Decoding Anomalies through Alpha Dynamics

Shuhao Ren

December 11, 2023

Abstract

This paper studies explanations of anomalies by analyzing how alphas of the characteristic-sorted portfolio *evolve* over months after the sorting period, referred to as the "alpha dynamic". In contrast, prior studies focus on the *average* of alphas after sorting. I develop new tests to statistically examine alpha dynamic patterns. I find that alpha dynamics provide new insights in evaluating whether anomalies (1) exist, (2) are profitable after considering trading costs, and (3) are likely due to mispricing or rational expectations. Upon incorporating the impacts of alpha dynamics into these questions, an analysis of 205 anomalies reveals that relying solely on t tests may miss many real anomalies. This problem becomes more severe with higher t cutoff values (e.g., 3.0). Also, the after-cost profitability has been significantly underestimated. Further, in about 60% of anomalies, the observed alpha dynamic pattern conforms to existing behavioral models rather than rational models. Examples of well-known categories include net share issuance, idiosyncratic volatility, and momentum.

1 Introduction

In the past few decades, researchers have reported hundreds of "anomalies".[1] Recently, there have been ongoing debates on interrelated explanations behind these anomalies. Key questions include: (1) Do the reported anomalies truly exist? (2) If they exist, would they be profitable after considering trading costs? (3) If they are profitable, are they due to mispricing or rational expectations?

For this paper, alpha is an average monthly excess return earned a specific number of months after a portfolio sort. For instance, if stocks are sorted only once a year, there would be twelve alphas associated with the portfolio sort. This paper examines whether the *evolution* of the alphas after sorting (henceforth, the *alpha dynamic*) provides new insights into these questions.

In contrast, most prior studies focus on the *average* of alphas after sorting. Here is a concrete example. Using the accruals as the characteristic, researchers sort firms in June every year and hold characteristic-sorted portfolios from July to the following June. Then they calculate monthly returns in excess of a given factor model's predictions and average these returns over all months. This method is equivalent to first calculating the alphas for each month following the sorting period and then *averaging* the alphas (henceforth, the *alpha mean*) over the first twelve months after sorting.[2]

While studying the alpha mean is useful, it does not reveal how alphas evolve over time after sorting. That is, the alpha earned (say) one month after sorting is blended with the alpha earned (say) six months after sorting. Are these two alphas the same? Theoretically, should they be? It is an open question whether or not the alpha dynamic sheds new light on anomalies. This paper is also different from the studies that examine whether alphas (or

[1] If a firm characteristic's return predictability on future stock returns cannot be fully explained by an asset pricing model, it is considered an anomaly relative to that model.

[2] Since alpha is an arithmetic average, the alpha mean is also an average. Additionally, when the alpha in one particular month after sorting is examined, the alpha mean is just that alpha itself. The term, "alpha mean" highlights that it is an average rather than a dynamic.

betas) vary over calendar time or other events.[3] Unlike previous papers, I study anomaly explanations by analyzing how alphas vary in event time following the portfolio sorting date for a large number of important anomalies.

To study the impact of alpha dynamics on anomaly explanations, I examine 205 published anomalies. I study their alpha dynamics by jointly investigating the alpha estimates over different months after sorting. I develop new empirical tests that allow me to statistically examine whether alphas are constant in the months after sorting and whether they exhibit certain patterns that can be implied by several existing economic models.[4] To detect different patterns, I analyze not only the alphas within the first twelve months, but also those at longer horizons. My main finding is that examining the patterns of alpha dynamics can provide new and important insights into all three questions listed at the beginning.

First, regarding the existence of anomalies, the alpha dynamic can detect non-zero alphas when the alpha mean may not. For example, a build-up of mispricing followed by a correction could imply positive alphas for a period of time followed by negative alphas. In this case, alphas would be non-zero, even while the alpha mean could be indistinguishable from zero. The tests that I propose involving the alpha dynamic can alleviate problems such as this.[5] Prior studies use alpha-mean tests (e.g., t tests) and argue that many anomalies do not exist if we change slightly the portfolio construction method or adjust the statistical threshold to address p-hacking concerns (Harvey, Liu, & Zhu, 2016; Hou, Xue, & Zhang, 2020).[6] I find that relying solely on alpha-mean tests may miss many real anomalies. Further, adjusting the statistical thresholds (e.g., raising t cutoff to 3.0) may miss many more real anomalies.

Second, I find that the appropriate holding period to evaluate after-cost profitability should be determined by both alpha dynamics and trading costs. Based on this rule, I provide

[3] For calendar time, see e.g., Boguth, Carlson, Fisher, and Simutin (2011) and Lewellen and Nagel (2006). For other events, McLean and Pontiff (2016) study publication dates and Engelberg, McLean, and Pontiff (2018) study corporate news and earnings announcement days.

[4] Alphas are considered constant if the difference between any two adjacent alphas after sorting is statistically insignificant from zero.

[5] I discuss other problems of alpha-mean tests and how alpha-dynamic tests alleviate them in Section 2.2.

[6] p-Hacking occurs when large t statistics result from searching for significant results among numerous meaningless characteristics.

a new and simple method to determine holding periods. I find that this method might be more appropriate to evaluate after-cost profitability and after-cost profitability may have been significantly underestimated. When compared to prior methods,[7] this method achieves a statistically significant improvement in after-cost alpha for about 20% of anomalies. And the improvement is about 0.3% per month across these improved anomalies.

Third, I observe that about 60% of the anomalies exhibit an alpha dynamic pattern that conforms to existing behavioral models rather than rational models. This suggests that they are likely to be at least partially due to mispricing. Examples of well-known categories include net share issuance, idiosyncratic volatility, and momentum.

Let me explain the simple economic rationale behind the first finding about detecting non-zero alphas. If anomalies do not exist, true alphas should be zero in any month after sorting.[8] This implies a null hypothesis that the alpha mean is zero and alphas are constant within any subset of months after sorting. Therefore, either a non-constant alpha dynamic or a non-zero alpha mean rejects the null and indicates the existence of non-zero alphas. This suggests that we can also use alpha dynamics to detect non-zero alphas.

Why can alphas be non-constant? Both rational and behavioral theories can imply non-constant alphas. For instance, Keloharju, Linnainmaa, and Nyberg (2021) show that several production-based models such as Berk, Green, and Naik (1999) can imply a monotone pattern of alphas after sorting because firm risks converge over time in those models. For behavioral theories, the changes in mispricing over time will be reflected as alphas. When the changes are not constant, alphas will also be non-constant. Models such as return extrapolation (e.g., Barberis, 2018) can imply not only monotone patterns of alphas but also a ripple pattern, characterized by alternating increases and decreases in alphas over time.[9]

[7] Researchers usually use one-month or twelve-month holding periods (A. Y. Chen & Velikov, 2023; A. Y. Chen & Zimmermann, 2021). Novy-Marx and Velikov (2016) determine holding periods by turnover rates.

[8] The true alpha represents the alpha in the population.

[9] More details are in Section 2.1. Furthermore, behavioral models usually assume a shock date when mispricing initially arises and study how alphas or prices evolve after the shock date. In contrast, I study portfolio sorting dates when firms are sorted based on a specific firm characteristic. Empirically, studying sorting dates is a tradition (e.g., van Binsbergen, Boons, Opp, & Tamoni, 2023). Economically, in Appendix B, I discuss more how shock dates and sorting dates might be related.

Alphas only exist relative to a particular benchmark asset pricing model (Fama, 1970). Rather than attempt to determine the correct benchmark model, I explore the new insights from alpha dynamics given an asset pricing model. While I use the CAPM to identify alpha dynamics, the methodology can be applied to other asset pricing models as well.[10]

I next explain why the holding period to evaluate after-cost profitability should be determined by alpha dynamics and trading costs. Since the alpha mean is an average of alphas, holding periods will affect the alpha mean when alphas are time-varying after sorting. In the accruals example, if alphas decay over time after the sort, holding the accruals portfolio longer results in a smaller alpha mean. In the meantime, a longer holding period also reduces portfolio turnover, thus lowering rebalancing costs. This leads to a trade-off. In a model with exponential decay of alphas over time, I find that the appropriate holding period should decrease with the initial size and decay rate of alphas and increase with trading costs.

Lastly, we are interested in whether alphas are due to mispricing or rational expectations. Existing behavioral explanations can imply a ripple pattern. I calibrate and show that models such as return extrapolation (Barberis, 2018) and inattention (Duffie, 2010) can imply a ripple pattern. For example, the intuition behind Barberis (2018) is that extrapolators overreact to both negative and positive returns. This results in an alternating build-up and resolution of overpricing and underpricing over time.[11] In contrast, existing rational explanations, to the best of my knowledge, do not imply such a pattern.

Empirically, I start by developing two new tests to statistically examine whether alphas are non-constant to determine if non-zero alphas exist. As some models can imply monotone patterns of alphas, the first test extends the monotonicity test by Paton and Timmermann (2010) and examines whether alphas exhibit an increasing or decreasing pattern.[12] Further, as some models can imply non-constant and non-monotone patterns such as the ripple pattern, I develop the second test. The intuition of the test is that if alphas are constant, the

[10]I use the CAPM as the baseline model because those theories are proposed relative to the CAPM.

[11]I explain more intuition behind the ripple pattern in those models in Section 2.1.3.

[12]Patton and Timmermann (2010) apply their method to examine the monotonicity relation in the cross-section. In contrast, I examine the monotonicity relation in the time series (event time after sorting).

timing to form a portfolio after sorting does not affect the alpha mean. If not, it implies that alphas are non-constant.

To compare with alpha-mean tests, I follow Hou et al. (2020) and construct value-weighted portfolios. Within a sample period (full-sample, before-sample, in-sample, and post-sample periods), I label those that fail the t test with a cutoff of 1.96 as suspicious anomalies. Among the remaining anomalies, I label those that fail the t test with a cutoff of 3.0 as HLZ suspicious anomalies. Similar to Hou et al. (2020), many anomalies do not exist based solely on t tests. However, in the full sample period, 21% of suspicious anomalies and 59% of HLZ suspicious anomalies may have non-zero alphas as they pass at least one of the two alpha-dynamic tests. In the post-sample periods, these numbers are 25% and 63%. Results suggest that alpha-mean tests alone may miss many real anomalies, and this problem becomes more severe as the statistical threshold rises. I further find that alpha dynamics remain useful even when several popular multi-factor models are used. Moreover, results cannot be explained by seasonality (e.g., Heston & Sadka, 2008), since I conduct portfolio sorts every month and track the buy-and-hold returns of each portfolio over time.

To evaluate after-cost profitability, I estimate the optimal holding period that maximizes the after-cost alpha over time.[13] To alleviate look-ahead bias and data mining concerns, I estimate optimal holding periods using known data but evaluate after-cost alpha based on out-of-sample returns. Compared to various prior methods, this new method improves after-cost alpha for a similar proportion of anomalies with a similar magnitude of improvement. With Monte Carlo simulations, I find these results cannot be generated by random variation.

To identify the ripple pattern, I analyze alphas over nine years after sorting. I divide the nine-year period into five consecutive unconnected subperiods. I introduce a new test to statistically assess the presence of either increasing, decreasing, or both patterns of alphas within each subperiod. When summarizing the results from the five subperiods, I apply

[13]While other elements such as characteristic frequencies and cost mitigation techniques may affect the after-cost profitability, the main focus of this analysis is to examine how holding periods impact after-cost profitability and the role of alpha dynamics and trading costs in determining these holding periods.

multiple comparison corrections. The ripple pattern is observed when both increasing and decreasing patterns of alphas are present in the nine-year period. Out of the 132 anomalies showing evidence of non-zero alphas in the full sample period, 79 (60%) exhibit the ripple pattern. Additionally, 41% of anomalies that exhibit the ripple pattern in the in-sample period do not exhibit it in the post-sample period.

Overall, alpha dynamics can provide new and important insights for multiple anomaly explanations. This implies that we should study the alpha dynamic alongside the alpha mean to better understand anomalies.

1.1 Related literature

This paper is first related to studies on explanations of anomalies. Most prior studies focus on the *alpha mean*: the existence of non-zero alphas (e.g., A. Y. Chen, 2021; A. Y. Chen & Zimmermann, 2020; Chordia, Goyal, & Saretto, 2020; Harvey, 2017; Harvey et al., 2016; Hou et al., 2020; Jacobs & Müller, 2020; Jensen, Kelly, & Pedersen, 2022; Kelly, Pruitt, & Su, 2019; Linnainmaa & Roberts, 2018; Martin & Nagel, 2021), after-trading-cost profitability (e.g., A. Y. Chen & Velikov, 2023; Novy-Marx & Velikov, 2016, 2019), and differentiating between mispricing and rational expectations (e.g., Engelberg et al., 2018; McLean & Pontiff, 2016). This paper complements this literature by showing that *alpha dynamics* can provide new and important insights into these anomaly explanations. McLean and Pontiff (2016, MP) compare the *alpha mean* before and after publication dates. In contrast, *alpha-dynamic* tests do not require two sample periods and can be applied to any period. While MP have not studied alpha explanations in the post-sample period, this paper finds that many anomalies exhibit alpha dynamics even in the post-sample period.

An exception is Bowles, Reed, Ringgenberg, and Thornock (2023, BRRT). They also study the existence of anomalies from the evolution of alphas. They use a unique database that has the precise timing of the release of accounting variables. They use this precise timing to construct portfolios and find a large improvement in performance. Unlike their study, the tests in this paper are different and more general in the way that they do not need access to a unique proprietary database or are restricted to accounting anomalies. Furthermore, while BBRT focus only on the existence of anomalies, this paper additionally investigates the after-cost profitability and differentiates between mispricing and rational expectations.

Studies of return momentum and reversal also examine alpha dynamics in spirit (e.g., De Bondt & Thaler, 1985; Jegadeesh & Titman, 1993). This paper examines not only momentum and reversal, but also anomalies from other categories. Perhaps more importantly, the methodology employed in this paper can potentially be applied to any anomaly.

This paper is also related to the growing literature on returns and prices multiple periods after the portfolio sorting period (e.g., Baba Yara, Boons, & Tamoni, 2020; Bessembinder, Cooper, & Zhang, 2021; Y. Chen & Kaniel, 2021; Chernov, Lochstoer, & Lundeby, 2022; Cho & Polk, 2019; Favero, Melone, & Tamoni, 2019; Hendershott, Menkveld, Praz, & Seasholes, 2022; Keloharju et al., 2021; van Binsbergen et al., 2023; van Binsbergen & Opp, 2019). This paper complements the literature by exploring questions regarding the existence, after-cost profitability, and differentiation between mispricing and rational expectations of anomalies. Baba Yara et al. (2020) study the implications of new and old sorts for factor models. van Binsbergen et al. (2023) classify anomalies into build-up and resolution anomalies. While they examine patterns of prices, I examine patterns of alphas. Further, they have not studied any of the three questions I examine.

2 Intuition

In this section, I use a simple model to provide economic intuition on why alpha dynamics provide new information in understanding various explanations of anomalies.

2.1 Motivation: patterns of alpha dynamics

Before discussing how alpha dynamics affect different anomaly explanations, let me first explain what patterns of alpha dynamics can be expected under different explanations of anomalies.

Let j be the number of months that have passed after portfolio sorting. Then j represents event time and event time 0 is the month of portfolio sorting. Denote α_j as the true alpha in event time j: the average of returns in excess of the factor-model prediction in the population. I omit the subscripts for characteristic X and time t in α_j to ease exposition, but it can be different both across characteristics and across time. In all that follows, I focus on alpha with respect to the market factor.

2.1.1 True alphas are zero

The first situation is when true alpha is zero in every period after sorting. That is, $\alpha_j = 0$ for all j. As shown in Figure 1a, within any subset of months after sorting, the alpha dynamic is *constant*.

2.1.2 Rational expectations

If true alphas do exist, one explanation is rational expectations. That is, the factor model used to assess anomalies may omit rational risk factors. The production-based asset pricing literature suggests that α_j can have a *non-constant* pattern. For example, in the models of Gomes, Kogan, and Zhang (2003) and Zhang (2005), production risk is mean-reverting. These models imply that α_j can have a strict decreasing function since high-risk firms today will become less risky in the future and low-risk firms today will become riskier in the future. Moreover, Keloharju et al. (2021) show that production-based models in Berk et al. (1999), Gomes et al. (2003), Hackbarth and Johnson (2015), and Zhang (2005) all imply a downward-sloping of alphas.[14]

2.1.3 Mispricing

Alternatively, alphas can be due to mispricing. In the behavioral literature, persistent mispricing may exist due to the limits of arbitrage. Mispricing can be an arbitrage opportunity that does not involve risk. It can also be collective and result from correlated errors in investor expectations (e.g., Barberis, Greenwood, Jin, & Shleifer, 2015; Stambaugh & Yuan, 2017). The corrections or build-ups of mispricing will be reflected as alphas. Therefore, if the rate of correction or build-up of mispricing is not constant over time, α_j will also have a *non-constant* pattern.

What non-constant patterns of alpha dynamics can be implied in behavioral models? To explore this, I present below the mechanisms of return extrapolation (Barberis, 2018) and inattention (Duffie, 2010; Hendershott et al., 2022). I also calibrate and plot the alpha dynamic implied by these models in Figure 2, using the parameters from the original papers.

Model 1: Barberis (2018):[15] **return extrapolation**. There are two types of investors, extrapolators and arbitrageurs. Extrapolators' belief on future price change is a weighted average of past price changes. And arbitrageurs have bounded rationality in the way that they do not have a full understanding of extrapolator demand. Extrapolators push the price away from the fundamental price and arbitrageurs drag the price back to the fundamental price.

Model 2: Duffie (2010): **fixed periods of inattention**. When there is a supply shock, only a few investors (all attentive investors and part of inattentive investors) can absorb the shock. Therefore, there is a large price recession to compensate investors who absorb the shock. These investors then lay off the risk over time when other inattentive investors

[14] Keloharju et al. (2021) describe the mechanisms of these papers in Section 2.1 and show the plots of the mechanisms in Figure 2.

[15] In Section 4.1 of the paper. The mechanism is similar to that in Hong and Stein (1999). Compared to the framework in Hong and Stein (1999), Barberis (2018) is more general as it models a well-known behavioral bias of return extrapolation that could persistently exist in the market.

come to the market and the price reverts over time. The model has one class of inattentive investors and they fix their periods of inattention.

