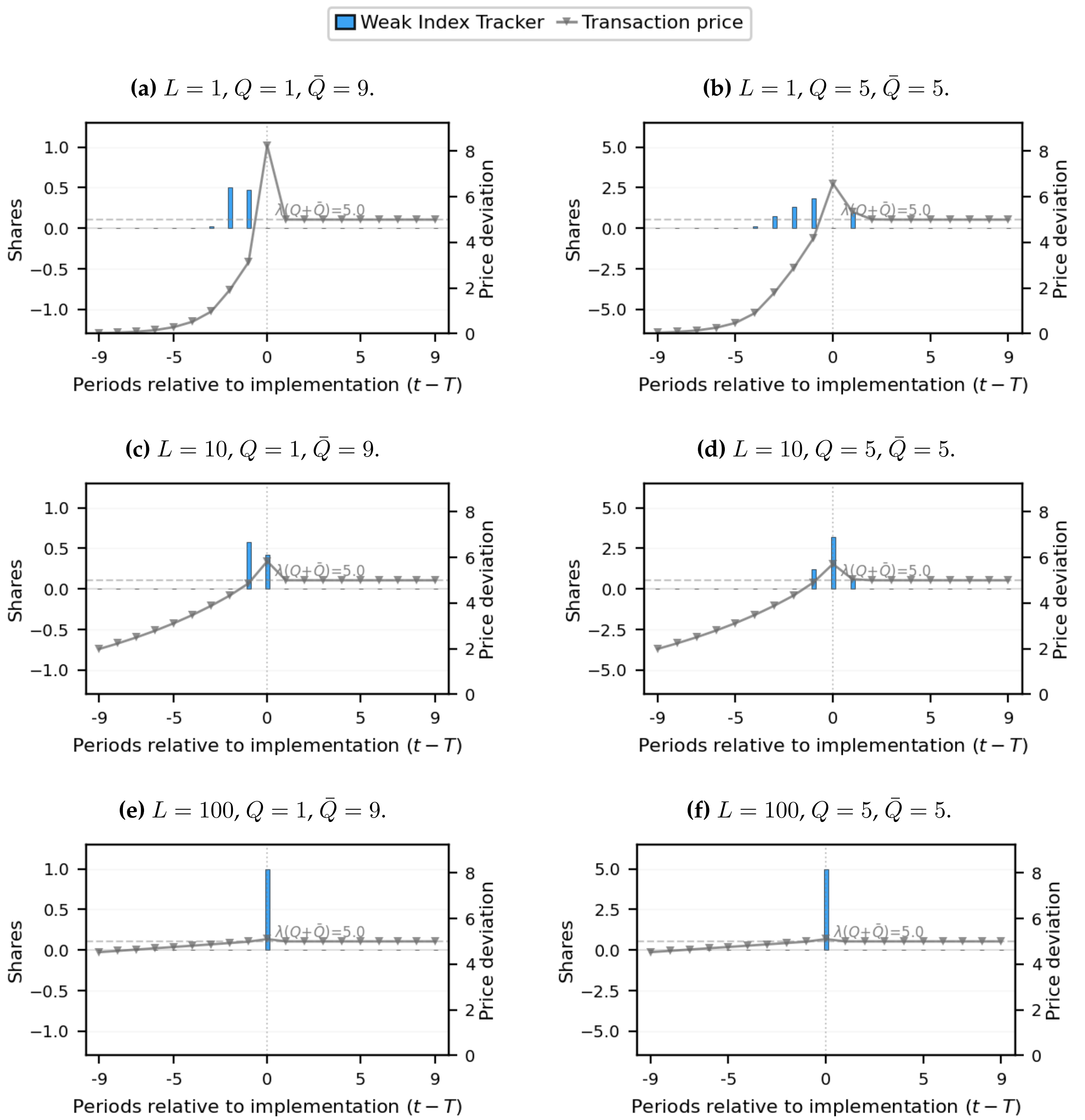
This section uses the model to reinterpret several empirical facts around index reconstitution. The first fact examines implementation-day trading volume relative to demand from strict index trackers, which speaks to apparent coordination among benchmark-linked investors. The second fact uses post-implementation day reversals to measure the avoidable component of price impact. Third, we use short interest to identify how intermediaries build and unwind positions around index reconstitutions. The fourth fact stud-

Figure 13: No-commitment trading schedule of a weak index tracker alongside a passive order flow.



Notes. The passive order flow is taken as exogenous at period T: it buys $\bar{Q}$ shares of N and sells $\bar{Q}$ shares of O. Bars: weak index tracker trade (blue). Gray line: transaction price (right axis). Dashed line: $\lambda_N(Q + \bar{Q}) = 5$. Panels 13(a), 13(c), and 13(e) show $Q = 1$, $\bar{Q} = 9$; panels 13(b), 13(d), and 13(f) show $Q = 5$, $\bar{Q} = 5$. Parameters: $T = 10$, $H = 9$, $\kappa = 10$, $\lambda_N = 0.5$, $\rho_N = 0.55$, $\lambda_O = 0.2$, $\rho_O = 0.2$, $\nu = 0.15$, $c = 0.3$.

Figure 14: Full-commitment trading schedule of a weak index tracker alongside a passive order flow.



Notes. The passive order flow is taken as exogenous at period T: it buys $\bar{Q}$ shares of N and sells $\bar{Q}$ shares of O. Bars: weak index tracker trade (blue). Gray line: transaction price (right axis). Dashed line: $\lambda_N(Q + \bar{Q}) = 5$. Relative to no commitment, the full-commitment weak-index-tracker schedule reacts less aggressively around the passive block; stronger speculator competition ($L = 100$) further compresses implementation-date net-flow spikes. Panels 14(a), 14(c), and 14(e) show $Q = 1$, $\bar{Q} = 9$; panels 14(b), 14(d), and 14(f) show $Q = 5$, $\bar{Q} = 5$. Parameters: $T = 10$, $H = 9$, $\kappa = 10$, $\lambda_N = 0.5$, $\rho_N = 0.55$, $\lambda_O = 0.2$, $\rho_O = 0.2$, $\nu = 0.15$, $c = 0.3$.

ies the link between variation in how predictable index reconstitutions are and abnormal trading volume. Finally, we combine reversals with the size of index-linked demand to provide a back-of-the-envelope calculation for the effective cost borne by index investors.

The main empirical setting is S&P 500 direct additions, events that involve stocks entering the S&P 500 from outside the S&P index universe (i.e., from outside the S&P 400 mid cap and S&P 600 small cap). These are useful in our setting, as we do not need to estimate the size of possibly offsetting demand shocks (as would be the case for e.g., migrations from the S&P 400 to the S&P 500). For some tests, we also use Russell index migrations to study cases where the expected index demand shock is negative (e.g., when stocks switch from the Russell 2000 to the Russell 1000). In addition, given the mechanical nature of Russell reconstitution, we use Russell indexes to isolate when stocks have anticipated and less-anticipated demand shocks.[3]

6.1 Coordinated Trading on the Implementation Day

The first fact builds on Chinco and Sammon (2024), showing that implementation-day trading is too large to be explained by the demand of strict index trackers (i.e., index funds and ETFs) and loose index trackers (identified using the method in Behmaram and Davis (2026)) alone. Quantifying trading from strict and loose trackers is important because this will be needed to quantify the avoidable component of trading costs for index fund investors.