Model 3: Hendershott et al. (2022): **stochastic arrival of inattentive investors**. The model has market makers, attentive investors, and multiple classes of inattentive investors[16]. The inattentive investors arrive stochastically. Each class of investors has random private-value shocks each period. And the shocks between attentive and inattentive investors can be perfectly canceled off. However, only part of inattentive investors will adjust their portfolios after the shocks. Therefore, attentive investors are compensated with a price recession to absorb the shocks of inattentive investors and lay off the risk when other inattentive investors come to adjust their portfolios and price reverts over time.

Based on these mechanisms, what patterns of alpha dynamics can be expected? First, alpha dynamics can have a downward-sloping pattern. In all mechanisms, there is a downward-sloping pattern of alpha dynamic within the first few periods after the initial shock. Since alphas represent changes in mispricing, this implies that the changes are becoming slower over time. Take the mechanism of Barberis (2018) as an example. There is a large positive cash flow shock in period 0 (and vice versa). As shown in Figure 2a, the decreasing positive alphas from period 1 to period 2 correspond to the build-up of overpricing because extrapolators overreact to the shock. As arbitrageurs keep pulling the price back to the fundamental price, and return extrapolators adjust their belief over time, the build-up rate slows down over time and eventually reaches zero. As a result, alphas have a decreasing pattern.

Moreover, the mechanisms of Barberis (2018) and Duffie (2010) can imply an upward-sloping pattern of alpha dynamics. Still, take the example of Barberis (2018). As shown in Figure 2a, alphas increase from period 2 to period 4 if we flip the sign of the negative alphas (same pattern from period 7 to period 9). This subperiod corresponds to the resolution of the overpricing. Initially, the rate of resolution is slow because extrapolators are still on the other side of arbitrageurs' trades. As extrapolators adjust their belief over time, the

[16] To provide the intuition of how inattention affects the pattern of alphas, I assume one class of inattentive investors when I simulate this model.

resolution rate becomes increases. The long-term reversal and value anomalies are examples of this resolution mechanism.

Furthermore, the mechanisms of Barberis (2018) and Duffie (2010) can imply a "ripple" pattern: increasing and decreasing alphas with a decaying magnitude. In Barberis (2018), the build-up and resolution of mispricing lead to opposite signs of alphas. Also, extrapolators overreact to both positive returns and negative returns, leading to alternating underpricing and overpricing. Therefore, prices jump above and below the fundamental price alternatively over time, resulting in a ripple pattern. And in Duffie (2010), the ripple pattern appears because some inattentive investors only adjust their portfolios every few periods. Therefore, prices surge above and plunge below their steady-state level over time due to imbalances in demand and supply. Overall, the ripple pattern is caused by investors' biased extrapolated beliefs or inattention.

2.2 Intuition for the tests on the existence of alphas

Section 2.1 shows that α_j remains constant when true alphas are zero. In contrast, α_j can be non-constant when non-zero alphas exist. This difference in patterns of α_j can be then used to test whether non-zero alphas exist.

2.2.1 On-paper alphas

With a specific combination of the number of skipped months after portfolio sorting (k) and holding period (h), the monthly true on-paper alpha of characteristic-sorted portfolios is given by:

$$\alpha_{k,h}^{op} = \frac{1}{h} \sum_{j=k+1}^{k+h} \alpha_j \tag{1}$$

Here, I use the traditional overlapping portfolio approach (e.g., Jegadeesh & Titman, 1993). Therefore, in each month, there will be h portfolios that are sorted from $k + h - 1$ to

k months ago. Each portfolio has a weight of $\frac{1}{h}$. The portfolios are rebalanced monthly, so there is an overlap in returns when $h > 1$. To deal with this overlap, I use a calendar-time portfolio approach to calculate average monthly returns.

Lemma 1 *When α_j remains constant in j, $\alpha^{op}_{k,h}$ is constant in k and h. And when α_j is non-constant in j, $\alpha^{op}_{k,h}$ is not constant in k and h.*

Lemma 1 is obvious based on Eq. 1. This implies that testing whether the alpha dynamic α_j is constant is the same as testing whether the on-paper alpha $\alpha^{op}_{k,h}$ is constant in k and h. The two tests in the next two sections are on the pattern of $\alpha^{op}_{k,h}$ in k and h based on Lemma 1.

2.2.2 The monotonicity test

The first test examines whether the alpha dynamic has a strictly increasing or strictly decreasing pattern within a subset of months after sorting. Testing monotonicity patterns is motivated by both rational production-based models (Keloharju et al., 2021) and behavioral models as discussed in Section 2.1. Specifically, the monotonicity test examines the following hypotheses:

$$\begin{aligned} &H0 : \alpha^{\mathrm{op}}_{\mathrm{0,h}} \text{ has a constant pattern in h.} \\ &H1 : \alpha^{\mathrm{op}}_{\mathrm{0,h}} \text{ has a strictly increasing or decreasing pattern in h.} \end{aligned} \tag{2}$$

If $\alpha^{op}_{0,h}$ is shown to have a strictly increasing or decreasing pattern, alphas exist. The reason to examine how on-paper alphas change in h is that this addresses the issue that we may not be able to directly observe when alphas disappear. For example, assume that alphas strictly decrease over time and disappear after 20 months. Suppose that we examine alphas in the first 60 months after sorting in the test. Then on-paper alphas will be zero for $k > 19$. Therefore, using k alone will fail to identify a strictly decreasing pattern. However, using h will be able to identify a strictly decreasing pattern since on-paper alphas decrease in h even when $h > 19$.

2.2.3 The optimization test

Non-constant patterns are not limited to monotone patterns. Instead, alphas may show other types of non-constant patterns.[17] The second test is designed to identify a more general non-constant pattern of alpha dynamics. Further, by conducting both tests, the issue of a lack of statistical power in each test is reduced.

According to Lemma 1, when α_j is constant in j, $\alpha^{op}_{k,h}$ is the same for all strategies based on k and h. In contrast, when α_j is non-constant, Let α_{j^*} be a maximum. To maximize the on-paper alpha, the problem is:

$$\max_{k,h} \ \alpha^{op}_{k,h} = \max_{k,h} \ \frac{1}{h} \sum_{j=k+1}^{k+h} \alpha_j \tag{3}$$

Proposition 1 *The optimal strategy to obtain the highest* $\alpha^{op}_{k,h}$ *is* $k^* = j^* - 1$ *and* $h^* = 1$.

Proof: See Appendix C.

Further, this optimal strategy also maximizes the wealth of investors that have a long investment horizon as well. This is because this strategy is optimal in any month theoretically.

Proposition 1 shows that there is at least one strategy (the optimal strategy) that can generate a higher on-paper alpha than the average of the on-paper alphas of alternative strategies on k and h when the alpha dynamic is non-constant:

[17] One example is the ripple pattern discussed in Section 2.1.

$$\begin{aligned}
&H0: \text{All strategies in k generate the same } \alpha^{\text{op}}_{\text{k},1}. \\
&H1: \text{There is a strategy in k that can generate a } \alpha^{\text{op}}_{\text{k},1} \text{ higher than the average of the } \alpha^{\text{op}}_{\text{k},1} \\
&\quad \text{of alternative strategies in k.}
\end{aligned} \tag{4}$$

It is important to note that $\alpha^{op}_{k,1}$ here is the true on-paper alpha in population rather than an estimate from a finite sample. This hypothesis design is to address two empirical challenges. First, we do not directly observe the true optimal k^* that generates the highest $\alpha^{op}_{k^*,1}$. However, the rejection of the hypothesis only requires that one strategy generates a higher alpha than the average of the alternatives. Empirically, I estimate the optimal k^* based on past known information before a month t and record the return of this optimization strategy in $t+1$. This alleviates look-ahead bias and data mining concerns. Then I examine whether the returns of this strategy in $t+1$ generate an alpha higher than the average of the alphas of alternative strategies in k.

The second challenge is that we cannot compare the returns of the optimization strategy with the returns of a strategy based on an arbitrary k, as we do not observe which k generates the highest true on-paper alpha. This challenge is addressed by comparing the alpha of the optimization strategy with the average of all alternative strategies. The intuition is that the optimization strategy should produce a higher alpha unless the true alphas are all the same.

Furthermore, alpha-dynamic tests alleviate some of the potential problems associated with alpha-mean tests. First, as in the build-up and resolution of the mispricing example discussed in Section 1, the choice of holding periods in portfolio construction often influences the results of alpha-mean tests and may lead to conflicting conclusions.[18] Given that the alpha dynamic is non-constant in that example, alpha-dynamic tests could serve to identify non-zero alphas. Moreover, alpha-dynamic tests study alphas jointly, thus alleviating the

[18] See Bessembinder, Burt, and Hrdlicka (2022), Hasler (2022), and Hou et al. (2020).

data mining concerns in the choice of holding periods in alpha-mean tests. Additionally, failing alpha-mean tests could be due to low statistical power. This problem would be more significant for value-weighting portfolios and in the post-publication period due to typically small alphas. However, alpha-dynamic tests may not encounter this issue simultaneously, as their statistical power depends on the differences between alphas after sorting and the volatility of these differences.

2.3 Intuition for the impact on after-cost profitability

It does not challenge market efficiency if characteristic-sorted portfolios generate alphas on paper, but agents cannot actually trade profitably because of transaction costs. To evaluate after-cost profitability, we need to choose an appropriate holding period. While prior studies use ad hoc holding periods or determine the holding period by the turnover rates, this section shows the intuition of why we should jointly consider alpha dynamic and trading costs to determine the holding period.

2.3.1 After-cost alphas

Let c represent the population mean of monthly rebalancing costs (in percentage). Further, assume c is exogenous. c can also be different across characteristics X and across time t. The monthly true after-cost alpha for a combination of k and h is then given by:

$$\alpha_{k,h}^{ac} = \frac{1}{h}\sum_{j=k+1}^{k+h} \alpha_j - \frac{c}{h} = \alpha_{k,h}^{op} - \frac{c}{h} \tag{5}$$

Based on Lemma 1, when the alpha dynamic α_j is constant, holding periods do not affect on-paper alpha $\alpha_{k,h}^{op}$. And holding periods will affect $\alpha_{k,h}^{op}$ when α_j is non-constant.

The appropriate k and h to evaluate after-cost profitability should be the optimal k^* and h^* that maximizes the $\alpha_{k,h}^{ac}$. Next, I will show how k^* and h^* are affected by alpha dynamics and trading costs.

2.3.2 Exponential decay in alphas

I model the behavior of alphas the same as that in Hendershott et al. (2022, HMPS). In HMPS, alphas are due to mispricing.[19] A shock happens each month t. Denote initial mispricing due to the shock as $-\delta_0$.[20] Denote the expected mispricing j months after the shock as P_j (in percentage). P_j is assumed to follow an Ornstein-Uhlenbeck (OU) process. Therefore, P_j follows an exponential decay and reverts to zero over time. In equilibrium, the decay rate is the inattentiveness of inattentive investors, λ ($\lambda > 0$). Then P_j can be expressed as:

$$P_j = -\delta_0 e^{-\lambda j} \qquad \text{for} \quad j \geq 0 \tag{6}$$

Let $\delta_0 > 0$.[21] Alphas are changes in P_j. Thus, the evolution of alphas can be expressed as:

$$\alpha_j = P_j - P_{j-1} = \delta_0(1 - e^{-\lambda})e^{-\lambda(j-1)} \qquad \text{for} \quad j \geq 1 \tag{7}$$

Eq. 7 indicates that alphas also reverts to zero over time.

2.3.3 Optimization

To maximize the after-cost alpha, the problem is:

$$\begin{aligned} \max_{k,h}\ \alpha_{k,h}^{ac} &= \max_{k,h}\ \frac{1}{h}\sum_{j=k+1}^{k+h}\alpha_j - \frac{c}{h} \\ &= \max_{k,h}\ \frac{1}{h}\delta_0 e^{-\lambda k}(1 - e^{-\lambda h}) - \frac{c}{h} \end{aligned} \tag{8}$$

Since $\alpha_{k,h}^{ac}$ decreases in k, $k^* = 0$. Replace $k = 0$ and take FOC on h:

[19] The following results can also apply to alphas resulting from rational expectations if omitted risk premia follow the same behavior as in Eq 7.

[20] This is the $-pG_t$ in Equation 12 of HMPS. In HMPS, the absolute value of δ_0 decreases in the fraction of attentive investors (m_F in HMPS) and increases in inattentiveness.

[21] The sign of δ_0 does not affect the conclusion. We can flip the sign of long-short portfolios so that alphas are positive when $\delta_0 < 0$.

$$foc = \frac{\delta_0(\lambda h + 1)e^{-\lambda h} + c - \delta_0}{h^2} \quad (9)$$

Since alphas decay over time by assumption, if α_1 is less than c, $\alpha^{ac}_{k,h}$ will be negative for any combination of k and h. Therefore, let us assume that $\alpha_1 > c$. Then $\delta_0(1 - e^{-\lambda}) > c$ according to Eq. 7 and $\delta_0 > c$. Since $\delta_0(\lambda h + 1)e^{-\lambda h}$ converges in h to zero and $\delta_0 > c$, foc cannot be always positive. Thus, there could be only two cases.

Proposition 2 *Case 1: When* $\frac{\lambda+1}{e^{\lambda}} \leq 1 - \frac{c}{\delta_0}$*, the optimal holding period* $h^* = 1$*.*

Proof: See Appendix C.

The intuition is that when trading costs c are tiny relative to λ or δ_0, trading costs can be ignored, and the problem is similar to the one for the on-paper alpha. Therefore, similar to Proposition 1, the optimal holding period is $h^* = 1$.

Proposition 3 *Case 2: When* $\frac{\lambda+1}{e^{\lambda}} > 1 - \frac{c}{\delta_0}$*, the optimal holding period* h^* *satisfies* $\frac{\lambda h^*+1}{e^{\lambda h^*}} = 1 - \frac{c}{\delta_0}$*. And* h^* *(1) decreases in* δ_0*, (2) decreases in* λ*, and (3) increases in* c*.*

Proof: See Appendix C.

Proposition 3 indicates that the closed-form solution does not exist when $\frac{\lambda+1}{e^{\lambda}} > 1 - \frac{c}{\delta_0}$,[22] but we can know how δ_0, λ, and c affect h^*. The intuition for (1) is that when initial mispricing δ_0 is larger, the decay in alphas will also be larger over time when everything else is the same. Then having a longer holding period will reduce the on-paper alpha and thereby reduce the after-cost alpha. The intuition behind (2) is similar. If the attentiveness or decay rate λ is greater, alphas decay more quickly. Finally, (3) is because when trading costs c are higher, a longer holding period can reduce rebalancing costs.

Overall, analysis indicates that when alphas decay exponentially over time, $k^* = 0$ and h^* depends on the initial mispricing level δ_0, rate of reversion of mispricing λ, and trading

[22] It is a Lambert's W function.

costs c. Prior studies use ad-hoc holding periods or turnover rates to determine holding periods. This may lead to a large underestimation of after-cost profitability.

Is k^* always zero? The assumption that alphas decay monotonically is motivated by the existing models. In reality, the alpha dynamic may also have other patterns. For example, when the alpha in the first period (α_1) is not the maximum among all possible α_j, Proposition 1 suggests that the optimal k to maximize the on-paper alpha will be greater than zero. Then the optimal k for after-cost alpha can also be different from zero.[23] Although such a scenario is possible, I do not find any model that implies such a pattern of alpha dynamics, making it difficult to model the behavior of alphas. Instead, I empirically examine whether k matters for after-cost profitability in Section 4.2.

2.4 Intuition for differentiating mispricing from rational expectations

Patterns of alpha dynamics provide new moments to evaluate asset pricing models. Among models that explain an anomaly, if behavioral models can imply a pattern that cannot be observed in rational models, such a pattern can be used to differentiate between existing behavioral and rational models.

As discussed in Section 2.1, existing behavioral models such as Barberis (2018) and Duffie (2010) feature a "ripple" pattern in the alpha dynamic. The ripple pattern is a pattern that contains both increasing and decreasing patterns in the alpha dynamic. As far as I am aware, existing rational models may imply a monotone pattern of alphas, but not a non-monotone ripple pattern. For example, production-based models (e.g., Zhang, 2005) imply a monotonically decreasing pattern of alpha dynamic as firms' production risk is mean-reversed. Moreover, although the risk of a firm may go up and down over time due to *random variation*, a rational model needs to explain why the risk of a firm goes up and down over time *on average* after portfolio sort dates.

[23]For example, when trading costs are extremely low, the effect of trading costs can be ignored.

Moreover, the mechanisms of return extrapolation (Barberis, 2018) and inattention (Duffie, 2010) are also general. As a result, they might be the underlying mechanism for any anomaly.

3 Data

3.1 Characteristic-sorted portfolios

I collect 205 firm or return-based characteristics from the website[24]. Authors of A. Y. Chen and Zimmermann (2021) create the website and kindly provide the data. The data is monthly and ranges from December 1925 to June 2022. Stock characteristics are taken from the Center for Research in Security Prices (CRSP) database. Returns of the factors are from Kenneth French's website and the q-factor library[25]. All firms that are listed on the NYSE, AMEX, and NASDAQ are included.

For each characteristic, I sort stocks into decile groups at the end of each month. Then I construct value-weighted long-short spread portfolios.[26] Then I track the monthly buy-and-hold returns of the spread portfolios. When a stock delists, I reinvest the amount of money in the stock (net of the delisting return) to the rest of the stocks in the portfolio with value weighting.

Finally, I investigate four periods, full-sample, before-sample, in-sample, and post-sample periods.

3.2 Trading cost measures

Stock-level trading cost measures are collected from Andrew Chen's website. A. Y. Chen and Velikov (2023) use these cost measures to estimate the profit after trading costs for each characteristic. They estimate effective spreads as measures of trading costs and argue that

[24] https://www.openassetpricing.com/

[25] http://global-q.org/factors.html

[26] A. Y. Chen and Zimmermann (2021) sign the characteristics so that the average return of the spread portfolios is positive within the sample period of the original paper. I use the same signs as A. Y. Chen and Zimmermann (2021).

their trading cost aims to measure the minimum amount by which prices would have been moved. Effective spreads are measured as twice the absolute difference between the midpoint of bid-ask spreads and the executed prices. The effective spread measures include one high-frequency (HF) measure and three low-frequency (LF) measures. Following A. Y. Chen and Velikov (2023), I employ half HF effective spreads as trading costs whenever possible and take half of the average of the three LF spreads as trading costs when HF spreads are missing.

Upon merging the trading costs data with stock returns data, I find that about 12% of the observations contain missing trading costs. Since I'm interested in whether portfolio construction choices will affect after-cost returns that consider trading costs, I will need trading cost estimates for any stock in any month. Otherwise, the estimate of portfolio rebalancing costs is imprecise. Therefore, I fill the missing trading cost of a stock in month t in the following sequence: If the trading cost of the stock in month $t-1$ is available, I use it to fill in the trading cost for the stock in month t. If the trading cost in the previous month is not available, I then sort stocks into deciles based on their firm size in month t. I use the average of the trading costs of the decile the stock is in month t to fill in the missing trading cost. Last, if the firm size of the stock is missing in month t, then I use the average of the trading costs of all stocks in month t to fill in the missing trading cost.

In Appendix E, I plot how effective spreads vary across time and stocks in Fig E1. I also plot the portfolio-level distribution of rebalancing costs and turnover in Fig E2 and Fig E3. These figures show that trading costs vary significantly across stocks, portfolios, and time.

4 Empirical design and results

In this section, I empirically test the impact of alpha dynamics on the explanations of anomalies based on the intuition discussed in Section 2.

4.1 Existence of alphas

4.1.1 Motivating evidence

The traditional method to test the existence of non-zero alphas is conducting t tests to investigate whether the alpha mean is significantly different from zero. To examine whether there is an additional contribution from the new alpha-dynamic tests, I first conduct the traditional t tests on 205 characteristic-sorted portfolios.

For each characteristic, I construct value-weighted long-short decile portfolios. Portfolios are then held for one month and rebalanced. The use of value weighting to evaluate anomalies is also used in Hou et al. (2020). Within a sample period (full-sample, before-sample, in-sample, post-sample periods), I label anomalies that fail the t test with a cutoff of 1.96 as suspicious anomalies. Moreover, in light of potential p-hacking, Harvey et al. (2016) recommend increasing the t-cutoff to 3.0. I label the anomalies that fail the t test with a cutoff of 3.0 as HLZ suspicious anomalies. The remaining anomalies are labeled as robust.

In the full sample period, I find that there are 92 suspicious, 41 HLZ suspicious, and 72 robust anomalies. The failure rate based on t tests is similar to Harvey et al. (2016) and Hou et al. (2020). Recently, researchers have also begun to worry that the portfolio construction method adopted by the original authors may be subject to data mining (Hasler, 2022). These studies indicate many anomalies may not truly exist.

Conversely, instead of using the same portfolio construction method for each anomaly, A. Y. Chen and Zimmermann (2021) use the portfolio construction method adopted by the original authors.[27] They find a 0% failure rate for this same set of 205 characteristics. Their study indicates that all these anomalies exist.

Overall, it remains uncertain whether alphas exist based solely on t tests from prior studies due to potential problems of t tests. Different portfolio construction methods and t cutoffs can affect t test outcomes, and low statistical power could also contribute to failing a t test. Motivated by these potential problems of t tests and the potential that alpha

[27] The construction decisions include the holding period, listing exchanges, the weighting scheme, etc.

dynamics can complement t tests as discussed in Section 2, I proceed to examine whether the monotonicity and the optimization test developed in Eq. 2 and 4 complement t tests in evaluating the existence of non-zero alphas.

4.1.2 The monotonicity test

The first test examines whether the alpha dynamic after the portfolio sorting period has a monotone pattern (strictly increasing or strictly decreasing) as proposed in Eq. 2.

I examine the pattern in the on-paper alphas of the eleven characteristic-sorted portfolios with a fixed number of months skipped after sorting $k = 0$ and holding periods of $h = 1, 6, 12, 18, ..., 60$. In Figure 3a, I demonstrate how calendar time t relates to k and h. Since it is unknown how long it will take for alphas to decay to zero, I take the average of alphas over time by varying holding periods. In this way, even when alphas monotonically decay to zero before the $60th$ month, this method still may be able to identify the monotonically decreasing pattern of the alpha dynamic. Moreover, with a holding period gap of six months, the difference in alpha may be large, thereby increasing statistical power.

To examine the monotone pattern in alpha dynamics, I extend the monotonicity test proposed by Patton and Timmermann (2010, PT). While PT use the test to examine monotone relations in the cross-section of stocks, I use it to examine monotone relations in the time series (event time). With $k = 0$ and $h = 1, 6, 12, 18, ..., 60$, denote their on-paper alphas as $\alpha^{op}_{1,t}$, $\alpha^{op}_{2,t}$, ..., $\alpha^{op}_{11,t}$. Let $\Delta_i = E[\alpha^{op}_{i,t}] - E[\alpha^{op}_{i+1,t}]$ for $i = 1, ..., 11$. Let $\Delta \equiv [\Delta_1, ..., \Delta_{11}]'$. I examine four tests proposed by PT. The first is to test a strict monotone relation (MR test) in Δ. The null is a constant or weakly decreasing pattern and the alternative is a strictly increasing pattern.

$$\begin{aligned} &H_0 : \Delta \leq 0 \\ &H_1 : \Delta > 0 \end{aligned} \tag{10}$$

The alternative hypothesis is the one that I would like to test. Therefore, I flip the long and short portfolios when I would like to examine whether alphas decrease over time and do not flip if I would like to examine whether alphas increase over time.

If statistical power is low, the MR test may fail to detect a monotone pattern for these anomalies. To address the issue of low statistical power, PT propose the next two tests. They examine whether at least some parts of the pattern of Δ are strictly positive (Up test) or negative (Down test). The Up and Down tests both have a null hypothesis of a constant pattern. Therefore, they are less restricted than the MR test and potentially detect any small deviation from a constant pattern. Specifically, for the Up test,

$$\begin{aligned} &H_0 : \Delta = 0 \\ &H_1^+ : \sum_{i=1}^{11} |\Delta_i| \mathbb{1}\{\Delta_i > 0\} > 0 \end{aligned} \tag{11}$$

where the indicator $\mathbb{1}\{\Delta_i > 0\}$ is one if $\Delta_i > 0$. And for the Down test,

$$\begin{aligned} &H_0 : \Delta = 0 \\ &H_1^- : \sum_{i=1}^{11} |\Delta_i| \mathbb{1}\{\Delta_i < 0\} > 0 \end{aligned} \tag{12}$$

where the indicator $\mathbb{1}\{\Delta_i < 0\}$ is one if $\Delta_i < 0$. Up and Down tests apply a non-parametric method. The distribution of the statistics in the tests is estimated from 1,000 bootstrapping replications.

The last test is Bonferroni bound. This test is more conservative as discussed by PT. It analyzes whether the minimum t-statistic on estimated $\Delta_i,\ i = 1, ..., 11$, falls below the critical value derived from a bound on the probability of a Type I error.

With the statistics on monotonicity, I investigate whether the alpha dynamic has a monotone pattern. Under a 5% significance level, I examine whether the statistics in either the MR test, the Up test, the Down test, or the Bonferroni bound test are significant. In the MR test and the Bonferroni bound test, I examine both the null of weakly decreasing and weakly increasing patterns. Thus, this exercise is designed to identify the anomalies whose alpha dynamic has a pattern that is strictly increasing, strictly decreasing, or if some parts are strictly increasing, or some parts are strictly decreasing.

4.1.3 The optimization test

The second test is proposed in Eq. 4. The test is to investigate whether all strategies in k generate the same CAPM alpha.

The first step is to obtain the returns of the optimization strategy. Motivated by Proposition 1, the holding period is set to one month ($h = 1$). To alleviate data mining concerns and look-ahead bias, I conduct a conditional analysis by running expanding regressions. The first 60 months are burn-out months. At the end of any month $t \geq 60$, I search for the optimal number of months to skip (k^*) that maximizes the on-paper alpha from the set of $\{0, 12, 24, ..., 60\}$ and is based on the information already known at t. This set of k ensures that the same subset of months is considered as in the monotonicity test. I form a spread portfolio based on k^* at the end of month t. And I calculate the return of the portfolio in $t + 1$. The return in $t + 1$ can then be considered out-of-sample returns. This process is repeated until the end of the sample.

In Figure 3a, I demonstrate how calendar time t relates to k and h. And in Figure 3b, I present the timeline for the estimation of k^*, formation of the portfolio, and measurement of the returns. This timeline is also used in all optimization strategies of the paper.

Next, I obtain the returns of the benchmark strategy. At the end of any month $t \geq 60$, I form a portfolio by equally weighting the portfolios based on $k = 0, 12, 24, ..., 60$. That is, I form a portfolio with equal weight on the portfolios that are sorted at the end of months t, $t - 12$, ..., and $t - 60$. I then calculate the return of this portfolio in month $t + 1$. This process is repeated until the end of the sample. The benchmark strategy is to obtain the average of the returns of alternative strategies in k consistent with the alternative hypothesis of Eq. 4. As discussed in Section 2.2.3, the intuition is that although the optimization strategy based on k^* produces the highest alpha before month t by design, it will, on average, generate the same alpha in $t+1$ as that of the benchmark strategy if the alpha dynamic is constant.

Finally, I compare the alpha of the optimization strategy to that of the benchmark strategy. Specifically, I run the following regression:

$$r_{opt,t} - r_{b,t} = a_{opt} + b_m r_{m,t} + \epsilon_{opt,t} \tag{13}$$

where $r_{opt,t}$, $r_{b,t}$, and $r_{m,t}$ are the returns of the optimization strategy, the benchmark strategy, and the market.

Under a 5% significance, if a_{opt} is positive and statistically significant, the optimization strategy generates a higher alpha than the average of the alphas of alternative strategies in k. This rejects the null hypothesis that all strategies generate the same CAPM alpha and indicates the existence of non-zero alphas.

The challenge then lies in how to infer the statistical significance of a_{opt}. I study the full-sample, before-sample, in-sample, and post-sample periods for each anomaly. Throughout the paper, all regressions require at least 20 observations. However, statistical inference may still be biased when the sample size is small. Furthermore, errors ($\epsilon_{opt,t}$) may have heteroskedasticity and autocorrelation issues.

To address these concerns, I use a bootstrapping approach to estimate the p-value of a_{opt}. Let $\{r_{i,t} \mid t = 1, ..., T;\ i = opt,\ b,\ m\}$ be the actual returns recorded for the optimization strategy, the benchmark strategy, and the market over T months for a sample period. I first use the stationary bootstrap of Politis and Romano (1994) to randomly draw (with replacement) a new sample of returns $\{\tilde{r}_{i,\tau}^{(b)} \mid \tau(1), ..., \tau(T);\ i = opt,\ b,\ m\}$. Here, $\tau(t)$ is the new time index, randomly drawn from the actual data $\{1, ..., T\}$. $\tau(t)$ is common across i to preserve cross-sectional dependencies in returns. The bootstrap replication number, denoted as b, ranges from 1 to 2,000. Furthermore, to account for time series dependencies, returns data are drawn in blocks, I choose the average block length to be 10 months. Within each bootstrapping replication, I estimate Eq. 17 and obtain an estimate of a_{opt}. Finally, I obtain a distribution of a_{opt}, and I calculate whether a_{opt} is greater than 0 in 95% of 2,000 replications.

4.1.4 Results

Results are shown in Table 2. In the full sample period, out of 92 suspicious and 41 HLZ suspicious anomalies, 19 (21%) suspicious and 24 (59%) HLZ suspicious anomalies pass alpha-dynamic tests. Therefore, true alphas are likely to exist in these anomalies. This suggests that alpha-dynamic tests have an additional contribution in identifying the existence of non-zero alphas relative to t tests. Furthermore, the results suggest that simply raising the t cutoff from 1.96 to 3.0 for all anomalies may miss many real anomalies.

Additionally, in the before-sample, in-sample, and post-sample periods, the numbers of suspicious anomalies that pass alpha-dynamic tests are 17, 28, and 37. Results suggest that the tests are useful across subperiods, especially in the post-sample periods.

Among all 205 anomalies, the total number of anomalies that present a non-constant pattern of alpha dynamics are 102, 39, 93, and 75 in the full sample, before-sample, in-sample, and post-sample periods. Results are consistent with McLean and Pontiff (2016). Since anomalies that are due to mispricing are likely to be arbitraged away after publication, fewer anomalies should have true alphas in the post-sample period.

Furthermore, the monotonicity test and the optimization test complement each other well. In each subperiod, both tests can detect anomalies that may have true alphas while the other test cannot. One reason is that although both tests examine the same subset of months after portfolio sorting, they use different information. The monotonicity test uses all alphas in the 60 months after sorting. In contrast, the optimization test only uses alphas in the first month, 12 months later, 24 months later,..., and 60 months later. Furthermore, the optimization test may identify non-constant patterns not limited to monotone patterns.

In Appendix E, Figure E4 shows the number of anomalies that generate the highest alpha at different k values. This figure shows that $k = 0$ generates the highest alphas for most of the anomalies. This suggests that alphas decay over time for the majority of the anomalies.

Therefore, the results cannot be explained by random variation in the patterns of alpha dynamics.

4.1.5 Robustness: Multi-factor models

The CAPM is used as an illustration. It might be also interesting to examine alpha-dynamic tests with multi-factor models. In this section, I examine alphas relative to Fama and French (1993, FF3) three-factor model, a four-factor model including the factors in the FF3 and momentum (FF3+MOM), Fama and French (2015, FF5) five-factor model, and Hou, Xue, and Zhang (2015, HXZ) model.

I examine the same alpha-dynamic tests (both the monotonicity test and the optimization test) for different models. Results are shown in Table 3. The classifications of suspicious, HLZ suspicious, and robust anomalies are similar. The t test now examines the alpha relative to different asset pricing models.

The total number of anomalies that pass alpha-dynamic tests for the CAPM, FF3, FF3+MOM, FF5, and HXZ are 102, 96, 85, 77, and 65 in the full sample period. These numbers drop when the model becomes bigger, consistent with bigger models digesting more anomalies.

The numbers of suspicious anomalies that pass alpha-dynamic tests for the CAPM, FF3, FF3+MOM, FF5, and HXZ are 19, 23, 24, 23, and 34 in the full sample period. Additionally, all models have similar numbers of suspicious anomalies that pass alpha-dynamic tests in the post-sample and in-sample periods. These results indicate that alpha-dynamic tests are useful not only for the CAPM but also for other models. Perhaps more interestingly, although the number of robust anomalies decreases with bigger models, the number of suspicious anomalies does not decrease. This suggests that although bigger models digest more anomalies, their performance was overstated.

4.2 After-cost profitability

If an anomaly does not have non-zero alphas, it is relatively meaningless to study its after-cost profitability. Therefore, I examine after-cost profitability only on the anomalies that either pass the t test or alpha-dynamic tests. In total, I examine 132, 65, 128, and 92 anomalies in the full-sample, before-sample, in-sample, and after-sample periods. This section examines whether prior studies underestimate after-cost profitability of anomalies.

To study the impact of alpha dynamics on after-cost profitability, I start by conducting an optimization strategy on the holding period h to maximize the after-cost alpha. Similar to the procedure of the optimization strategy for on-paper alphas, I conduct conditional analysis by running expanding regressions to alleviate look-ahead bias and data mining concerns. Motivated by Section 2.3, I first restrict the number of months skipped $k = 0$. At the end of any month t, the optimal holding period h^* is searched from the set of $\{1, 3, 6, 9, 12\}$ and is based on the information already known at t. This set is chosen to enable a comparison with previous studies, as prior studies mainly use $h = 1$ and $h = 12$. Then I construct spread portfolios based on h^* with the overlapping portfolio approach, and record the returns between $t + 1$ and $t + h^*$.

In Figure 3a, I demonstrate how calendar time t relates to k and h. And in Figure 3b, I present the timeline for the estimation of h^*, formation of the portfolio, and measurement of the returns.

One important difference is that I estimate the optimal h^* every five years. This means that I use the same h^* in the following five years after an estimation. This reduces the computing time in calculating the rebalancing costs.

4.2.1 Benchmark strategies

To investigate whether alpha dynamics impact after-cost profitability, I first compare the performance with a benchmark strategy that always takes $k = 0$ and $h = 12$ (henceforth, H12). Since h is searched from $\{1, 3, 6, 9, 12\}$ and H12 already uses the longest holding period among alternative holding periods, trading costs cannot be further reduced by considering optimal h^*. Therefore, any superior performance from the optimization strategy should only come from the impact of alpha dynamics on the on-paper alphas. That is, comparing the performance of the optimization strategy with that of H12 allows for a clear understanding of the impact of alpha dynamics on after-cost profitability.

To have a complete analysis of whether prior studies underestimate after-cost profitability, I also consider three other benchmark strategies. The second benchmark strategy always takes $k = 0$ and $h = 1$ (henceforth, H1). Using this benchmark is because H1 is the most frequently used portfolio construction methods in the literature (A. Y. Chen & Zimmermann, 2021).

Furthermore, the third benchmark strategy determines the holding period by the turnover rate rule in Novy-Marx and Velikov (2016, henceforth, the NV method). The rule is that if each of the long and short sides, on average, turns over less than once a year, a one-month holding period is used. In other cases, a twelve-month holding period is used.

Finally, the fourth benchmark strategy takes the holding period used in the original paper (henceforth, the CV method). A. Y. Chen and Velikov (2023) use this method to evaluate after-cost profitability. Therefore, this benchmark is to compare with there results.[28]

Next, I will describe the performance metrics I use to compare the performance.

4.2.2 Performance metrics

I first use the following regression to explore whether an optimization strategy generates a higher after-cost alpha than the benchmark strategy.

$$r_{s,t} - r_{b,t} = a + b_m r_{m,t} + \epsilon_{s,t} \tag{14}$$

where $r_{s,t}$ is the after-cost return of the optimization strategy for a characteristic-sorted

[28]In the original papers, holding periods were distributed as follows: one month (104 anomalies), three months (7 anomalies), six months (2 anomalies), twelve months (91 anomalies), and thirty-six months (1 anomaly).

portfolio in month t, $r_{b,t}$ is the after-cost return of a benchmark strategy, and $r_{m,t}$ is the after-cost market return. If a is positive and significant at the 5% significance level, the optimization strategy generates a higher after-cost CAPM alpha than the benchmark strategy. Further, if an investor only invests in a single anomaly, then a is the performance metric of interest.