To measure demand from strict index trackers, we identify plain-vanilla index mutual funds and ETFs tracking the S&P 500. We use CRSP mutual fund classifications, fund names, active share and realized tracking error to classify funds. For example, we identify S&P 500 funds using the CRSP/Lipper S&P 500 flag when available, as well as name-based matches for funds whose names explicitly reference the S&P 500. For this exercise, we exclude style, sector, factor, ESG, non-value weighted, and enhanced-index products. We also drop funds whose holdings counts or realized tracking errors are inconsistent with tracking the S&P 500 (e.g., stocks with a stated benchmark to the S&P 500 but hold less than 400 securities). This procedure gives us a conservative measure of strict index-fund demand for each index.

We then supplement this measure using the broader shadow-indexing measure from Behmaram and Davis (2026), who expand on Pavlova and Sikorskaya (2023) by construct-

[3]We do not use Russell tests when examining reversals or, in general, when tests involve prices. The mechanical Russell reconstitution involves ranking stocks based on market capitalization, which is directly affected by prices. That is, returns in the days and weeks leading up to the rank date can directly affect anticipated index-linked demand.

ing a *holdings*-based measure of realized passive ownership adjusted for Active Share (Cremers and Petajisto, 2009). This broader measure is useful because many benchmark-linked investors do not explicitly identify themselves as index funds, but still hold portfolios that closely resemble an index and therefore may trade around index events to minimize tracking error.

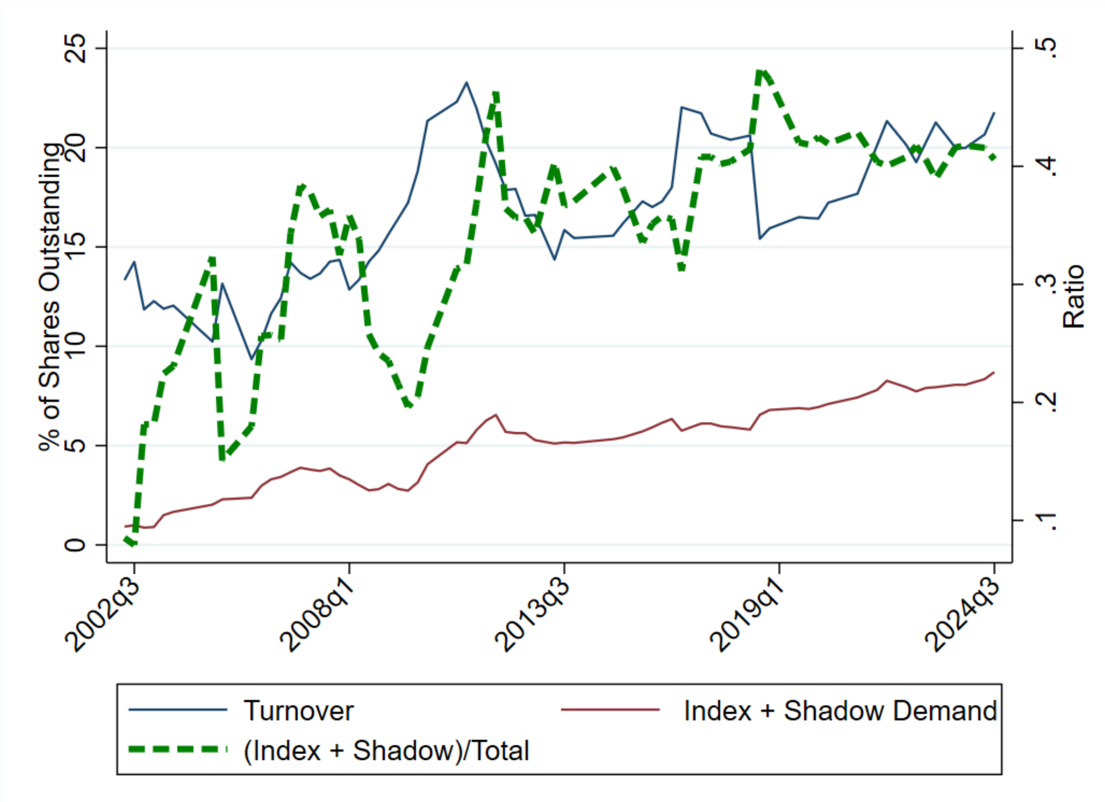
Figure 15 shows the comparison of S&P 500 direct addition trading activity and predicted index-linked demand. The blue solid line shows total implementation-day turnover. The red solid line shows estimated demand from strict S&P 500 index funds plus shadow S&P 500 index trackers. The dashed green line shows the ratio of index-plus-shadow demand to total turnover. Two facts stand out. First, total turnover on the implementation day is extremely large, often exceeding 15% to 20% of shares outstanding. Second, even after including shadow-index demand, benchmark-linked demand explains only part of total implementation-day trading (less than 50%). The red line shows that, if all strict and shadow index trackers traded on implementation day, we would expect turnover of less than 10% of shares outstanding. The ratio of index demand to implementation-day turnover, plotted in the dashed green line, shows this more clearly: the ratio rises over time but remains well below one and is closer to 0.4.

Our interpretation of these patterns is that implementation-day trading is not limited to strict index funds. Of course, one might argue that some of the additional volume is driven by intermediation, but given that most trading on these days takes place in the closing auction, or at trades linked to the closing auction price (Chinco and Sammon, 2024), we believe this is unlikely to be the main driver of the difference. Our model provides an interpretation for the "missing" passive investors: many non-strict index trackers and shadow indexers not captured by the Behmaram and Davis (2026) measure, as well as other benchmark-linked investors, choose to trade on the same day as strict trackers. This is exactly the type of coordination emphasized by the model. When tracking-error concerns are sufficiently high, or when intermediary competition is sufficiently strong, benchmark-linked investors optimally concentrate trading on the implementation day.

6.2 Post-Implementation Reversals and the Avoidable Component of Price Impact

The second fact is that post-implementation price reversals have fallen over time. In standard execution models, we typically find that larger trades generate larger temporary price impact and therefore larger reversals. Our model predicts, however, that the relationship between trading volume and reversals depends on the degree of interme-

Figure 15: Implementation-Day Turnover and Benchmark-Linked Demand



Notes. This figure shows implementation-day trading and demand from strict and shadow index trackers for S&P 500 direct additions. The blue line shows total implementation-day turnover, measured as trading volume scaled by shares outstanding. The red line shows estimated index-plus-shadow demand, also scaled by shares outstanding. The dashed green line shows the ratio of index-plus-shadow demand to total implementation-day turnover. Strict index trackers are identified using CRSP mutual fund classifications, fund names, holdings counts, and realized tracking error. Shadow-index demand is based on the holdings-based approach in Behmaram and Davis (2026). The figure presents five-year moving averages.

diation. As competition among speculators increases – i.e., as L increases – more of the implementation-day demand is pre-positioned. This spreads price impact over time and reduces the temporary component of price impact that reverts – all of the shares demanded by index trackers is provided with little to no price impact on the implementation day.