In addition, we may want to know which strategy should be used to be combined with the market factor so that the highest Sharpe ratio is obtained. $r_{s,t}$ may be exposed to the same risk factors as $r_{b,t}$. Then even if one strategy has a higher CAPM alpha than the other, they will generate the same Sharpe ratio if they are optimally combined with the market factor.

In contrast, $r_{s,t}$ may be exposed to additional risk factors that $r_{b,t}$ is not exposed to. This is possible if $r_{s,t}$ contains some time-series information that $r_{b,t}$ does not contain. If so, $r_{s,t}$ is the strategy that will generate a higher Sharpe ratio if it is optimally combined with the market factor. A second performance metric is obtained from the regression of:

$$r_{s,t} = a_s + b_s r_{b,t} + b_m r_{m,t} + \epsilon_{s,t} \tag{15}$$

Then a strategy is considered to outperform a benchmark strategy if a_s is positive and significant at a 5% significance level. In other words, allocating some positive weight to the strategy can improve the investment opportunity of the investors already trading the benchmark strategy and the market factor.

4.2.3 Results

First, in Panel A of Table 4, I present the number of anomalies for which the optimization strategy outperforms the benchmark strategies. Under a 5% significance level, I find that 31, 6, 30, and 12 anomalies generate a higher after-cost alpha than that of H12 in the full-sample, before-sample, in-sample, and post-sample periods, respectively. Further, 26, 5, 28, and 10 anomalies cannot be fully explained by the returns of H12 and the market factor.

These results suggest that alpha dynamics affect after-cost profitability.

In Panel B, I present the results when H1 is the benchmark. Although fewer anomalies (16) generate a higher after-cost alpha in the full sample period, more anomalies (34) cannot be explained by H1 and the market factor. This indicates that the appropriate holding periods to evaluate after-cost profitability are different across anomalies.

In Panel C and D, I present the results when the benchmark is NV and CV, respectively. Results are similar to those for H1 and H12. For economic magnitudes, the average improvement in monthly after-cost alpha (a) is 0.30% when NV is the benchmark and 0.28% when CV is the benchmark in the full-sample period. Post-sample improvements are stronger. The average improvement in monthly after-cost alpha is 0.46% when NV is the benchmark and 0.44% when CV is the benchmark. These results suggest that prior studies have significantly underestimated after-cost profitability of anomalies.

Can these results be generated from random variation? I further examine this question by conducting simulations following a similar procedure as in Bessembinder, Burt, and Hrdlicka (2021). In Appendix D, I describe how I conduct the simulations. The basic idea is to assume a data-generating process that no strategy based on k and h can outperform H1 based on statistics a and a_s. Then I conduct the same optimization strategies with simulated returns and investigate how many anomalies can outperform H1 in each simulation. Results show that random variation alone cannot explain the results in Panel A of Table 4. For example, under a t-cutoff of 1.96, 2.00, 2.50, 3.00, 3.50, and 4.00, the maximum numbers of anomalies that have positive and statistically significant a from 2,000 simulations are 10, 10, 5, 4, 2, 1 in the full-sample period. In contrast, these numbers are 16, 15, 10, 10, 7, and 7 from the actual data when the optimization strategy searches for h only.

4.2.4 Robustness: Search for both k and h

In this section, I empirically examine whether the number of months skipped k also affects after-cost profitability. Instead of restricting $k = 0$, I search for both k and h in the optimization strategy.

Results are shown in Table E1. I compare the results with the optimization strategy that searcher for h only. The number of anomalies that outperform the benchmark strategies H1 and H12 is lower when searching for both k and h. Results suggest that k does not have a significant impact on after-cost profitability.

4.3 Mispricing or rational expectations

If alphas exist, they could be due to mispricing or rational expectations. In this section, I focus on the anomalies that show evidence of the existence of non-zero alphas from either t tests or alpha-dynamic tests. As discussed in Section 2.4, behavioral explanations such as Barberis (2018) and Duffie (2010) predict that the alpha dynamic may exhibit a pattern with decaying ripples (both increasing and decreasing patterns exist in the alpha dynamic) as shown in Figure 2.

The next challenge is to identify such a ripple pattern empirically. The monotonicity test proposed by Patton and Timmermann (2010, PT) cannot be directly used to test ripple patterns since PT's tests can only test monotonicity patterns. Instead, to detect possible ripple patterns in alpha dynamics, I extend the methodology of estimating the optimal strategy for on-paper alphas in Section 4.1.3. There are two differences. First, I divide all the months after portfolio sorting into multiple consecutive unconnected subperiods. Next, I examine the pattern of the alpha dynamic in each subperiod. Second, I compare the alpha of the optimal strategy with the alphas at the start and end of the subperiod. This is to examine whether the maximum alpha in this subperiod occurs at the start, end, or middle. In this way, I can determine if the alpha dynamic has an increasing pattern, a decreasing pattern, or both.

To provide the intuition, denote a subperiod as $[k_{start}+1,\ k_{end}+1]$. At the end of each month t, the optimal strategy searches for the optimal k^* that generates the highest alpha from $[k_{start}, k_{end}]$. By running expanding regressions, this search is based on known information before month t. Then at the end of month t, a long-short portfolio is constructed based on k^* and the return in month $t+1$ is calculated. Figure 3a demonstrates how calendar time t relates to k and h. And Figure 3b presents the timeline for the estimation of k^*, formation of the portfolio, and measurement of the returns.

Last, I examine whether the CAPM alpha of the optimal strategy is statistically higher than the CAPM alphas at the start and end of the subperiod ($\alpha_{k_{start}+1}$ and $\alpha_{k_{end}+1}$). The alphas at the start and end of the subperiod are estimated separately from the strategy that always takes $k = k_{start}$ and $h = 1$ and the strategy that always takes $k = k_{end}$ and $h = 1$. Specifically, to compare the alphas between the optimal strategy and those at the start and end of the subperiod, I run the following regression:

$$r_{opt,t} - r_{edge,t} = \phi_0 + \phi_1 r_{m,t} + \epsilon_t \tag{16}$$

where $r_{opt,t}$ is the return of the optimization strategy and $r_{edge,t}$ is either the return on the left edge of the subperiod or the return on the right edge of the subperiod. $r_{m,t}$ is the market return and ϵ_t is the residule. If ϕ_0 is statistically significant and positive, it implies that the optimization strategy generates a higher CAPM alpha.

If the optimal strategy generates a CAPM alpha higher than that on the left edge of the subperiod $\alpha_{k_{start}+1}$, it implies that $\alpha_{k_{start}+1}$ is not the highest within the subperiod. This suggests that the alpha dynamic has at least one increasing pattern within the subperiod. With the same logic, if the optimal strategy generates a CAPM alpha higher than that on the right edge of the subperiod $\alpha_{k_{end}+1}$, then $\alpha_{k_{end}+1}$ is not the highest. This suggests that the alpha dynamic has at least one decreasing pattern within the subperiod.

Figure 5 presents this intuition with examples. The subfigure at the top shows two potential patterns of alphas when the optimal strategy generates a higher alpha than $\alpha_{k_{start}+1}$. The subfigure shows that there is at least one increasing pattern within the subperiod. And the subfigure at the bottom shows two potential patterns of alphas when the optimal strategy generates a higher alpha than $\alpha_{k_{end}+1}$. The subfigure shows that there is at least one decreasing pattern within the subperiod.

Once again, the method is to determine whether a pattern is predictable within a subperiod. If the pattern is unpredictable, the optimal strategy will not generate a higher alpha than the alphas at the start or end of a subperiod. In other words, the pattern identified by this test is unlikely to be generated from random variation in firm risks.

Specifically, I examine five subperiods in event time, [1,13], [13,37], [37,61], [61,85], and [85,109].[29] Since I examine five subperiods and compare two differences in CAPM alphas in each subperiod, there are ten hypotheses in total. To deal with multiple hypothesis testing, I adjust the p values by the Benjamini-Hochberg method to ensure that the expected proportion of false discovery rate is no greater than 5%. The Benjamini-Hochberg method is also used by Harvey et al. (2016) and Keloharju et al. (2021).

Figure 4 gives an example. The pattern is similar to that in Figure 2a. There is a decreasing pattern in the first subperiod [1,13]. Additionally, the method can identify an increasing pattern in the third subperiod [37,61], as well as both an increasing and decreasing pattern in the fourth subperiod [61,85]. Therefore, a ripple pattern can be identified based on my definition.

In Table 5, I present the pattern of some well-known anomalies within the nine years following portfolio sorting. The table shows that the ripple pattern appears in accruals ($Accruals$), idiosyncratic volatility ($IdioRisk$), momentum ($Mom12m$), and net share issuance ($SharIss1Y$) in the full sample period. Characteristic-sorted portfolios based on book-to-market show a strictly increasing pattern in the first twelve months after sorting. This is consistent with the increasing resolution of mispricing discussed in Section 2.1.3. The increasing pattern is also consistent with the finding in Giglio, Kelly, and Kozak (2023) when they estimate portfolios' term structure. They have, however, not formally tested the pattern or studied anomaly explanations.

[29] In the first subperiod, k can be chosen from the set of 0, 3, 6, 9, 12. For the rest of the subperiods, k is chosen from the smallest possible value to the largest possible value with a step of 6. For example, in the second subperiod, k is chosen from the set of 12, 18, 24, 30, 36.

In Table 6, I show all the anomalies that show the ripple pattern in different sample periods. Many anomalies exhibit a ripple of the alpha dynamic. The total number of anomalies that show this pattern in the full-sample, before-sample, in-sample, and post-sample subperiods are 79 (60%), 27 (41%), 69 (54%), and 41 (45%). Therefore, this test is useful in identifying a large number of anomalies that are potentially due to mispricing. The table also shows that other anomalies in the categories of idiosyncratic volatility, net share issuance, and momentum exhibit the ripple pattern as well.

Further, many anomalies such as accruals ($Accruals$) and net share issuance ($ShareIss1Y$) exhibit the ripple pattern in the in-sample period, but not in the post-sample period. A possible explanation could be that arbitrageurs correct mispricing following the publication of anomalies (McLean & Pontiff, 2016).

5 Other implications: portfolio construction in asset pricing

Besides explaining anomalies, studying the alpha dynamic can also shed light on optimization in portfolio construction. This is because alpha means are investigated jointly when studying alpha dynamics. The usual method in the literature that studies the alpha mean cannot reveal which portfolio construction method is optimal. This is because those alpha means are studied independently and one alpha mean may be higher than the other simply due to random variation.

Due to the limitation of the alpha mean, researchers have the concern that original authors may have data mined the results with the portfolio construction method they use (e.g., Hasler, 2022). For example, using different holding periods in portfolio construction is basically studying different subsets of months after sorting, and therefore different alpha means. For an anomaly, if the used portfolio construction method that takes a number of months skipped k and a holding period h is not the optimal method, this raises the data mining concern. Since $k = 0$ and $h = 1$ (H1) and $k = 0$ and $h = 12$ (H12) are the most used portfolio construction methods, we would like to know whether they are optimal. Section 4.2 shows that neither H1 nor H12 is optimal for all anomalies when after-cost alphas are considered in any subperiod. This section further studies whether H1 or H12 is optimal when on-paper alphas are considered.

Furthermore, alpha may not be the only metric of interest. Alpha is a mean and has not considered risk. However, both researchers and investors might also be interested in metrics that consider risk. In this section, I also study the information ratio (IR), which is alpha divided by the volatility of the residual. It is important because the optimal factor construction method should maximize the information ratio. This concept is mentioned in many studies such as Barillas and Shanken (2017). The intuition is that the maximum obtainable Sharpe ratio of the new model that includes the new factor is equal to the sum of the maximum obtainable Sharpe ratio of the old model and the information ratio of the new factor.

5.1 Optimization on the on-paper alpha

Optimization strategies are similar to that of the optimization test in Section 4.1.3. Instead of searching for k from the set of $\{0, 12, ..., 60\}$, I search for k from the set of $\{0, 3, 6, 9, 12\}$ to compare with H1 and H12. Figure 3a demonstrates how calendar time t relates to k and h. And Figure 3b presents the timeline for the estimation of k^*, formation of the portfolio, and measurement of the returns.

The performance metrics are similar to those in Eq. 14 and the a_s in Eq. 15. The difference is that all returns in the regression this time are on-paper returns. a^{op} measures whether the optimization strategy generates a higher on-paper alpha and a_s^{op} measures whether the optimization strategy can be completely explained by a benchmark strategy.

$$r_{s,t}^{op} - r_{b,t}^{op} = a^{op} + b_m r_{m,t}^{op} + \epsilon_{s,t}^{op} \tag{17}$$

$$r^{op}_{s,t} = a^{op}_s + b_s r^{op}_{b,t} + b_m r^{op}_{m,t} + \epsilon^{op}_{s,t} \tag{18}$$

Results are shown in Table 7. Results indicate that H12 is not optimal for many anomalies. Out of 139 anomalies, only a few (9) anomalies outperform H1 based on a in the full sample period. However, many more (33) anomalies outperform H1 based on a_s in the full sample period. Although the number is small based on a, we can also refer to results in the monotonicity test as additional evidence. If the pattern of the alpha dynamic in the first twelve months after sorting is strictly increasing, this suggests that H1 is not optimal based on Proposition 1 since alpha in the first month after sorting α_1 is not a maximum. The book-to-market characteristic is one of the examples that have an increasing pattern of alpha dynamic in the first twelve months after sorting.

5.2 Optimization on the information ratio

Results from on-paper alphas suggest that many anomalies cannot be fully explained by traditional portfolio construction methods H12. However, H12 is the traditional method used for constructing factors (e.g., Fama & French, 1993; Hou et al., 2015). In this section, I examine whether H12 is optimal for the size and book-to-market factors in FF3 based on the IR. For book-to-market, I investigate both the version that uses the market equity in December of the prior year ($BMdec$) and the version that uses the latest market equity (BM).

Even when the alpha dynamic is non-constant, the IR dynamic can be constant. If alphas are due to omitted risk premium, then alphas can be non-constant due to different exposures to an omitted risk factor. However, IR will be constant since the exposure appears in both the numerator and denominator of IR. Therefore, if the IR dynamic is non-constant, either because alpha is due to mispricing or because old and new alphas reflect risk compensation to different omitted factors.

I estimate IR relative to the CAPM. And I conduct a similar optimization strategy as before. The difference is that, in the optimization strategy, I start by searching for k from $\{0, 3, 6, 9, 12\}$ and $h = 1$ to maximize IR instead of alphas. Then I compare the IR of the optimization strategy with that of H12.

Results are shown in Table 8. For both size and BM, there is a significant increase in IR compared to H12. When considering the optimization strategy, the IR almost triples in the full sample period. For example, the IR for the optimization strategy is 0.118 for size and 0.077 for BM. In contrast, these values for H12 are 0.042 and 0.0.027, respectively.

Furthermore, holding periods h may also impact the IR. Since the overlapping portfolio approach is used, a longer holding period means holding multiple portfolios each month. This leads to a diversification effect among the portfolios. To investigate the effect of h, I conduct optimization strategies that search for h only or search for both k and h. The results in Table 8 indicate that these strategies do not outperform the strategy that searches for k only.

Overall, results suggest other dynamics such as the IR dynamic is also useful and the current factor construction method may not be optimal and can be improved.

6 Conclusion

To explain anomalies, prior studies focus on the average of alphas within a subset of months after the portfolio sorting period (the alpha mean). This paper's contribution is to investigate whether studying how alphas evolve over time after sorting (the alpha dynamic) helps better understand anomalies.

Results indicate that alpha dynamics have a significant impact on anomaly explanations. First, alpha-mean tests have several problems, and relying solely on alpha-mean tests may miss many anomalies. Alpha-dynamic tests alleviate these problems and help better determine the existence of non-zero alphas. Furthermore, determining the holding period based on both alpha dynamics and trading costs significantly improves after-cost profitability. Therefore, this method might be more appropriate to evaluate after-cost profitability. Moreover, for a large proportion of anomalies, alpha dynamic patterns are consistent with existing behavioral models rather than rational models. The findings of this study suggest that we have more real anomalies than we had thought, that there are more anomalies that are profitable when trading costs are considered, and that a large proportion of these anomalies might be at least partly due to mispricing.

Overall, results have important implications for academics seeking to understand anomalies, firm managers estimating discount rates, and investors considering asset allocations and trading strategies. Therefore, to better understand anomalies, we should study both the alpha mean and the alpha dynamic. For future research, alpha dynamics can also be applied to other frequencies beyond monthly frequency. Moreover, alpha are not the only measure of interest. For example, factor construction matters when there is a dynamic in the information ratio.

References

Baba Yara, F., Boons, M., & Tamoni, A. (2020). Persistent and transitory components of characteristics: Implications for asset pricing. *Available at SSRN 3529140*.

Barberis, N. (2018). Psychology-based models of asset prices and trading volume. In *Handbook of behavioral economics: Applications and foundations 1* (pp. 79–175). Elsevier.

Barberis, N., Greenwood, R., Jin, L., & Shleifer, A. (2015). X-capm: An extrapolative capital asset pricing model. *Journal of financial economics*, *115*(1), 1–24.

Barillas, F., & Shanken, J. (2017). Which alpha? *The Review of Financial Studies*, *30*(4), 1316–1338.

Berk, J. B., Green, R. C., & Naik, V. (1999). Optimal investment, growth options, and security returns. *The Journal of Finance*, *54*(5), 1553–1607.

Bessembinder, H., Burt, A., & Hrdlicka, C. M. (2021). Time series variation in the factor zoo. *Aaron Paul and Hrdlicka, Christopher M., Time Series Variation in the Factor Zoo (December 22, 2021)*.

Bessembinder, H., Burt, A., & Hrdlicka, C. M. (2022). Factor returns and out-of-sample alphas: Factor construction matters. *Aaron Paul and Hrdlicka, Christopher M., Factor Returns and Out-of-Sample Alphas: Factor Construction Matters (November 19, 2022)*.

Bessembinder, H., Cooper, M. J., & Zhang, F. (2021). Return horizon and mutual fund performance. *Michael J. and Zhang, Feng, Return Horizon and Mutual Fund Performance (November 9, 2021)*.

Boguth, O., Carlson, M., Fisher, A., & Simutin, M. (2011). Conditional risk and performance evaluation: Volatility timing, overconditioning, and new estimates of momentum alphas. *Journal of Financial Economics*, *102*(2), 363–389.

Bowles, B., Reed, A. V., Ringgenberg, M. C., & Thornock, J. R. (2023). Anomaly time. *Available at SSRN 3069026*.

Chen, A. Y. (2021). The limits of p-hacking: Some thought experiments. *The Journal of Finance*, *76*(5), 2447–2480.

Chen, A. Y., & Velikov, M. (2023). Zeroing in on the expected returns of anomalies. *Journal of Financial and Quantitative Analysis*, *58*(3), 968–1004.

Chen, A. Y., & Zimmermann, T. (2020). Publication bias and the cross-section of stock returns. *The Review of Asset Pricing Studies*, *10*(2), 249–289.

Chen, A. Y., & Zimmermann, T. (2021). Open source cross-sectional asset pricing. *Critical Finance Review, Forthcoming*.

Chen, Y., & Kaniel, R. (2021). Firm characteristics and stock price levels: A long-term discount rate perspective. *Available at SSRN 3868318*.

Chernov, M., Lochstoer, L. A., & Lundeby, S. R. (2022). Conditional dynamics and the multihorizon risk-return trade-off. *The Review of Financial Studies*, *35*(3), 1310–1347.

Cho, T., & Polk, C. (2019). Putting the price in asset pricing. *Available at SSRN 3499681*.

Chordia, T., Goyal, A., & Saretto, A. (2020). Anomalies and false rejections. *The Review of Financial Studies*, *33*(5), 2134–2179.

De Bondt, W. F., & Thaler, R. (1985). Does the stock market overreact? *The Journal of Finance*, *40*(3), 793–805.

Duffie, D. (2010). Presidential address: Asset price dynamics with slow-moving capital. *The Journal of Finance*, *65*(4), 1237–1267.

Engelberg, J., McLean, R. D., & Pontiff, J. (2018). Anomalies and news. *The Journal of Finance*, *73*(5), 1971–2001.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25*(2), 383–417.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22.

Favero, C. A., Melone, A., & Tamoni, A. (2019). Factor models with drifting prices. *Available at SSRN 3418352*.

Giglio, S., Kelly, B. T., & Kozak, S. (2023). *Equity term structures without dividend strips data* (tech. rep.). National Bureau of Economic Research.

Gomes, J., Kogan, L., & Zhang, L. (2003). Equilibrium cross section of returns. *Journal of Political Economy*, *111*(4), 693–732.

Hackbarth, D., & Johnson, T. (2015). Real options and risk dynamics. *The Review of Economic Studies*, *82*(4), 1449–1482.

Harvey, C. R. (2017). Presidential address: The scientific outlook in financial economics. *The Journal of Finance*, *72*(4), 1399–1440.

Harvey, C. R., Liu, Y., & Zhu, H. (2016). . . . and the cross-section of expected returns. *The Review of Financial Studies*, *29*(1), 5–68.

Hasler, M. (2022). Looking under the hood of data-mining. *Available at SSRN 4279944*.

Hendershott, T., Menkveld, A. J., Praz, R., & Seasholes, M. (2022). Asset price dynamics with limited attention. *The Review of Financial Studies*, *35*(2), 962–1008.

Heston, S. L., & Sadka, R. (2008). Seasonality in the cross-section of stock returns. *Journal of Financial Economics*, *87*(2), 418–445.

Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of finance*, *54*(6), 2143–2184.

Hou, K., Xue, C., & Zhang, L. (2015). Digesting anomalies: An investment approach. *The Review of Financial Studies*, *28*(3), 650–705.

Hou, K., Xue, C., & Zhang, L. (2020). Replicating anomalies. *The Review of Financial Studies*, *33*(5), 2019–2133.

Jacobs, H., & Müller, S. (2020). Anomalies across the globe: Once public, no longer existent? *Journal of Financial Economics*, *135*(1), 213–230.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, *48*(1), 65–91.

Jensen, T. I., Kelly, B., & Pedersen, L. H. (2022). Is there a replication crisis in finance? *The Journal of Finance*.

Kelly, B. T., Pruitt, S., & Su, Y. (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*, *134*(3), 501–524.

Keloharju, M., Linnainmaa, J. T., & Nyberg, P. (2021). Long-term discount rates do not vary across firms. *Journal of Financial Economics*, *141*(3), 946–967.

Lewellen, J., & Nagel, S. (2006). The conditional capm does not explain asset-pricing anomalies. *Journal of financial economics*, *82*(2), 289–314.

Linnainmaa, J. T., & Roberts, M. R. (2018). The history of the cross-section of stock returns. *The Review of Financial Studies*, *31*(7), 2606–2649.

Martin, I. W., & Nagel, S. (2021). Market efficiency in the age of big data. *Journal of Financial Economics*.

McLean, R. D., & Pontiff, J. (2016). Does academic research destroy stock return predictability? *The Journal of Finance*, *71*(1), 5–32.

Novy-Marx, R., & Velikov, M. (2016). A taxonomy of anomalies and their trading costs. *The Review of Financial Studies*, *29*(1), 104–147.

Novy-Marx, R., & Velikov, M. (2019). Comparing cost-mitigation techniques. *Financial Analysts Journal*, *75*(1), 85–102.

Patton, A. J., & Timmermann, A. (2010). Monotonicity in asset returns: New tests with applications to the term structure, the capm, and portfolio sorts. *Journal of Financial Economics*, *98*(3), 605–625.

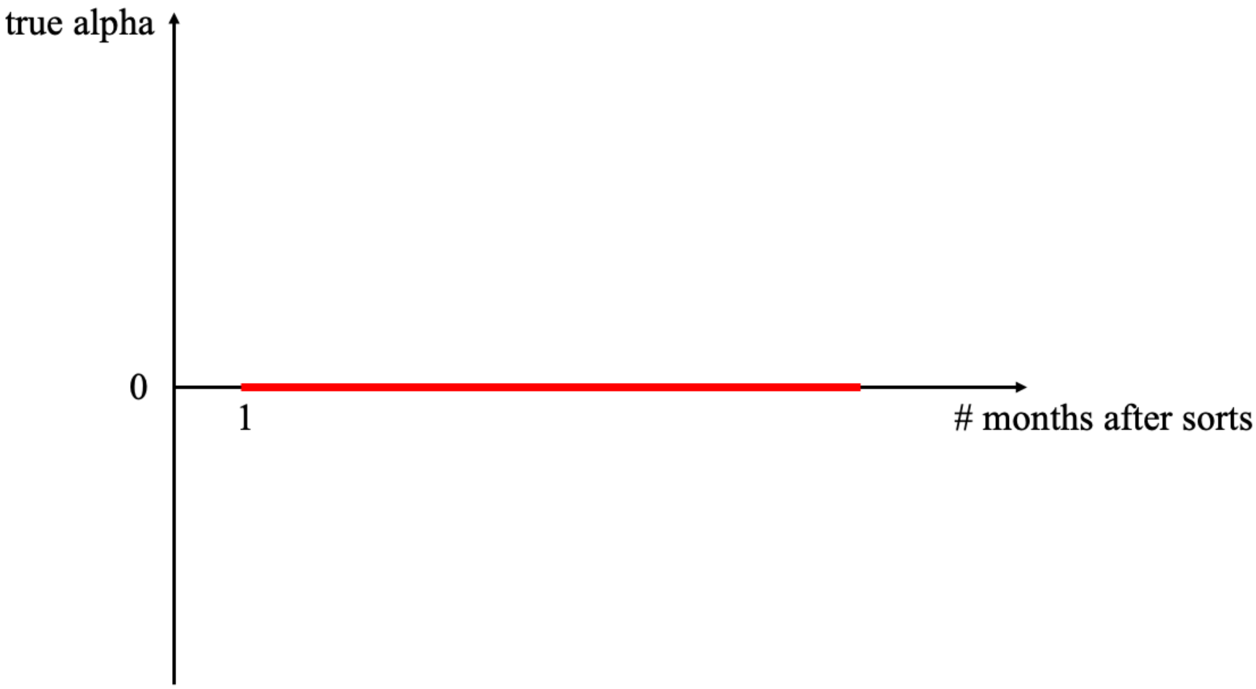
Politis, D. N., & Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical association*, *89*(428), 1303–1313.

Stambaugh, R. F., & Yuan, Y. (2017). Mispricing factors. *The review of financial studies*, *30*(4), 1270–1315.

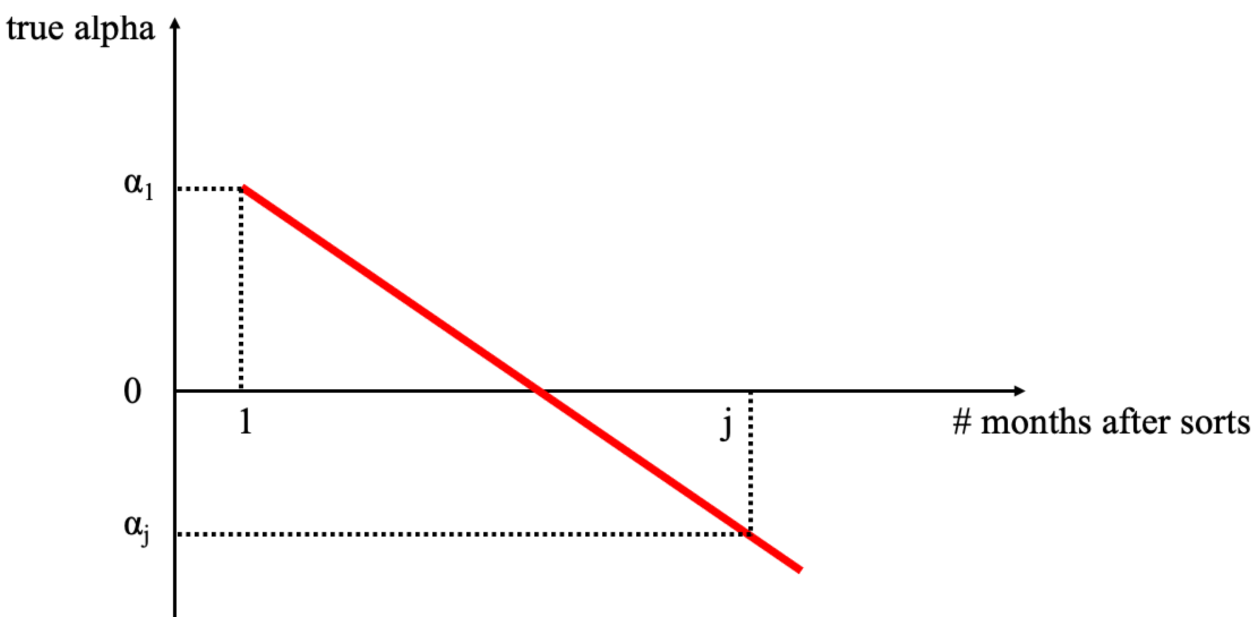
van Binsbergen, J. H., Boons, M., Opp, C. C., & Tamoni, A. (2023). Dynamic asset (mis) pricing: Build-up versus resolution anomalies. *Journal of Financial Economics*, *147*(2), 406–431.

van Binsbergen, J. H., & Opp, C. C. (2019). Real anomalies. *The Journal of Finance*, *74*(4), 1659–1706.

Zhang, L. (2005). The value premium. *The Journal of Finance*, *60*(1), 67–103.



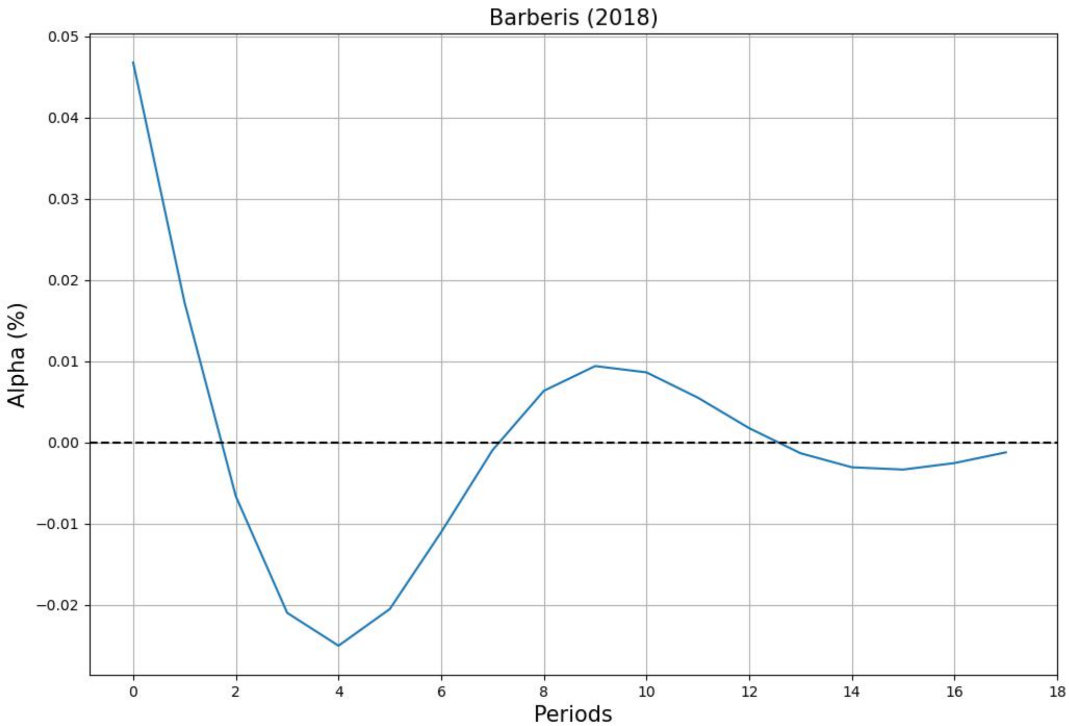
(a) The alpha dynamic when alphas do not exist



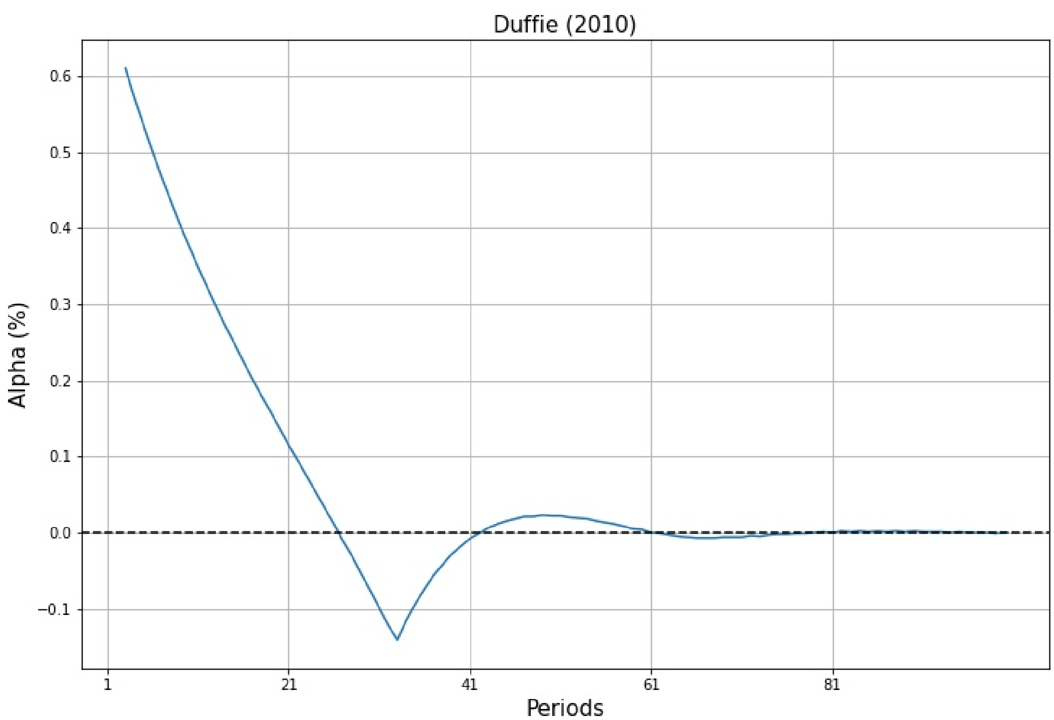
(b) An example of non-constant alpha dynamic

Figure 1: Patterns of alpha dynamics

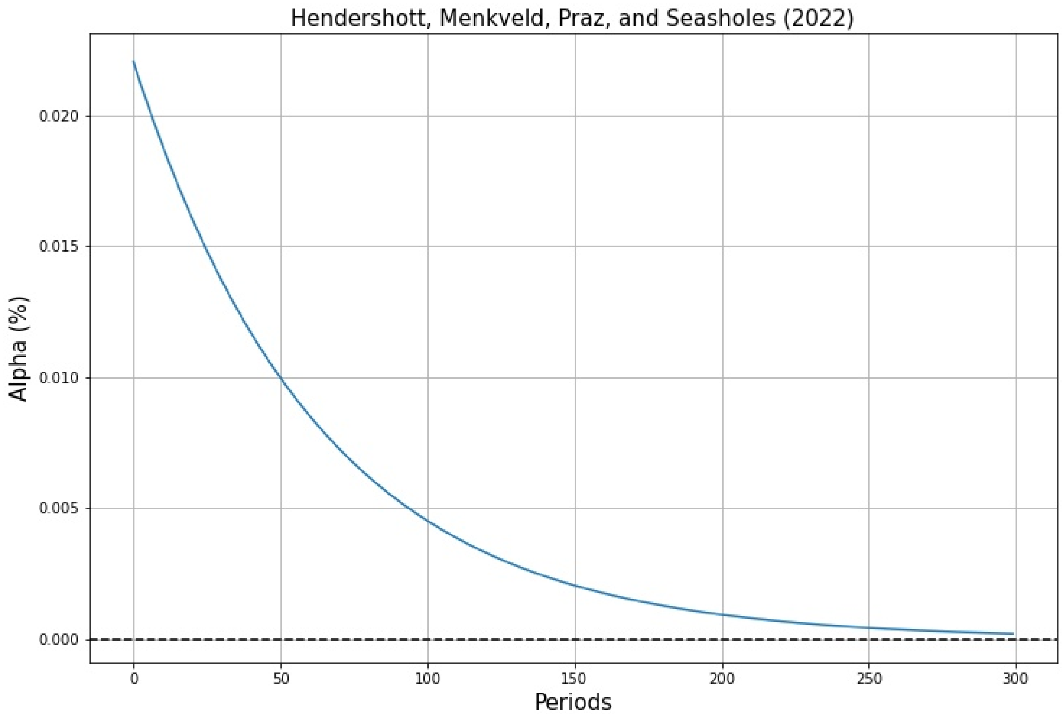
This figure shows the patterns of alpha dynamics. The y axis represents true alphas. x axis represents the number of months after the portfolio sorting period. The true alpha represents the population mean. Figure (a) shows how alphas after sorting evolve under the null hypothesis that alphas do not exist. And Figure (b) shows an example of how true alphas after sorting evolve under a non-constant alpha dynamic.