The model also clarifies the right empirical object to quantify the amount and cost of intermediation taking place. While the raw reversal (i.e., return) is informative, the model's analogue is the reversal scaled by the size of the index demand shock:

$$\Gamma(L) \equiv \frac{P_{\text{post}} - P_{\text{implementation}}}{Q}.$$

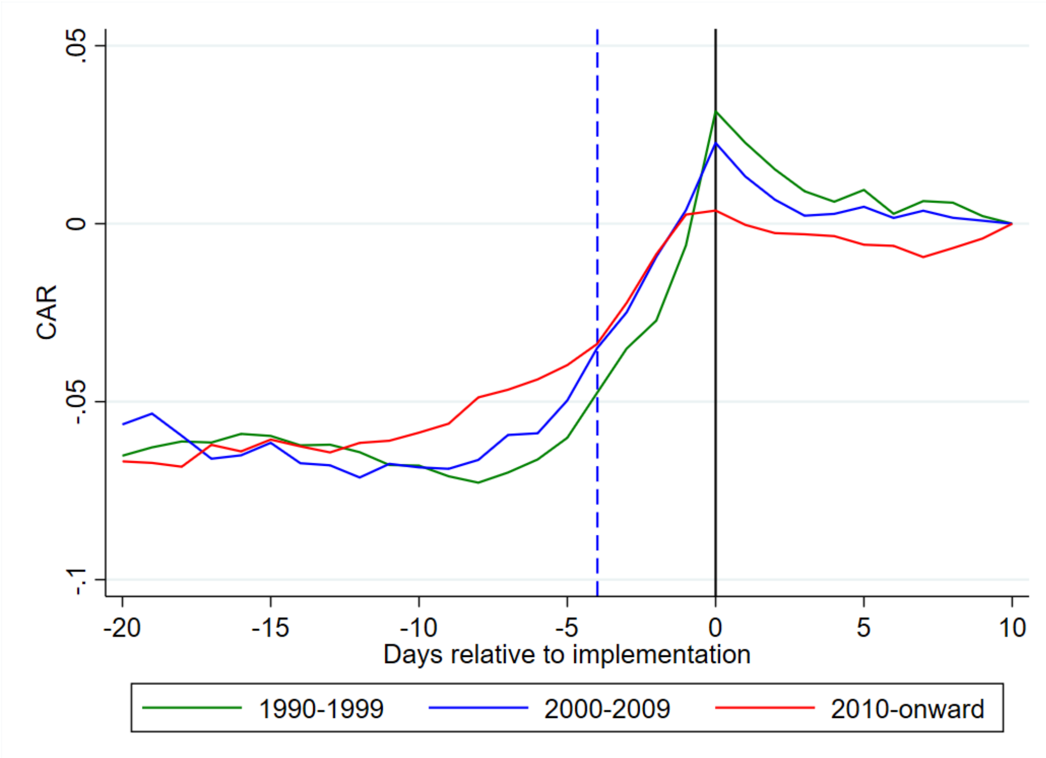
This object measures the avoidable component of price impact per unit of benchmark-linked demand, not the full implementation shortfall faced by index investors. If index demand grows over time while reversals shrink, then $\Gamma(L)$ must have fallen substantially. The decline in reversals *despite* growing index-linked demand therefore points to a large increase in intermediation capacity or competition among intermediaries.

Figure 16 shows basic return evidence for S&P 500 direct additions. Specifically, we plot cumulative abnormal returns around implementation by decade, normalizing prices so that the post-event price is the same 10 days after the implementation day ($t = +10$). This normalization makes it easier to see how the reversal has changed over time. In the 1990s, prices rise into implementation and then reverse noticeably afterward. In the 2000s, the reversal is smaller even though index-linked demand is larger. In fact, the 2000s is when many index ETFs were introduced and saw rapid growth as an investment vehicle. In the 2010-and-onward period, the post-implementation reversal is essentially zero, despite still larger index-linked demand. Importantly, the *total* price movement is roughly similar across time – the starting point at $t = -20$ is almost the same across decades even though returns have been normalized to zero at $t = +10$. The major change over decades is how much of the implementation-day price effect takes place before implementation. For example, in the 2010 and onward sample, a larger share of the return happens before the *announcement* itself.

Figure 17 shows the time-series version of the same fact. Panel A plots five-year moving averages of post-event reversals for S&P 500 direct additions, using five-day and ten-day post-implementation-day windows. In the 1980s, there was continuation, while starting in the 1990s, there was post-implementation reversal. Reversals started to shrink around 2000 and are nearly zero by the end of our sample. Panel B divides the same reversal by the size of the index-linked demand shock estimated in Figure 15. We focus on the post-2002 period for this sample based on data availability for the size of the index-linked demand shock. Panel B shows that the decline in reversals is even more striking after this normalization. Index-fund demand has increased substantially, yet the reversal per unit of demand has fallen to almost zero. In addition, we know from Section 6.1 that our measure of index-linked demand accounts for a little less than half of all demand as measured by implementation-day turnover, most of which comes from the closing auction. Accounting for the hidden index-linked demand would push the reversal per unit of demand even closer to zero for every period.

This joint pattern is difficult to reconcile with the view that larger index demand mechanically creates larger temporary price distortions. It is instead exactly what the model predicts when intermediary competition increases. As L rises, speculators compete more

Figure 16: Price Paths Around S&P 500 Direct Additions by Decade



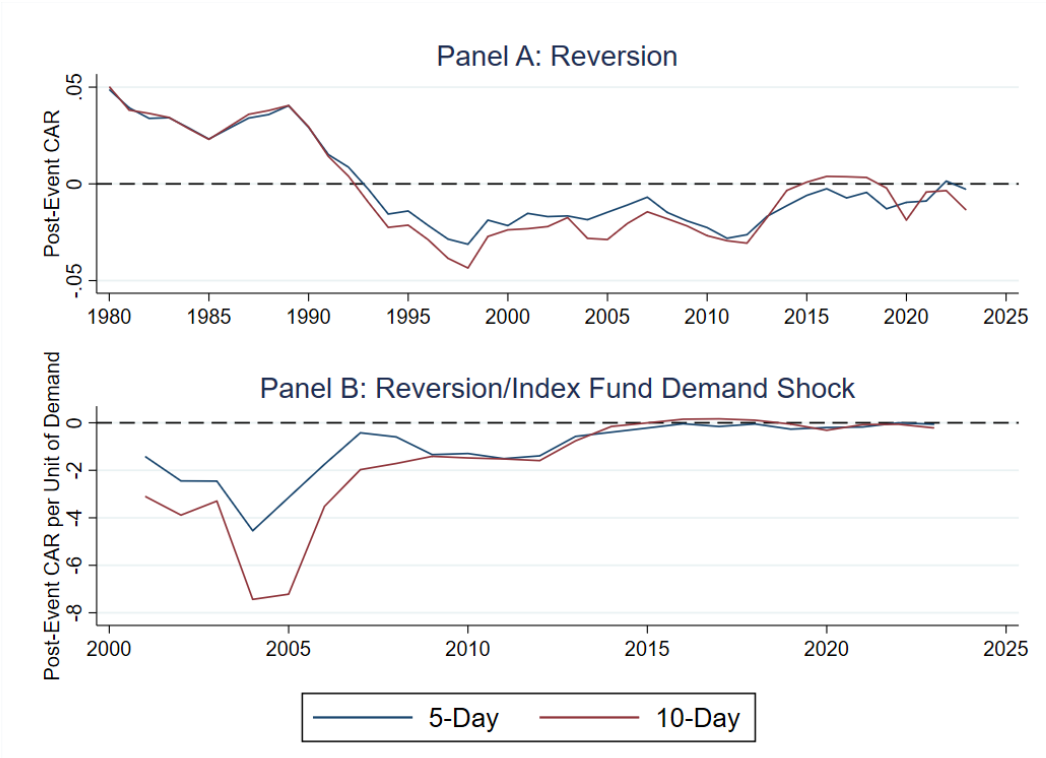
Notes. This figure shows cumulative abnormal returns around S&P 500 direct additions, separated by decade. Event time is measured relative to the implementation day of the index reconstitution. The solid vertical line marks the implementation day, and the dashed vertical line marks the average announcement day. Cumulative abnormal returns are normalized so that each decade's average price path equals zero at day $t = +10$.

aggressively to supply liquidity at implementation. This pushes the temporary component of price impact down, even as the quantity traded at implementation rises. Thus, in light of the model, the empirical fact that quantities have increased while reversals have decreased suggests that the market for intermediating index events has become much more competitive over time.