(a) The alpha dynamic of Barberis (2018)



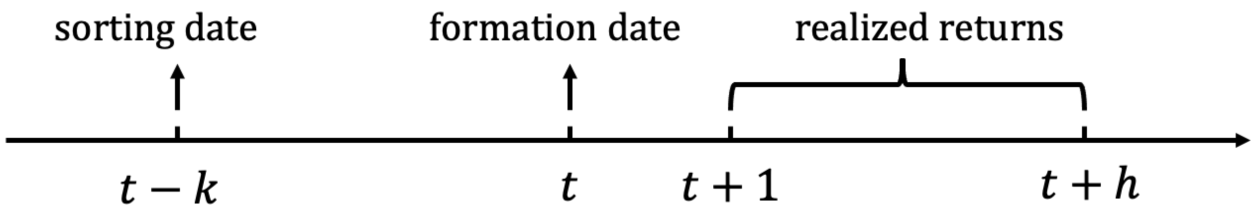
(b) The alpha dynamic of Duffie (2010)



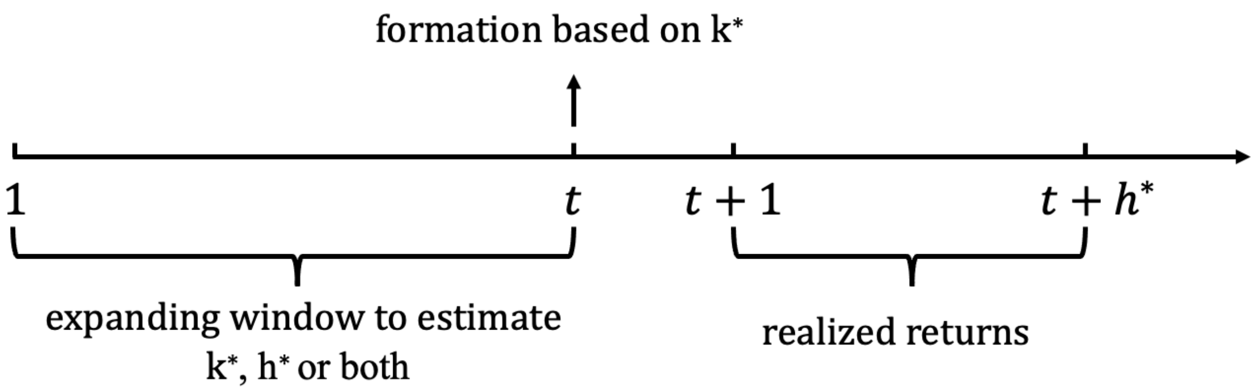
(c) The alpha dynamic of Hendershott, Menkveld, Praz, and Seasholes (2022)

Figure 2: Dynamics of alphas in Barberis (2018), Duffie (2010), and Hendershott, Menkveld, Praz, and Seasholes (2022)

The figures show the patterns of alpha dynamics implied by the models of Barberis (2018), Duffie (2010), and Hendershott, Menkveld, Praz, and Seasholes (2022). The y axis represents true alphas. x axis represents the number of periods after the initial shocking dates when mispricing arises. Alphas equal changes of mispricing over time.



(a) Relationship between t, k, and h



(b) Timeline for estimation, portfolio formation, and measurement of returns

Figure 3: Timelines

Figure 3a shows the relationship between calendar time t, the number of months skipped after sorting k and holding period h. Figure 3b shows the timeline for estimation, portfolio formation, and measurement of returns. When searching only for the optimal k^*, $h = 1$. Additionally, when searching for the optimal h^*, I measure returns from $t+1$ and $t+h^*$ using the overlapping portfolio approach. All optimization strategies in the paper use this timeline.

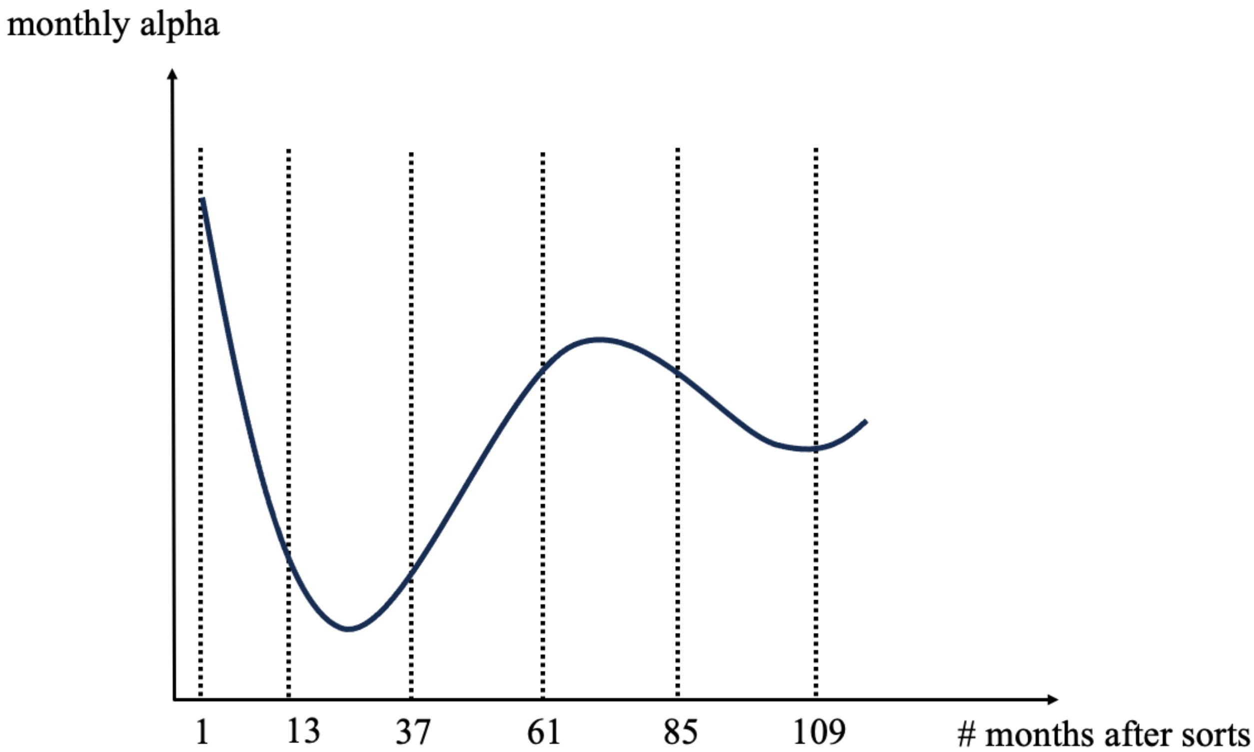
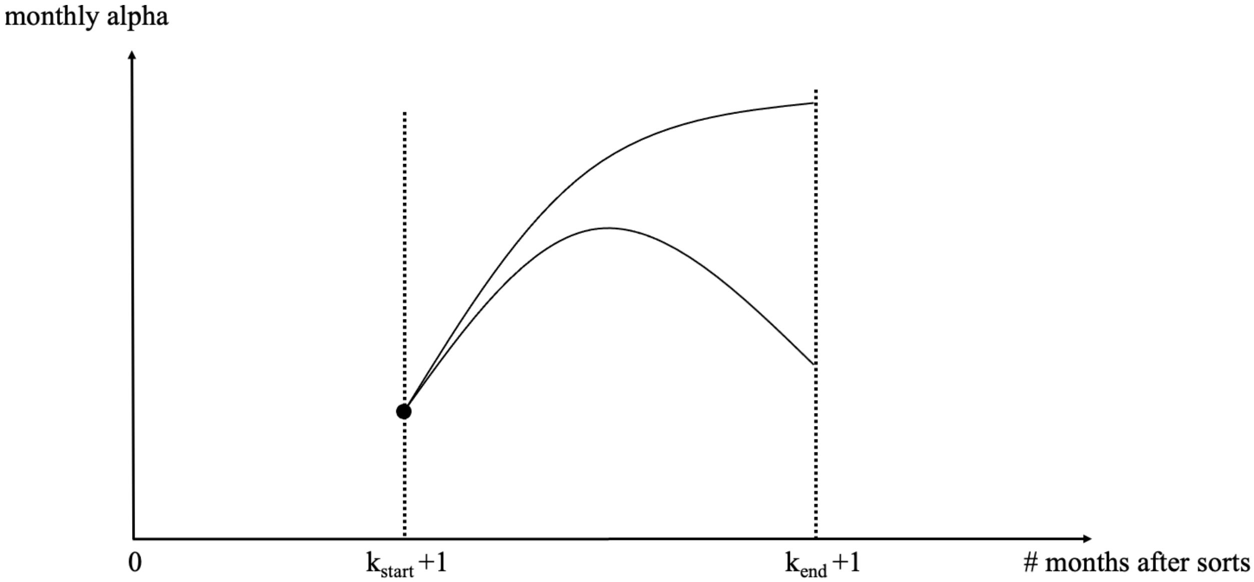
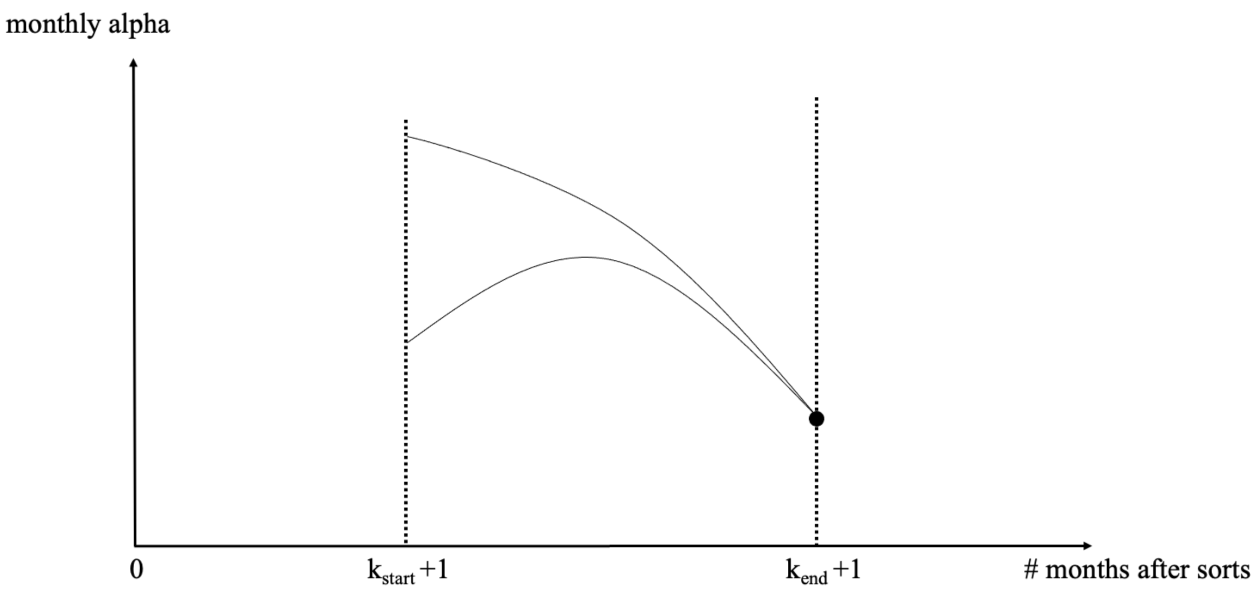


Figure 4: Examining the ripple pattern in the alpha dynamic

The figure provides an example to illustrate how I identify the ripple pattern in the alpha dynamic. The y axis represents monthly alphas, while the x axis represents the number of months after the sorting date. I examine five consecutive, unconnected subperiods after sorting: [1,13], [13,37], [37,61], [61,85], and [85,109]. Within each subperiod, I analyze whether there is an increasing pattern, a decreasing pattern, or both in the alpha dynamic. A ripple pattern occurs when the alpha dynamic exhibits both increasing and decreasing patterns over the nine years (109 months) following sorting.



(a) Examples of alpha dynamics when the optimization strategy generates a higher alpha than that at the start of a subperiod



(b) Examples of alpha dynamics when the optimization strategy generates a higher alpha than that at the end of a subperiod

Figure 5: Potential patterns within a subperiod

The figures show the potential patterns of alpha dynamics within a subperiod of months after portfolio sorting by comparing the alpha of the optimization strategy with those at the start and end of the subperiod. The y axis represents true alphas. x axis represents the number of months after the sorting date. The top figure shows examples of patterns of alpha dynamics within the subperiod $[k_{start}+1, k_{end}+1]$ when the optimization strategy generates a higher alpha than $\alpha_{k_{start}+1}$. The bottom figure shows examples of patterns of alpha dynamics within the subperiod $[k_{start}+1, k_{end}+1]$ when the optimization strategy generates a higher alpha than $\alpha_{k_{end}+1}$.

Table 1: A summary of the contributions of alpha dynamics to anomaly explanations

Explanations	New Information from Alpha Dynamics
Existence of Alphas	The alpha dynamic can detect non-zero alphas when the alpha mean may not.
After-cost Profitability	Holding periods should be determined by jointly considering the alpha dynamic and trading costs.
Mispricing versus Rational Expectations	Existing behavioral models can exhibit a "ripple" pattern of the alpha dynamic. Such a pattern is not implied in any existing rational models. Therefore, it can be used to distinguish between existing behavioral models and rational models. If there is a ripple pattern, it is likely due at least in part to mispricing.

Table 2: alpha-dynamic tests on the existence of non-zero alphas

This table presents the number of anomalies that pass alpha-dynamic tests under the 5% significance level. Alphas are relative to the CAPM. Within a sample period and with a holding period of one month, I label those that fail the t test with a cutoff of 1.96 as suspicious anomalies. Among the remaining anomalies, I label those that fail the t test with a cutoff of 3.0 as HLZ suspicious anomalies, while the rest as robust anomalies. Panel A shows the results for all anomalies. Panel B shows the results for suspicious anomalies. Panel C shows the results for HLZ suspicious anomalies. And Panel D shows the results for robust anomalies (72 anomalies). $Mono$ shows the results from the monotonicity test. *Opt* shows the results from the optimization test. $Total$ presents the number of anomalies that either passes the monotonicity test or the optimization test. The last two columns ("#" and "%") compare the results of alpha-dynamic tests with those of t tests. "#" shows the number of anomalies within a category and "%" shows the percentage of anomalies passing alpha-dynamic tests within a category.

Panel A: all anomalies					
Period	Mono	Opt	Total		
Full-sample	77	75	102		
Before-sample	28	23	39		
In-sample	68	72	93		
Post-sample	52	45	75		
Panel B: suspicious anomalies					
Period	Mono	Opt	Total	#	%
Full-sample	13	8	19	92	21%
Before-sample	14	8	17	137	12%
In-sample	16	19	28	105	27%
Post-sample	23	18	37	150	25%
Panel C: HLZ suspicious anomalies					
Period	Mono	Opt	Total	#	%
Full-sample	15	15	24	41	59%
Before-sample	7	9	14	32	44%
In-sample	7	7	12	35	34%
Post-sample	21	18	27	43	63%
Panel D: robust anomalies					
Period	Mono	Opt	Total	#	%
Full-sample	49	52	59	72	82%
Before-sample	7	6	8	16	50%
In-sample	45	46	53	65	82%
Post-sample	8	9	11	12	92%

Table 3: alpha-dynamic tests with different benchmark models

This table presents the total number of anomalies that pass alpha-dynamic tests relative to different asset pricing models. Columns 2 to 6 separately examine the alphas relative to the CAPM, Fama and French (1993, FF3) three-factor model, a four-factor model including the factors in the FF3 and momentum (FF3+MOM), Fama and French (2015, FF5) five-factor model, and Hou, Xue, and Zhang (2015, HXZ) model. Within a sample period and with a holding period of one month, I label those that fail the t test with a cutoff of 1.96 as suspicious anomalies. Among the remaining anomalies, I label those that fail the t test with a cutoff of 3.0 as HLZ suspicious anomalies, while the rest as robust anomalies. Panel A shows for results for all anomalies. Panel B shows the results for suspicious anomalies. Panel C shows the results for HLZ suspicious anomalies. And Panel D shows the results for robust anomalies.

Panel A: all anomalies					
Period	CAPM	FF3	FF3+MOM	FF5	HXZ
Full-sample	102	96	85	77	65
Before-sample	39	40	32	23	14
In-sample	93	90	80	66	60
Post-sample	75	67	68	54	51
Panel B: suspicious anomalies					
Full-sample	19	23	24	23	34
Before-sample	17	14	12	11	10
In-sample	28	29	29	23	32
Post-sample	37	31	33	34	36
Panel C: HLZ suspicious anomalies					
Full-sample	24	14	22	19	17
Before-sample	14	15	11	8	3
In-sample	12	11	16	13	12
Post-sample	27	20	23	11	10
Panel D: robust anomalies					
Full-sample	59	59	39	35	14
Before-sample	8	11	9	4	1
In-sample	53	50	35	30	16
Post-sample	11	16	12	9	5

Table 4: Impact of alpha dynamics on after-cost profitability

The table presents the number of anomalies for which the optimization strategy outperforms the benchmark strategies and the magnitude of the improvement. I only consider anomalies that pass either the t test or alpha-dynamic tests in each sample period. In total, I examine 132, 65, 128, and 92 anomalies in the full-sample, before-sample, in-sample, and post-sample periods. In Panel A, the benchmark strategy always takes a holding period of twelve months (H12). In Panel B, the benchmark strategy always takes a holding period of one month (H1). In Panel C, the benchmark strategy determines the holding period based on the turnover rate rule of Novy-Marx and Velikov (2016) (NV). And in Panel D, the benchmark strategy takes the holding period that is used in the original paper (CV). The performance metrics are a and a_s. They are defined in Equation 14 and 15. $a > 0$ implies that the optimization strategy generates a higher after-cost alpha than a benchmark strategy. And $a_s > 0$ implies that the returns of the optimization strategy cannot be completely explained by a benchmark strategy and the market factor. a and a_s are expressed in monthly percentages.