6.3 Short Interest Dynamics

The evidence on reversals suggests that intermediation has become more competitive over time. In this subsection, we examine securities-lending data to study how that intermediation occurs. In the model, spreading trades across more periods is valuable because it reduces temporary price impact. One way speculators do this is by acquiring positions outside of the implementation day and selling into benchmark-linked demand at implementation. One way to conduct this trade is through short selling on the implementation

Figure 17: Post-Implementation Reversals Around S&P 500 Direct Additions



Notes. This figure shows post-implementation reversals for S&P 500 direct additions. Panel A plots five-year moving averages of post-event cumulative abnormal returns measured over five- and ten-day windows after implementation. Panel B plots the same reversals divided by the estimated index-fund demand shock estimated in Figure15. The model implies that the demand-normalized reversal is the empirical analogue of the avoidable component of price impact per unit of index demand. A value closer to zero indicates less temporary price impact.

day and unwinding afterward. Specifically, for index additions, speculators can borrow shares on the implementation day, sell those shares to index trackers buying at implementation, and then buy shares back over time to cover the short position. In this sense, the post-implementation period gives intermediaries additional time to spread out their trades. In other words, rather than all intermediation occurring before the event, some of it can occur after the event through short covering. This mechanism is also useful to examine for data purposes – speculator purchasers will be mixed in with other investors' purchases, but speculator sales will likely show as short sales, distinguishing them from "long sales" from investors. In this sense, securities lending and short interest data can provide insight into speculator activity more specifically.

For this purpose, we use Markit data from 2010 to 2023. The data include short interest, borrowing fees, and lendable quantity. We also compute utilization, defined as short interest divided by lendable quantity. This last measure is useful because index inclusion can itself increase lendable supply: once index funds and ETFs own the stock, some of those shares may become available for lending. Thus, both the numerator and denominator of short-interest measures can move around when there are index reconstitutions. Because shorting settles at T+2 trading days, we align shorting with event timing by shifting the Markit data backward by two trading days.

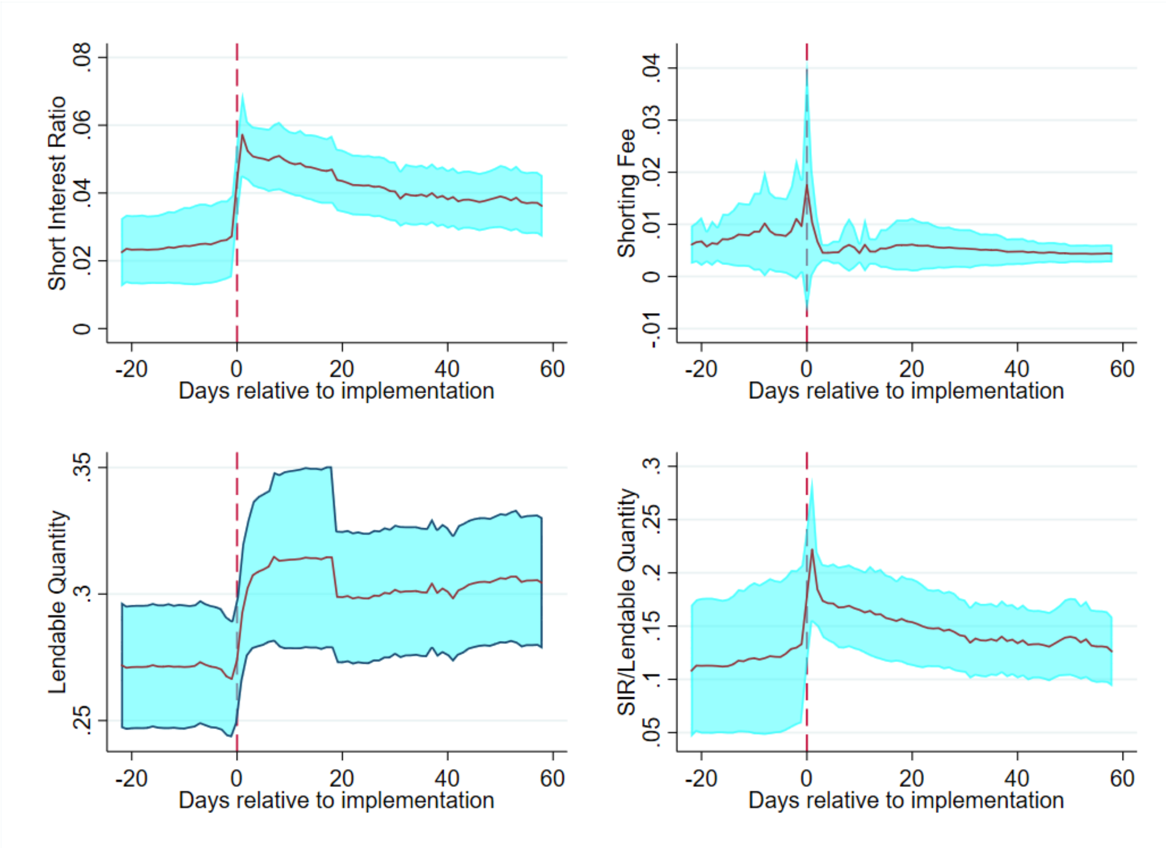
Figure 18 shows the pattern for S&P 500 direct additions. Short interest rises sharply around implementation and then declines gradually afterward, consistent with speculators entering into short positions to satisfy index-linked implementation demand. Lendable quantity also increases after implementation, consistent with the idea that index funds are willing to lend out their shares. Borrow fees spike around implementation and then revert, consistent with the increased supply of shares from passive funds driving down the fee for shorting. However, shorting increases even relative to this increase in supply. Therefore, the overall patterns in Figure 18 are consistent with intermediaries borrowing shares, selling them into benchmark-linked demand, and then covering over time.

Russell migrations provide an additional useful test because the sign of the net demand shock can differ across migration types. Stocks moving from the Russell 1000 to the Russell 2000 experience a positive demand shock because more dollars track the Russell 2000 than the Russell 1000 relative to the total capitalization of each index. Figure 19 shows that short interest rises around implementation and then declines afterward, similar to the evidence in Figure 18, which are also positive net index-tracker demand shocks.

The reverse migration is especially useful. Stocks moving from the Russell 2000 to the Russell 1000 experience a negative index-linked demand shock. In this case, the model predicts the opposite pattern: rather than borrowing and shorting more shares to supply positive demand at implementation, intermediaries should reduce short positions at implementation. Figure 20 shows this is exactly the case empirically. Short interest falls around implementation. This symmetry is important because it shows that the short-interest evidence is not simply a generic feature of any index reconstitution. Instead, it captures speculator activity and moves in the direction predicted by the sign of benchmark-linked demand.

Taken together, the securities-lending evidence supports the model's liquidity provision mechanism. The model highlights that the post-implementation period is also useful for intermediation. Specifically, these are another set of trading periods that intermediaries can use to spread price impact. Intermediaries spread price impact by trading before and after the event, with shorting serving as a measure of intermediary activity.

Figure 18: Short Interest and Lending Markets for S&P 500 Direct Additions

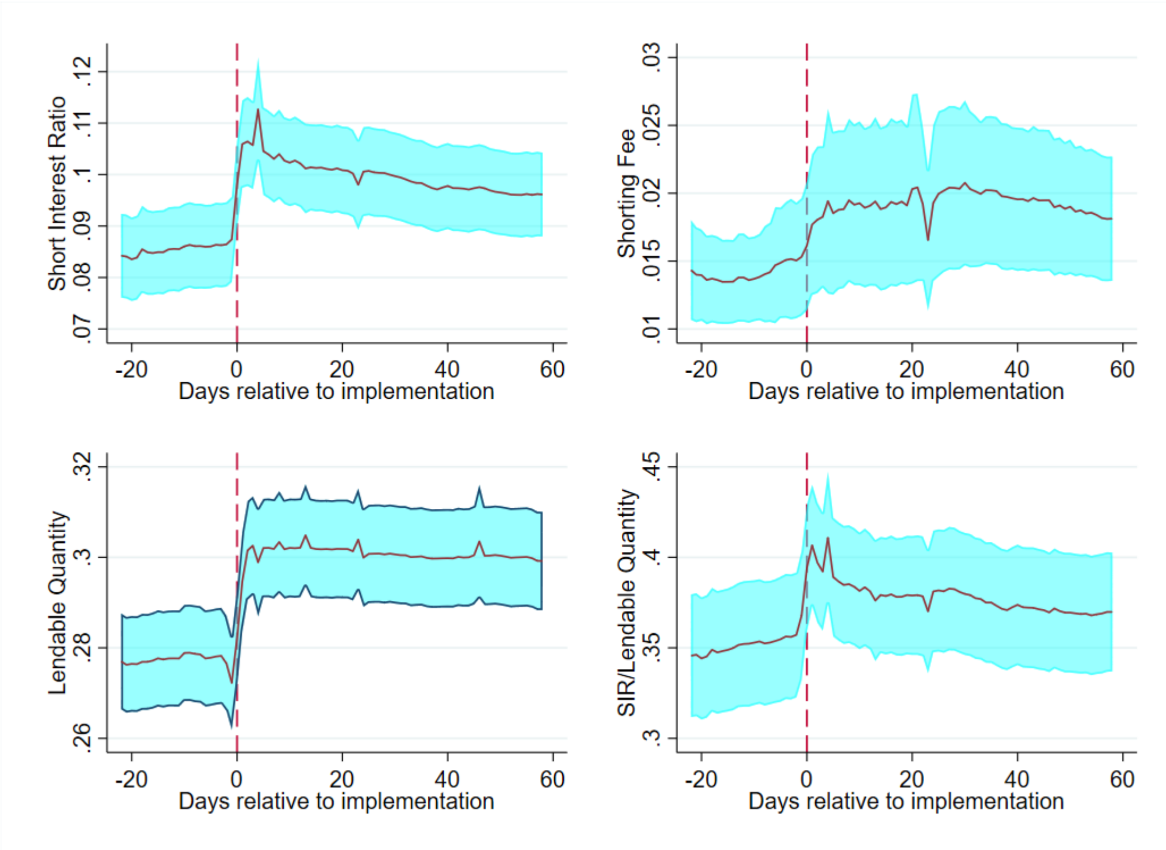


Notes. This figure shows securities-lending variables for S&P 500 direct additions around implementation days of index reconstitutions. The four panels show the short-interest ratio, indicative borrow (shorting) fee, the lendable quantity, and short interest divided by the lendable quantity. Event time is measured relative to the implementation day. The red line shows the average event-time path, and the shaded region shows confidence intervals. The dashed vertical line marks the implementation day. The sample uses Markit securities-lending data from 2010 to 2023.

6.4 Predictability and Last-Minute Intermediation Around Russell Migrations

The extended model with multiple pre-implementation periods emphasizes the importance of predictability. When the demand shock is known far in advance, intermediaries have more time to pre-position. When there is little time to pre-position, intermediation must occur in a compressed time frame – effectively, there are fewer pre-implementation periods to take advantage of. Based on this logic, we expect less predictable events to generate more trading concentrated between the reconstitution announcement and the implementation date, as well as more post-event trading (given the shorting and post-event covering mechanism described above).

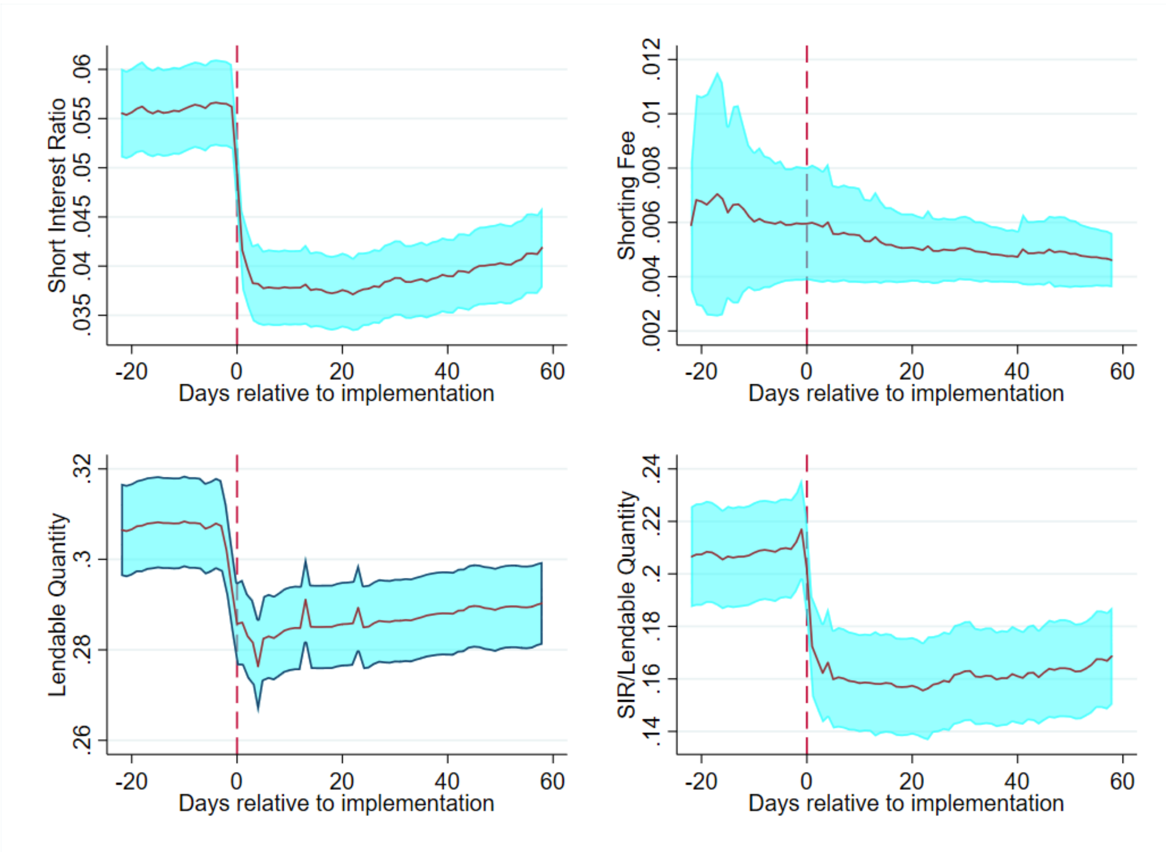
Figure 19: Short Interest and Lending Markets Around Russell 1000 to Russell 2000 Migrations