Panel A: Benchmark is H12												
	a						a_s					
Period	N	Mean	STD	25%	50%	75%	N	Mean	STD	25%	50%	75%
Full-sample	31	0.25	0.19	0.12	0.20	0.28	26	0.27	0.2	0.12	0.19	0.29
Before-sample	6	0.44	0.18	0.30	0.38	0.60	5	0.50	0.22	0.32	0.44	0.66
In-sample	30	0.36	0.26	0.19	0.26	0.47	28	0.37	0.27	0.20	0.28	0.43
Post-sample	12	0.42	0.28	0.22	0.34	0.56	10	0.45	0.28	0.27	0.32	0.72
Panel B: Benchmark is H1												
	a						a_s					
Period	N	Mean	STD	25%	50%	75%	N	Mean	STD	25%	50%	75%
Full-sample	16	0.36	0.25	0.22	0.24	0.44	34	0.27	0.19	0.13	0.23	0.43
Before-sample	4	0.59	0.41	0.29	0.49	0.79	9	0.37	0.34	0.10	0.37	0.45
In-sample	19	0.39	0.22	0.24	0.37	0.50	29	0.40	0.29	0.15	0.28	0.63
Post-sample	6	0.47	0.41	0.15	0.31	0.75	8	0.34	0.11	0.27	0.33	0.41
Panel C: Benchmark is NV												
	a						a_s					
Period	N	Mean	STD	25%	50%	75%	N	Mean	STD	25%	50%	75%
Full-sample	26	0.30	0.21	0.16	0.23	0.32	37	0.30	0.17	0.18	0.27	0.40
Before-sample	7	0.51	0.33	0.28	0.32	0.66	9	0.51	0.28	0.32	0.41	0.75
In-sample	27	0.34	0.20	0.18	0.30	0.46	31	0.41	0.28	0.20	0.30	0.55
Post-sample	7	0.46	0.37	0.24	0.34	0.62	10	0.32	0.12	0.26	0.31	0.39

Table 4: Impact of alpha dynamics on after-cost profitability

Panel D: Benchmark is CV												
	a						a_s					
Period	N	Mean	STD	25%	50%	75%	N	Mean	STD	25%	50%	75%
Full-sample	21	0.28	0.20	0.16	0.23	0.30	30	0.26	0.13	0.16	0.26	0.30
Before-sample	5	0.36	0.17	0.26	0.31	0.32	6	0.40	0.21	0.29	0.36	0.44
In-sample	23	0.31	0.18	0.17	0.27	0.42	29	0.34	0.22	0.19	0.28	0.45
Post-sample	8	0.44	0.39	0.12	0.29	0.71	7	0.35	0.18	0.27	0.34	0.45

Table 5: Alpha dynamic patterns of well-known anomalies

This table presents the patterns of alphas of a few well-known anomalies in the first nine years of months after portfolio sorting, accruals, idiosyncratic volatility, momentum, net share issuance, and book-to-market. Start and End are the start and end number of months away from the portfolio sorting period. For example, the first row examines the alpha dynamic pattern within the first twelve months after the portfolio sorting period. *increasing*, *decreasing*, and *both* are statistically significant patterns under a 5% significance level. *increase* means there is at least an increasing pattern within the subperiod. *decrease* means there is at least a decreasing pattern within the subperiod. And *both* means there are both increasing and decreasing patterns within the subperiod. *none* means there is no statistically significant pattern within the subperiod.

Start	End	Accruals	Idio Vol	Mom12m	Net Share Issuance	B/M
1	13	decrease	decrease	decrease	none	increase
13	37	decrease	both	both	decrease	none
37	61	both	both	both	both	none
61	85	both	both	both	both	none
85	109	decrease	both	both	both	none

Table 6: Anomalies that exhibit the ripple pattern in alpha dynamics.

This table presents the anomalies that exhibit a statistically significant ripple pattern in alpha dynamics under a 5% false discovery rate. I consider the anomalies that pass either the t test or alpha-dynamic tests in each sample period. In total, I examine 132, 65, 128, and 92 anomalies in the full-sample, before-sample, in-sample, and post-sample periods. The ripple pattern is a pattern that features both increasing and decreasing patterns in the whole alpha dynamic. Detailed descriptions of anomaly acronyms can be found in Appendix A.

Period	Total Number	Acronym
Full-sample	79	'AM' 'AbnormalAccruals' 'Accruals' 'AnalystRevision' 'AnnouncementReturn' 'AssetGrowth' 'CBOperProf' 'CF' 'ChAssetTurnover' 'ChEQ' 'ChInv' 'ChInvIA' 'ChNWC' 'CompEquIss' 'Coskewness' 'CustomerMomentum' 'DelCOA' 'DelEqu' 'DelFINL' 'DelNetFin' 'EntMult' 'EquityDuration' 'FEPS' 'FirmAgeMom' 'ForecastDispersion' 'GP' 'GrSaleToGrInv' 'High52' 'IO_ShortInterest' 'IdioRisk' 'IdioVol3F' 'IdioVolAHT' 'Illiquidity' 'IndMom' 'IndRetBig' 'IntMom' 'InvGrowth' 'InvestPPEInv' 'Leverage' 'MaxRet' 'Mom12m' 'Mom12mOffSeason' 'Mom6m' 'Mom6mJunk' 'MomOffSeason06YrPlus' 'MomRev' 'MomSeason06YrPlus' 'MomSeason11YrPlus' 'MomVol' 'NOA' 'NetDebtFinance' 'NetEquityFinance' 'NetPayoutYield' 'OPLeverage' 'OScore' 'OperProf' 'OperProfRD' 'OrgCap' 'PS' 'RD' 'REV6' 'RIO_Volatility' 'ResidualMomentum' 'SP' 'ShareIss1Y' 'Size' 'SmileSlope' 'VolMkt' 'VolSD' 'XFIN' 'betaVIX' 'dNoa' 'grcapx' 'grcapx3y' 'roaq' 'sfe' 'std_turn' 'zerotrade' 'zerotradeAlt1' 'zerotradeAlt12'
Before-sample	27	'ChInv' 'ChNWC' 'CompEquIss' 'DelCOA' 'FirmAgeMom' 'GrSaleToGrInv' 'High52' 'IdioRisk' 'IdioVol3F' 'IndMom' 'IndRetBig' 'InvestPPEInv' 'MaxRet' 'Mom12m' 'Mom12mOffSeason' 'Mom6m' 'MomOffSeason06YrPlus' 'MomSeason' 'MomSeasonShort' 'MomVol' 'NetEquityFinance' 'ShareIss1Y' 'VolMkt' 'VolSD' 'std_turn' 'zerotrade'
In-sample	69	'AM' 'AbnormalAccruals' 'Accruals' 'AnnouncementReturn' 'AssetGrowth' 'BMdec' 'CBOperProf' 'ChAssetTurnover' 'ChEQ' 'ChInv' 'CompositeDebtIssuance' 'DelCOA' 'DelEqu' 'DelFINL' 'DelNetFin' 'EntMult' 'EquityDuration' 'FEPS' 'FirmAgeMom' 'GP' 'GrSaleToGrInv' 'High52' 'IdioRisk' 'IdioVol3F' 'IdioVolAHT' 'IndRetBig' 'IntMom' 'InvGrowth' 'InvestPPEInv' 'MaxRet' 'Mom12m' 'Mom12mOffSeason' 'Mom6m' 'Mom6mJunk' 'MomOffSeason' 'MomRev' 'MomSeason11YrPlus' 'MomSeason16YrPlus' 'MomSeasonShort' 'MomVol' 'NOA' 'NetEquityFinance' 'NetPayoutYield' 'OPLeverage' 'OperProf' 'OperProfRD' 'OrgCap' 'ProbInformedTrading' 'REV6' 'RIO_MB' 'RIO_Volatility' 'ResidualMomentum' 'ShareIss1Y' 'ShareIss5Y' 'Size' 'VolMkt' 'VolSD' 'XFIN' 'betaVIX' 'dNoa' 'roaq' 'sfe' 'std_turn' 'tang' 'zerotrade' 'zerotradeAlt1' 'zerotradeAlt12'

Period	Total Number	Acronym
Post-sample	41	'AnalystRevision' 'AnnouncementReturn' 'ChInv' 'ChInvIA' 'CompEquIss' 'Coskewness' 'EarningsForecastDisparity' 'EquityDuration' 'FEPS' 'FirmAgeMom' 'ForecastDispersion' 'High52' 'IO_ShortInterest' 'IdioRisk' 'IdioVolAHT' 'IndRetBig' 'MaxRet' 'Mom12m' 'Mom12mOffSeason' 'Mom6m' 'MomOffSeason06YrPlus' 'MomSeason11YrPlus' 'MomVol' 'NetEquityFinance' 'NetPayoutYield' 'OperProf' 'OperProfRD' 'PS' 'RD' 'REV6' 'SP' 'Size' 'SmileSlope' 'Tax' 'VolMkt' 'VolSD' 'XFIN' 'grcapx3y' 'roaq' 'std_turn'

Table 7: Optimization on the on-paper alphas

This table presents the results for the number of anomalies for which the optimization strategy outperforms the benchmark strategies. The benchmark strategies are those that always take a twelve-month holding period (H12) or always take a one-month holding period (H1). I only consider anomalies that pass either the t test or alpha-dynamic tests in each sample period. In total, I examine 132, 65, 128, and 92 anomalies in the full-sample, before-sample, in-sample, and post-sample periods. The performance metrics are a^{op} and a_s^{op} as in Equation 17 and 18. a^{op} examines whether the optimization strategy generates a higher on-paper alpha than a benchmark strategy. And a_s^{op} examines whether the returns of the optimization strategy can be completely explained by a benchmark strategy.

	Benchmark strategy			
	H12		H1	
Period	$a^{op} > 0$	$a_s^{op} > 0$	$a^{op} > 0$	$a_s^{op} > 0$
Full-sample	45	46	9	33
Before-sample	15	16	4	15
In-sample	49	48	11	34
Post-sample	20	15	5	15

Table 8: Optimization on the information ratio

This table compares the results between the optimization strategy and the benchmark strategy that takes $k = 0$ and $h = 12$ (H12) based on the information ratio (IR). k is the number of skipped months and h is the holding period. I consider *Size* and *Book-to-market* factors. *BM* uses the latest market equity to construct the book-to-market. And *BMdec* uses the market equity in December of the prior year to construct book-to-market. k is searched from $\{0, 3, 6, 9, 12\}$ and h is searched from $\{1, 3, 6, 9, 12\}$ to maximize IR over time. k^* searches for k only and restricts $h = 1$. h^* searches for h only and restricts $k = 0$. And k^*h^* searches for both k and h. T is the sample length.

Panel A: Size					
Period	T	k^*	h^*	k^*h^*	H12
Full-sample	1070	0.118	0.118	0.118	0.042
In-sample	514	0.075	0.075	0.075	0.038
Post-sample	556	0.101	0.101	0.101	0.027
Panel B: BM					
Full-sample	618	0.077	0.028	0.075	0.027
Before-sample	53	0.002	-0.009	0.003	-0.007
In-sample	132	0.112	0.086	0.112	0.086
Post-sample	433	0.025	-0.018	0.023	-0.017
Panel C: BMdec					
Full-sample	739	0.069	0.064	0.064	0.065
Before-sample	48	0.045	0.039	0.032	0.052
In-sample	330	0.103	0.091	0.099	0.093
Post-sample	361	-0.019	-0.016	-0.019	-0.019

A Appendix: Descriptions of anomalies

Acronym	Authors	Year	Description
AbnormalAccruals	Xie	2001	Abnormal Accruals
AbnormalAccrualsPercent	Hafzalla, Lundholm, Van Winkle	2011	Percent Abnormal Accruals
AccrualQuality	Francis, LaFond, Olsson, Schipper	2005	Accrual Quality
AccrualQualityJune	Francis, LaFond, Olsson, Schipper	2005	Accrual Quality in June
Accruals	Sloan	1996	Accruals
Activism2	Cremers and Nair	2005	Active shareholders
AdExp	Chan, Lakonishok and Sougiannis	2001	Advertising Expense
AMq	Fama and French	1992	Total assets to market (quarterly)
AnalystValue	Frankel and Lee	1998	Analyst Value
AnnouncementReturn	Chan, Jegadeesh and Lakonishok	1996	Earnings announcement return
AOP	Frankel and Lee	1998	Analyst Optimism
AssetGrowth	Cooper, Gulen and Schill	2008	Asset growth
AssetLiquidityMarket	Ortiz-Molina and Phillips	2014	Asset liquidity over market
AssetLiquidityMarketQuart	Ortiz-Molina and Phillips	2014	Asset liquidity over market (qtrly)
AssetTurnover	Soliman	2008	Asset Turnover
Beta	Fama and MacBeth	1973	CAPM beta
BetaBDLeverage	Adrian, Etula and Muir	2014	Broker-Dealer Leverage Beta
betaCC	Acharya and Pedersen	2005	Illiquidity-illiquidity beta (beta2i)
betaCR	Acharya and Pedersen	2005	Illiquidity-market return beta (beta4i)
BetaDimson	Dimson	1979	Dimson Beta
BetaFP	Frazzini and Pedersen	2014	Frazzini-Pedersen Beta
betaNet	Acharya and Pedersen	2005	Net liquidity beta (betanet,p)
betaRR	Acharya and Pedersen	2005	Return-market return illiquidity beta
BetaTailRisk	Kelly and Jiang	2014	Tail risk beta
BidAskTAQ	Hou and Loh	2016	Bid-ask spread (TAQ)
BM	Rosenberg, Reid, and Lanstein	1985	Book to market using most recent ME
BMdec	Fama and French	1992	Book to market using December ME
BMq	Rosenberg, Reid, and Lanstein	1985	Book to market (quarterly)
BookLeverageQuarterly	Fama and French	1992	Book leverage (quarterly)
BrandCapital	Belo, Lin and Vitorino	2014	Brand capital to assets
CapTurnover	Haugen and Baker	1996	Capital turnover
CapTurnover_q	Haugen and Baker	1996	Capital turnover (quarterly)
Cash	Palazzo	2012	Cash to assets
cashdebt	Ou and Penman	1989	CF to debt
CBOperProf	Ball et al.	2016	Cash-based operating profitability
CBOperProfLagAT_q	Ball et al.	2016	Cash-based oper prof lagged assets qtrly
CF	Lakonishok, Shleifer, Vishny	1994	Cash flow to market
CFq	Lakonishok, Shleifer, Vishny	1994	Cash flow to market quarterly
ChangeInRecommendation	Jegadeesh et al.	2004	Change in recommendation

ChAssetTurnover	Soliman	2008	Change in Asset Turnover
ChEQ	Lockwood and Prombutr	2010	Growth in book equity
ChInv	Thomas and Zhang	2002	Inventory Growth
ChInvIA	Abarbanell and Bushee	1998	Change in capital inv (ind adj)
ChNAnalyst	Scherbina	2008	Decline in Analyst Coverage
ChNCOA	Soliman	2008	Change in Noncurrent Operating Assets
ChNCOL	Soliman	2008	Change in Noncurrent Operating Liab
ChNNCOA	Soliman	2008	Change in Net Noncurrent Op Assets
ChPM	Soliman	2008	Change in Profit Margin
ConsNegRet	Watkins	2003	Consistently negative return
ConsRecomm	Barber et al.	2002	Consensus Recommendation
ConvDebt	Valta	2016	Convertible debt indicator
Coskewness	Harvey and Siddique	2000	Coskewness
CredRatDG	Dichev and Piotroski	2001	Credit Rating Downgrade
currat	Ou and Penman	1989	Current Ratio
CustomerMomentum	Cohen and Frazzini	2008	Customer momentum
DelayAcct	Callen, Khan and Lu	2013	Accounting component of price delay
DelayNonAcct	Callen, Khan and Lu	2013	Non-accounting component of price delay
DelBreadth	Chen, Hong and Stein	2002	Breadth of ownership
DelSTI	Richardson et al.	2005	Change in short-term investment
depr	Holthausen and Larcker	1992	Depreciation to PPE
DivInit	Michaely, Thaler and Womack	1995	Dividend Initiation
DivOmit	Michaely, Thaler and Womack	1995	Dividend Omission
DivSeason	Hartzmark and Salomon	2013	Dividend seasonality
DivYield	Naranjo, Nimalendran, Ryngaert	1998	Dividend yield for small stocks
DivYieldAnn	Naranjo, Nimalendran, Ryngaert	1998	Last year's dividends over price
DivYieldST	Litzenberger and Ramaswamy	1979	Predicted div yield next month
dNoa	Hirshleifer, Hou, Teoh, Zhang	2004	change in net operating assets
DolVol	Brennan, Chordia, Subra	1998	Past trading volume
EarningsConsistency	Alwathainani	2009	Earnings consistency
EarningsForecastDisparity	Da and Warachka	2011	Long-vs-short EPS forecasts
EarningsPredictability	Francis, LaFond, Olsson, Schipper	2004	Earnings Predictability
EarningsSmoothness	Francis, LaFond, Olsson, Schipper	2004	Earnings Smoothness
EarningsSurprise	Foster, Olsen and Shevlin	1984	Earnings Surprise
EarningsValueRelevance	Francis, LaFond, Olsson, Schipper	2004	Value relevance of earnings
EarnSupBig	Hou	2007	Earnings surprise of big firms
EBM	Penman, Richardson and Tuna	2007	Enterprise component of BM
EBM_q	Penman, Richardson and Tuna	2007	Enterprise component of BM
EP	Basu	1977	Earnings-to-Price Ratio
EPq	Basu	1977	Earnings-to-Price Ratio
ExchSwitch	Dharan and Ikenberry	1995	Exchange Switch
FEPS	Cen, Wei, and Zhang	2006	Analyst earnings per share
fgr5yrLag	La Porta	1996	Long-term EPS forecast