Notes. This figure shows securities-lending variables for Russell 1000 to Russell 2000 migrations around implementation days of index reconstitutions. These migrations generally correspond to positive net benchmark-linked demand because more assets track the Russell 2000 than the Russell 1000. The four panels show the short-interest ratio, indicative borrow (shorting) fee, the lendable quantity, and short interest divided by lendable quantity. Event time is measured relative to the implementation day. The red line shows the average event-time path, and the shaded region shows confidence intervals. The dashed vertical line marks the implementation day. The sample uses Markit securities-lending data from 2010 to 2023.

Russell migrations are useful for studying this idea because index assignment is rule-based and depends on market capitalization ranks. Some migrations are "slam dunks" by the rank date: the stock is far enough over the cutoff that the migration was likely anticipated far in advance. Other migrations are close calls. For these stocks, changes in returns for one or more stocks could reshuffle rankings and change whether stocks stay in their current index or move to a different index. As a result, the market learns relatively late (and closer to or on the actual rank date) whether the stock will switch indexes. We expect that close-call switchers should require more last-minute intermediation than slam-dunk switchers, and therefore have more trading volume just before and after the implementation date.

Figure 20: Short Interest and Lending Markets Around Russell 2000 to Russell 1000 Migrations



Notes. This figure shows securities-lending variables for Russell 2000 to Russell 1000 migrations around implementation days of index reconstitutions. These migrations generally correspond to negative net benchmark-linked demand because more assets track the Russell 2000 than the Russell 1000. The four panels show the short-interest ratio, indicative borrow (shorting) fee, the lendable quantity, and short interest divided by lendable quantity. Event time is measured relative to the implementation day. The red line shows the average event-time path, and the shaded region shows confidence intervals. The dashed vertical line marks the implementation day. The sample uses Markit securities-lending data from 2010 to 2023.

Table 1 tests this idea. The sample includes Russell migration candidates from 2013 to 2024, focusing on years in which the net demand shock exceeds 1% of shares outstanding. We compare three groups: close switchers within 50 ranks of the cutoff, close non-switchers within 50 ranks of the cutoff, and slam-dunk switchers more than 50 ranks away from the cutoff.[4] We estimate a regression that examines turnover for these three groups. The dependent variables are turnover and abnormal turnover in three windows: from the rank date to the day before the implementation date, the implementation day itself, and the 40 days after the implementation date. Abnormal turnover is defined as the

[4]For more details on how these cutoffs are computed in the "banding" regime, see, e.g., Coles et al. (2022).

residual from an OLS regression of turnover on the stock's pre-rank return, lagged Russell index membership, and year fixed effects. We regress each turnover variable and in each of the three periods on dummies to capture group membership, as well as pre-rank turnover as a control. The omitted group is the slam dunk switchers.

The main result is that close switchers have significantly higher turnover and abnormal turnover before and after the implementation date (relative to the slam dunk switchers), though the post implementation raw turnover is not statistically significant. The implementation-day coefficient is small, which is a useful placebo test: close-call switchers do not just (all else equal) have more trading volume at every point in time. Rather, Table 1 suggests that when index assignment is resolved late, the intermediation process is more compressed and leads to more trading around the implementation day (both before and after). However, the implementation day trading is very similar, consistent with the idea that the implementation day volume is driven by index-linked demand, which is the same for switchers (as a percentage of shares outstanding) regardless of whether the stock was close or far from the rank cutoff.[5]

Table 1 also shows that the stocks that do not end up switching have much smaller turnover on all days surrounding implementation, including the implementation day itself. Since these stocks stay in their existing index, there is essentially no index-linked demand shock. This means that there is no demand shock to drive up turnover on the implementation day, and no interest from speculators in the days surrounding implementation.

6.5 How Large Are the Costs to Index Investors?

The previous facts show that intermediation for index reconstitutions has become more competitive and that the temporary component of price impact has fallen. We now use the model to translate these reversals into a simple measure of the cost of intermediation for index fund investors.

As mentioned above, the model implies that the post-implementation reversal captures the avoidable component of the trading cost. Permanent price impact is not avoidable: the market must ultimately absorb the benchmark-linked demand shock, and this will move prices. By contrast, temporary price impact is the part that reverts after implementation and could be avoided through smoother trading. A simple empirical analogue

[5]The Russell indexes (as well as almost all value-weighted stock indexes) simply hold a constant fraction of (float-adjusted) shares outstanding, regardless of rank. While higher ranked stocks will have larger weights, each stock in an index-tracker holds the same fraction of shares outstanding for each stock in the index (float adjustments tweak this slightly). See Sammon and Shim (2026) for more details.

Table 1: Turnover Around Close-Call Russell Migrations

	Rank Day to Implemen. Day -1		Implemen. Day		Implemen. Day $+1$ to $+40$ Days	
	Turnover	Abnormal Turnover	Turnover	Abnormal Turnover	Turnover	Abnormal Turnover
Within 50 Ranks + Switch	**0.138****	**0.214*****	5.13E-04	0.00437	0.200	0.312***
	(0.059)	**(0.054)**	(0.005)	(0.006)	(0.111)	(0.095)
Within 50 Ranks + No Switch	-0.302*	-0.245*	-0.0265	-0.0236	-0.313**	-0.230*
	(0.140)	(0.135)	(0.020)	(0.021)	(0.103)	(0.112)
Pre-Rank Turnover	0.299***	0.251***	0.00963***	0.00743***	0.542***	0.477***
	(0.053)	(0.055)	(0.001)	(0.001)	(0.042)	(0.040)
Within 50 Ranks + Switch × Pre-Rank Turnover	-0.0611	-0.0993**	-0.000328	-0.00215	-0.200***	-0.253***
	(0.041)	(0.039)	(0.002)	(0.002)	(0.036)	(0.043)
Within 50 Ranks + No Switch × Pre-Rank Turnover	0.169	0.180**	-0.00549	-0.00509	0.121	0.132
	(0.098)	(0.076)	(0.006)	(0.006)	(0.085)	(0.080)
Observations	368	368	368	368	368	368
R-squared	0.696	0.553	0.794	0.780	0.689	0.609
FE				Year		

Notes. This table studies turnover around Russell migration candidates close to the Russell 1000/Russell 2000 cutoff. "Within 50 Ranks + Switch" is an indicator for stocks within 50 ranks of the cutoff on the rank date that switch indexes. "Within 50 Ranks + No Switch" is an indicator for stocks within 50 ranks of the cutoff that do not switch. The omitted group is "slam-dunk" switchers, or stocks that are more than 50 ranks away from the cutoff (and thus are certain to switch). The dependent variables are turnover and abnormal turnover measured from the rank date to the day before the implementation date, on the implementation day, and from one to 40 trading days after the implementation date. Abnormal turnover is residual turnover after controlling for pre-rank returns, lagged Russell index membership, and year fixed effects. Standard errors are clustered by year. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

of the avoidable cost is therefore

$$\text{Avoidable Cost}_e \approx \text{Benchmark-linked dollars traded}_e \times \text{Post-event reversal}_e.$$

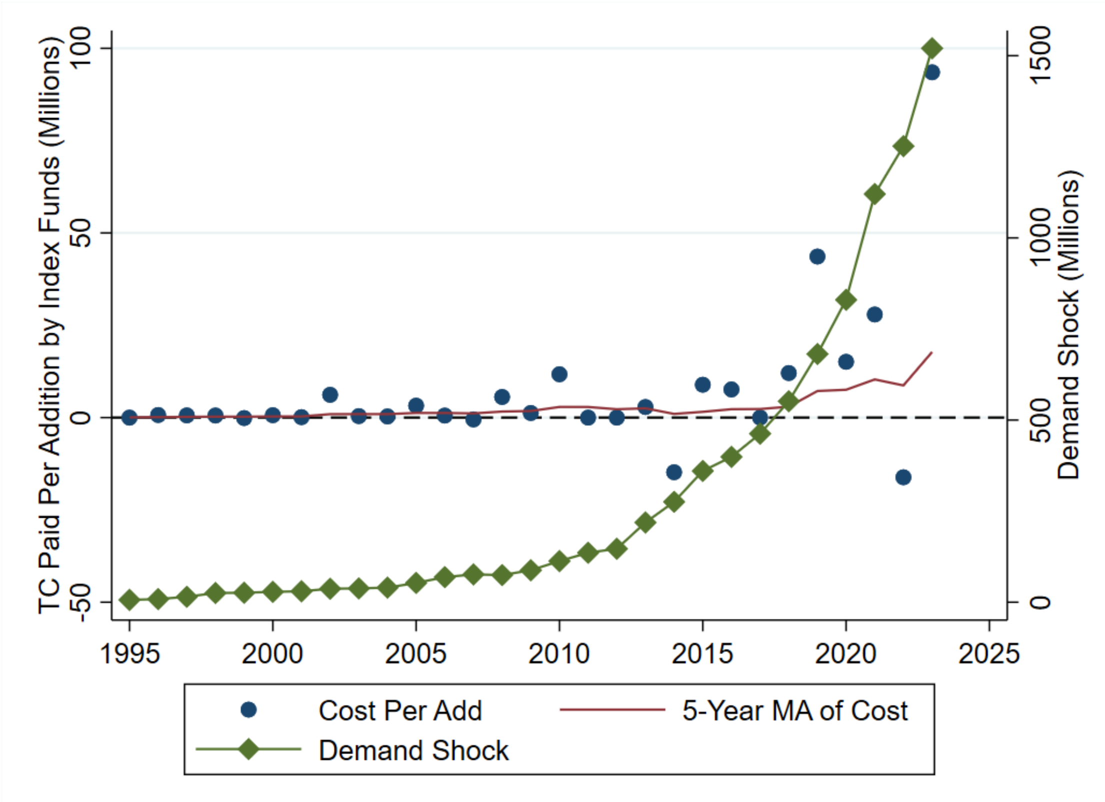
Equivalently, in the notation of the model,

$$\text{Avoidable Cost}_e \approx Q_e \times \left(P_{\text{implementation},e} - P_{\text{post},e}\right).$$

This measure is not the total implementation shortfall. If speculators acquire shares in advance and sell them to index trackers at implementation, some of the total cost may be embedded in the pre-event price. The point is instead that the reversal measures the temporary and avoidable part of price impact associated with concentrated implementation-day trading.

Figure 21 plots the resulting cost estimate for S&P 500 direct additions over time. The green line shows the index-linked demand shock, which has grown dramatically as indexing has expanded. The blue points show the estimated trading cost paid per addition by index funds, and the red line shows the 5-year moving average of this cost. The key fact is that benchmark-linked demand has grown enormously, but the estimated cost has not grown proportionally. In recent years, the cost is roughly 1% of the benchmark-linked trade.

Figure 21: Estimated Avoidable Trading Costs for S&P 500 Direct Additions



Notes. This figure shows a back-of-the-envelope estimate of avoidable trading costs for S&P 500 direct additions. The green line shows the index-linked demand shock, measured in millions of dollars and plotted on the right axis. The blue points show the estimated cost paid per addition by index funds, measured in millions of dollars and plotted on the left axis. The red line is a 5-year moving average of the cost estimate. The cost is computed as benchmark-linked dollars traded multiplied by the post-implementation reversal.

Our estimated cost per trade – especially given the size of these trades – is extremely small. For example, Frazzini et al. (2012) find that price impact is roughly 40 bps for a trade about 10% of ADV. By contrast, S&P 500 additions can involve benchmark-linked trades over 2500% of ADV, and yet the price impact is only about 100 bps. In this sense, the surprising fact is not that index investors face trading costs – any trade must generate costs in some form – it is that index investors can trade such enormous quantities at a cost that appears very low relative to standard benchmarks.

Again, the model is useful because it can resolve this seeming puzzle. As index demand has grown, the incentive to intermediate index events has also grown. More intermediaries enter the trade, competition increases, and the temporary component of price impact falls. This is why rising Q and falling reversals are jointly informative. In a model without increasing intermediary competition, larger demand shocks should *mechanically* generate larger reversals. In our model, the fact that reversals fall while Q rises indicates that the market has become substantially more efficient at intermediating index demand.

6.6 Implications for Index Design

Collectively, the empirical evidence implies that predictable index demand is potentially costly to index fund investors, but predictability also allows intermediaries to prepare and reduce the temporary component of price impact. The relevant counterfactual is therefore not a world in which index funds trade for free. The relevant counterfactual is a world in which index trackers must trade large quantities without intermediaries providing shares to satisfy demand.

This has implications for index design – and we suspect that index providers are aware of these channels. When index changes are predictable and the implementation window is long enough, intermediaries can pre-position and supply liquidity at the close. This can reduce implementation-day price impact and lower the avoidable component of trading costs for index investors. In this sense, transparency and commitment can be valuable because they coordinate liquidity. For example, starting in 1989, S&P started pre-announcing index changes. Similarly, Russell has increased the gap between the ranking date and the implementation date, another way to give intermediaries more time to acquire shares and spread out trading over more days.

At the same time, this logic does not imply that increased predictability is always good. If the pre-positioning window is too short, intermediaries may not have enough time to acquire inventory and smooth price impact. In those cases, implementation-day price impact and post-event reversals may reappear, as explained in Section 5.2 and shown in Murray and Sammon (2026) for post-IPO fast-track additions to CRSP indexes. The reason is that there simply are not enough periods after the IPO day and before the implementation day, and shorting is restricted due to IPO lock-ups, preventing speculators from smoothing trading over the days after implementation. Similarly, if predictability attracts informed trading or creates adverse selection, index providers may face additional costs. The point of the model is to clarify this tradeoff. Concentrated implementation-day trading can be optimal when tracking-error concerns are important and intermediaries are competitive, but the benefits depend on giving intermediaries enough time and certainty to supply liquidity.

7 Conclusions

We show that apparent front-running around index reconstitution events is a form of liquidity provision that benefits index investors. Our model also rationalizes the empirical disconnect between trading volume and price changes around such events, whereby large reconstitution-day trading volumes, with over 20% of shares outstanding trading, occur with small implementation-day returns.

The model delivers four main takeaways. First, speculators act as endogenous liquidity providers rather than pure predators. They pre-position before implementation, trade against benchmark-linked demand at reconstitution, and unwind short positions after implementation. This mechanism can generate price pressure before the event, but reduces implementation-day price impact and post-implementation reversal. Second, competition between front-runners plays a central role. As competition among speculators increases, index trackers—including those that only loosely track the index—optimally concentrate rebalancing on the implementation date. Third, the model identifies post-implementation price reversal as the measure of the avoidable cost borne by index-linked investors. Fourth, we establish a liquidity complementarity between strict and weak index trackers. In highly competitive markets, loose index trackers increasingly concentrate their rebalancing at implementation when the size of the passive rebalancing volume increases.

We use the model to interpret several empirical facts. Implementation-day volume is much larger than strict index-fund demand alone, suggesting that the implementation date serves as a coordination point for strict index trackers, shadow indexers and benchmarked investors. Moreover, post-implementation reversals have fallen sharply over time for S&P 500 direct additions, even as benchmark-linked demand has grown. While the raw reversal is interesting in its own right, the model implies that the relevant object is the reversal scaled by the size of the demand shock, and this demand-normalized reversal has also fallen substantially over the past two decades. Under the lens of the model, this suggests competition among speculators has increased and that speculators, as a result, effectively provide liquidity at a lower cost.

Additionally, securities-lending data provide direct evidence on the way speculators are able to spread their trades to reduce price impact. Around positive-demand additions, short interest and borrow fees rise around implementation and then decline as speculators cover their shorts. For negative-demand Russell migrations, the pattern reverses, with short interest falling around implementation. Finally, close-call Russell switchers have more trading before and after implementation than slam-dunk switchers, consistent with the idea that less predictable demand gives speculators less time to pre-position and leads to more concentrated trading.

These results have practical implications for how to interpret, and potentially reduce, the costs of index rebalancing. If speculators acquire shares in advance and sell them to index trackers at implementation, some of the effective trading cost is embedded in the pre-event price path. The model implies that post-event reversals are the right way to measure the *avoidable* component of price impact. Using this logic, the striking fact is not that index investors face trading costs. Any large trade must generate transaction costs in some form. The striking fact is that, in recent years, index funds appear able to trade enormous quantities – in some cases more than 20 times average daily volume – with costs of roughly 1% of the trade. Relative to standard price-impact estimates for institutional trades, this is surprisingly *cheap*, not surprisingly expensive. Apparent front-running can improve outcomes for index fund investors if competition between speculators is high enough, because it shifts risk to speculators that are willing to supply immediacy at implementation and allows index investors to minimize tracking error without incurring substantial price impact.

Using our framework, we also shed light on index design and the welfare implications of fast-track index additions after initial public offerings (IPOs). Index investors can rebalance their portfolio at small costs when the incoming constituent is very liquid, or when competitive speculators can intermediate the required shares by building inventory before implementation or shorting shares at implementation, covering the short positions afterward. In the case of fast-track additions, the incoming constituents are likely illiquid and hard to short. If the time from the IPO to the index addition is short, index investors will suffer from implementation-day price impact and post-event reversals, as empirically shown by Murray and Sammon (2026). In contrast, issuers and original owners benefit from fast-track additions, especially when competitive speculators bid up IPO prices.

REFERENCES

Admati, A. R. and Pfleiderer, P. (1988). A theory of intraday patterns: Volume and price variability. *Review of Financial Studies*, 1(1):3–40.

Almgren, R. and Chriss, N. (2001). Optimal execution of portfolio transactions. *Journal of Risk*, 3:5–40.

Arnott, R. D., Brightman, C., Kalesnik, V., and Wu, L. (2023). Earning alpha by avoiding the index rebalancing crowd. *Financial Analysts Journal*, 79(2):76–97.

Behmaram, A. and Davis, C. (2026). Indexing and the elasticity of stock demand.

Brunnermeier, M. K. and Pedersen, L. H. (2005). Predatory trading. *Journal of Finance*, 60(4):1825–1863.

Chinco, A. and Sammon, M. (2024). The passive ownership share is double what you think it is. *Journal of Financial Economics*, 157:103860.

Coles, J. L., Heath, D., and Ringgenberg, M. C. (2022). On index investing. *Journal of Financial Economics*, 145(3):665–683.

Cremers, K. M. and Petajisto, A. (2009). How active is your fund manager? a new measure that predicts performance. *Review of Financial Studies*, 22(9):3329–3365.

Frazzini, A., Israel, R., and Moskowitz, T. J. (2012). Trading costs of asset pricing anomalies. *Fama-Miller Working Paper, Chicago Booth Research Paper*, (14-05).

Gabaix, X. and Koijen, R. S. (2021). In search of the origins of financial fluctuations: The inelastic markets hypothesis. Technical report, National Bureau of Economic Research.

Greenwood, R. and Sammon, M. (2025). The disappearing index effect. *Journal of Finance*, 80(2):657–698.

Harris, L. and Gurel, E. (1986). Price and volume effects associated with changes in the S&P 500 list: New evidence for the existence of price pressures. *Journal of Finance*, 41(4):815–829.

Harvey, C. R., Mazzoleni, M. G., and Melone, A. (2025). The unintended consequences of rebalancing. Technical report, National Bureau of Economic Research.

Huberman, G. and Stanzl, W. (2004). Price manipulation and quasi-arbitrage. *Econometrica*, 72(4):1247–1275.

Koijen, R. S. and Yogo, M. (2019). A demand system approach to asset pricing. *Journal of Political Economy*, 127(4):1475–1515.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, pages 1315–1335.

Lo, A. W. (2004). The adaptive markets hypothesis: Market efficiency from an evolutionary perspective. *Journal of Portfolio Management, Forthcoming*.

Murray, C. and Sammon, M. (2026). Primary capital market transactions and index funds. *Review of Asset Pricing Studies*, page raaf013.

Pavlova, A. and Sikorskaya, T. (2023). Benchmarking intensity. *Review of Financial Studies*, 36(3):859–903.

Petajisto, A. (2011). The index premium and its hidden cost for index funds. *Journal of Empirical Finance*, 18(2):271–288.

Sammon, M. and Shim, J. J. (2026). Index rebalancing and stock market composition: Do indexes time the market? *Journal of Financial Economics*, 177:104229.

Shim, J. J. and Wang, Y. (2025). Commitment power and trading frequency in optimal execution. *Available at SSRN 4978304*.

Shleifer, A. (1986). Do demand curves for stocks slope down? *Journal of Finance*, 41(3):579–590.