FirmAge	Barry and Brown	1984	Firm age based on CRSP
FR	Franzoni and Marin	2006	Pension Funding Status
FRbook	Franzoni and Marin	2006	Pension Funding Status
Frontier	Nguyen and Swanson	2009	Efficient frontier index
Governance	Gompers, Ishii and Metrick	2003	Governance Index
GP	Novy-Marx	2013	gross profits / total assets
GPlag	Novy-Marx	2013	gross profits / total assets
GrAdExp	Lou	2014	Growth in advertising expenses
GrSaleToGrOverhead	Abarbanell and Bushee	1998	Sales growth over overhead growth
GrSaleToGrReceivables	Abarbanell and Bushee	1998	Change in sales vs change in receiv
Herf	Hou and Robinson	2006	Industry concentration (sales)
HerfAsset	Hou and Robinson	2006	Industry concentration (assets)
HerfBE	Hou and Robinson	2006	Industry concentration (equity)
High52	George and Hwang	2004	52 week high
IdioVol3F	Ang et al.	2006	Idiosyncratic risk (3 factor)
IdioVolAHT	Ali, Hwang, and Trombley	2003	Idiosyncratic risk (AHT)
Illiquidity	Amihud	2002	Amihud's illiquidity
IndIPO	Ritter	1991	Initial Public Offerings
IndMom	Grinblatt and Moskowitz	1999	Industry Momentum
IndRetBig	Hou	2007	Industry return of big firms
IntanCFP	Daniel and Titman	2006	Intangible return using CFtoP
IntanEP	Daniel and Titman	2006	Intangible return using EP
IntanSP	Daniel and Titman	2006	Intangible return using Sale2P
IntrinsicValue	Frankel and Lee	1998	Intrinsic or historical value
invest	Chen and Zhang	2010	Capex and Inventory Change
Investment	Titman, Wei and Xie	2004	Investment to revenue
IO_ShortInterest	Asquith Pathak and Ritter	2005	Inst own among high short interest
KZ	Lamont, Polk and Saa-Requejo	2001	Kaplan Zingales index
LaborforceEfficiency	Abarbanell and Bushee	1998	Laborforce efficiency
Leverage_q	Bhandari	1988	Market leverage quarterly
MaxRet	Bali, Cakici, and Whitelaw	2010	Maximum return over month
MeanRankRevGrowth	Lakonishok, Shleifer, Vishny	1994	Revenue Growth Rank
Mom12m	Jegadeesh and Titman	1993	Momentum (12 month)
Mom12mOffSeason	Heston and Sadka	2008	Momentum without the seasonal part
Mom6mJunk	Avramov et al	2007	Junk Stock Momentum
MomOffSeason	Heston and Sadka	2008	Off season long-term reversal
MomOffSeason06YrPlus	Heston and Sadka	2008	Off season reversal years 6 to 10
MomOffSeason11YrPlus	Heston and Sadka	2008	Off season reversal years 11 to 15
MomOffSeason16YrPlus	Heston and Sadka	2008	Off season reversal years 16 to 20
MomRev	Chan and Ko	2006	Momentum and LT Reversal
MomSeason	Heston and Sadka	2008	Return seasonality years 2 to 5
MomSeasonShort	Heston and Sadka	2008	Return seasonality last year
MomVol	Lee and Swaminathan	2000	Momentum in high volume stocks

MRreversal	De Bondt and Thaler	1985	Medium-run reversal
nanalyst	Elgers, Lo and Pfeiffer	2001	Number of analysts
NetDebtFinance	Bradshaw, Richardson, Sloan	2006	Net debt financing
NetDebtPrice	Penman, Richardson and Tuna	2007	Net debt to price
NetDebtPrice_q	Penman, Richardson and Tuna	2007	Net debt to price
NetEquityFinance	Bradshaw, Richardson, Sloan	2006	Net equity financing
NetPayoutYield	Boudoukh et al.	2007	Net Payout Yield
NetPayoutYield_q	Boudoukh et al.	2007	Net Payout Yield quarterly
NOA	Hirshleifer et al.	2004	Net Operating Assets
NumEarnIncrease	Loh and Warachka	2012	Earnings streak length
OperProf	Fama and French	2006	operating profits / book equity
OperProfLag	Fama and French	2006	operating profits / book equity
OperProfRDLagAT	Ball et al.	2016	Oper prof R&D adj lagged assets
OPLeverage	Novy-Marx	2010	Operating leverage
OptionVolume1	Johnson and So	2012	Option to stock volume
OrderBacklog	Rajgopal, Shevlin, Venkatachalam	2003	Order backlog
OrgCap	Eisfeldt and Papanikolaou	2013	Organizational capital
OScore	Dichev	1998	O Score
PatentsRD	Hirschleifer, Hsu and Li	2013	Patents to RD expenses
pchcurrat	Ou and Penman	1989	Change in Current Ratio
pchdepr	Holthausen and Larcker	1992	Change in depreciation to PPE
pchgm_pchsale	Abarbanell and Bushee	1998	Change in gross margin vs sales
pchquick	Ou and Penman	1989	Change in quick ratio
PM	Soliman	2008	Profit Margin
PM_q	Soliman	2008	Profit Margin
Price	Blume and Husic	1972	Price
PriceDelayRsq	Hou and Moskowitz	2005	Price delay r square
PriceDelaySlope	Hou and Moskowitz	2005	Price delay coeff
PriceDelayTstat	Hou and Moskowitz	2005	Price delay SE adjusted
ProbInformedTrading	Easley, Hvidkjaer and O'Hara	2002	Probability of Informed Trading
PS	Piotroski	2000	Piotroski F-score
PS_q	Piotroski	2000	Piotroski F-score
quick	Ou and Penman	1989	Quick ratio
RD_q	Chan, Lakonishok and Sougiannis	2001	R&D over market cap quarterly
rd_sale	Chan, Lakonishok and Sougiannis	2001	R&D to sales
RDAbility	Cohen, Diether and Malloy	2013	R&D ability
RDcap	Li	2011	R&D capital-to-assets
RDIPO	Gou, Lev and Shi	2006	IPO and no R&D spending
realestate	Tuzel	2010	Real estate holdings
ResidualMomentum	Blitz, Huij and Martens	2011	Momentum based on FF3 residuals
ResidualMomentum6m	Blitz, Huij and Martens	2011	6 month residual momentum
retConglomerate	Cohen and Lou	2012	Conglomerate return
RetNOA	Soliman	2008	Return on Net Operating Assets

RetNOA_q	Soliman	2008	Return on Net Operating Assets
ReturnSkew	Bali, Engle and Murray	2015	Return skewness
ReturnSkew3F	Bali, Engle and Murray	2015	Idiosyncratic skewness (3F model)
ReturnSkewQF	Bali, Engle and Murray	2015	Idiosyncratic skewness (Q model)
REV6	Chan, Jegadeesh and Lakonishok	1996	Earnings forecast revisions
RevenueSurprise	Jegadeesh and Livnat	2006	Revenue Surprise
RIO_Turnover	Nagel	2005	Inst Own and Turnover
RIO_Volatility	Nagel	2005	Inst Own and Idio Vol
RoE	Haugen and Baker	1996	net income / book equity
roic	Brown and Rowe	2007	Return on invested capital
salerec	Ou and Penman	1989	Sales to receivables
secured	Valta	2016	Secured debt
securedind	Valta	2016	Secured debt indicator
sfe	Elgers, Lo and Pfeiffer	2001	Earnings Forecast to price
ShareIss1Y	Pontiff and Woodgate	2008	Share issuance (1 year)
ShareIss5Y	Daniel and Titman	2006	Share issuance (5 year)
ShareRepurchase	Ikenberry, Lakonishok, Vermaelen	1995	Share repurchases
ShortInterest	Dechow et al.	2001	Short Interest
sinAlgo	Hong and Kacperczyk	2009	Sin Stock (selection criteria)
sinOrig	Hong and Kacperczyk	2009	Sin Stock (original list)
Size	Banz	1981	Size
skew1	Xing, Zhang and Zhao	2010	Volatility smirk near the money
Spinoff	Cusatis, Miles and Woolridge	1993	Spinoffs
std_turn	Chordia, Subra, Anshuman	2001	Share turnover volatility
STreversal	Jegadeesh	1989	Short term reversal
tang	Hahn and Lee	2009	Tangibility
tang_q	Hahn and Lee	2009	Tangibility quarterly
Tax_q	Lev and Nissim	2004	Taxable income to income (qtrly)
UpRecomm	Barber et al.	2002	Up Forecast
VarCF	Haugen and Baker	1996	Cash-flow to price variance
VolMkt	Haugen and Baker	1996	Volume to market equity
VolSD	Chordia, Subra, Anshuman	2001	Volume Variance
WW	Whited and Wu	2006	Whited-Wu index
WW_Q	Whited and Wu	2006	Whited-Wu index
zerotrade	Liu	2006	Days with zero trades
zerotradeAlt1	Liu	2006	Days with zero trades
ZScore_q	Dichev	1998	Altman Z-Score quarterly

B Appendix: Sorting Dates and Shock Dates

The behavioral models described in Section 2.1.3 all assume a shock date when mispricing initially arises and study how alphas or prices evolve after the shock date. In contrast, I study how alphas evolve after the sorting dates, which correspond to the dates when firms are sorted based on a specific characteristic. In this section, I discuss how sorting dates are related to shock dates.

First, while the exact relationship between shock dates and sorting dates may not be directly observed, these models can still imply a non-constant alpha dynamic after sorting dates. It is because, in those models, the alpha dynamic is non-constant over any subset of periods before mispricing is completely resolved as shown in Figure 2. If a characteristic is associated with remaining mispricing at sorting dates ($j = 0$), then the alpha dynamic after sorting α_j can be non-constant according to these models.

Further, we may also be able to predict the pattern of the alpha dynamic based on the models. That is, where does $j = 0$ fall in these models? First, in the models that explain anomalies endogenously like Model 1 (Barberis, 2018), there is a clear linkage between shock dates and sorting dates. Further, in exogenous models that do not explain anomalies, exogenous investment opportunities (mispricing) appear on shock dates. For example, Model 2 (Duffie, 2010) and Model 3 (Hendershott et al., 2022) describe how prices evolve when some rational traders do not trade on investment opportunities immediately. Since mispricing is exogenous, we can interpret trading on the anomaly characteristics as investment opportunities and shock dates as sorting dates. That is, $j = 0$ should be the shock date. For instance, if some traders do not adjust their characteristic-sorted portfolios based on firm size immediately at the end of each month, alpha dynamics after sorting dates may display a non-constant pattern based on Models 2 and 3.

C Appendix: Proofs

Proof of Proposition 1. $\alpha^{op}_{k,h} = \frac{1}{h}\sum_{j=k+1}^{k+h} \alpha_j \leq \frac{1}{h} h\alpha_{j^*} = \alpha^{op}_{j^*-1,1}$. ■

Proof of Proposition 2. Since, $foc = \frac{\delta_0(\lambda h+1)e^{-\lambda h}+c-\delta_0}{h^2}$, $foc(h = 1) = \delta 0(\lambda+1)e^{-\lambda}+c-\delta_0$. Then $foc(h = 1) \leq 0 \Leftrightarrow \frac{\lambda+1}{e^{\lambda}} \leq 1 - \frac{c}{\delta_0}$.

As $\delta_0(\lambda h + 1)e^{-\lambda h}$ decreases in h, $foc < 0$ when $h > 1$ if $\frac{\lambda+1}{e^{\lambda}} \leq 1 - \frac{c}{\delta_0}$. Therefore, after-cost alpha $\alpha^{ac}_{k,h}$ strictly decreases in h and $h^* = 1$ when $\frac{\lambda+1}{e^{\lambda}} \leq 1 - \frac{c}{\delta_0}$. ■

Proof of Proposition 3. Take $foc = 0$:

$$\frac{\lambda h^* + 1}{e^{\lambda h^*}} = 1 - \frac{c}{\delta_0} \tag{19}$$

Since $\delta_0(\lambda h + 1)e^{-\lambda h}$ decreases in h, $foc > 0$ when $h < h^*$ and $foc < 0$ when $h > h^*$. Therefore, h^* maximizes $\alpha^{ac}_{k,h}$ and solves the problem. ■

D Appendix: Simulated performance under random variation

In Section 4.2, I show that the benchmark strategy that always takes the number of months skipped $k = 0$ and holding period $h = 1$ (H1) is outperformed by an optimization strategy for many anomalies when after-cost alphas are compared.

One possible concern is whether the improvement in after-cost alphas can be obtained with random variation. This section is to investigate this concern with simulations. The null hypothesis is that H1 is optimal and no strategy can outperform it based on after-cost alphas.

The null hypothesis is examined as follows. I conduct 2,000 simulations. In each simulation, I first simulate a time series of after-cost market returns calibrated to the sample mean and standard deviation of after-cost market returns over the sample period. The sample period is between January 1936 and December 2021.

For each characteristic, I estimate a_d, b_d, and the residual volatility by regressing the actual returns of H1 on the actual after-cost market excess returns (not the simulated returns):

$$r_{b,t} = a_d + b_d r_{m,t} + \epsilon_{b,t} \tag{20}$$

I then generate a simulated time series of returns of the benchmark strategy $r_{b,t}$ for each characteristic with simulated market returns. The date range of the simulated returns for a characteristic match the actual returns of the characteristic-sorted portfolios.

Next, within each characteristic, I estimate a_s, b_s, b_m and the residual volatility from the following regression:

$$r_{s,t} = a_s + b_s r_{b,t} + b_m r_{m,t} + \epsilon_{s,t} \tag{21}$$

$r_{s,t}$ are after-cost returns of a strategy that restricts $k = 0$ and takes different $h = 1, 3, 6, 9, 12$ throughout the sample. Returns of each $r_{s,t}$ and $r_{b,t}$ have the same length.

After estimating $\hat{b}_s$, $\hat{b}_m$, and the volatility of $\hat{\epsilon}_{s,t}$ with the actual data, I create simulated time series of returns for the strategies based on $\hat{b}_s$, $\hat{b}_m$, estimated residual volatility, and simulated returns of $r_{b,t}$ and $r_{m,t}$:

$$r_{s,t} = \hat{b}_s r_{b,t} + \hat{b}_m r_{m,t} + \hat{\epsilon}_{s,t} \tag{22}$$

That is, I demean the intercept, a_s. This is to generate a data-generating process that neither strategy outperforms H1. In the meantime, the correlation structure as well as other moments are preserved. Let us denote this data-generating process as DGP_{as}.

In Section 4.2.2, I run another regression that examines whether a strategy generates a higher alpha than a benchmark strategy:

$$r_{s,t} - r_{b,t} = a + b_m r_{m,t} + \epsilon_{s,t} \tag{23}$$

If I use DGP_{as} to generate simulated returns of the other strategies, a problem is that a in Eq 23 may not be zero. As I have demeaned the intercept, a_s, $r_{s,t} - r_{b,t} = a = (\hat{b}_s - 1)r_{b,t}$. That is, a will not be zero with DGP_{as} if $\hat{b}_s - 1 \neq 0$ and the mean of $r_{b,t}$ is not zero. To address this issue, I also simulate returns of the other strategies based on the following DGP, which I denote as DGP_a:

$$r_{s,t} - r_{b,t} = \hat{b}_m r_{m,t} + \hat{\epsilon}_{s,t} \tag{24}$$

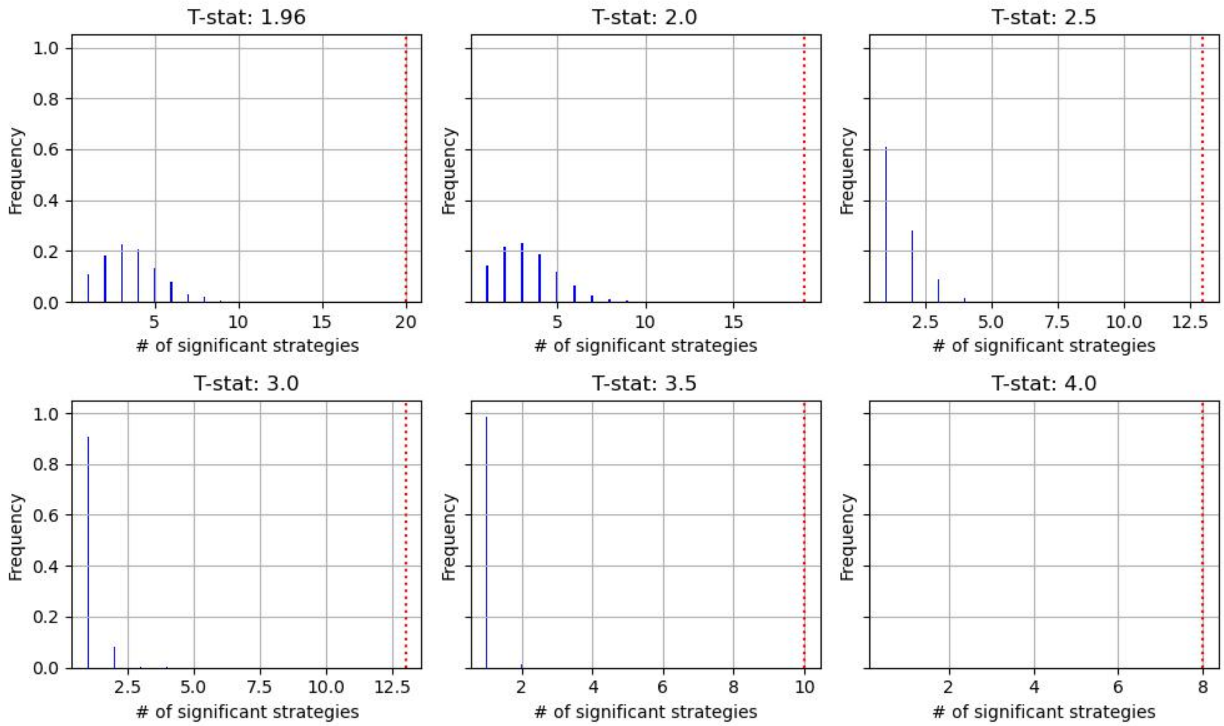
Here, $\hat{b}_m$ and $\hat{\epsilon}_{s,t}$ is generated based on estimated b_m and the volatility of $\epsilon_{s,t}$ in Eq. 23 with the actual data.

Then I use the same procedure as described in Section 4.2 to obtain the return series of the optimization strategy with simulated returns.

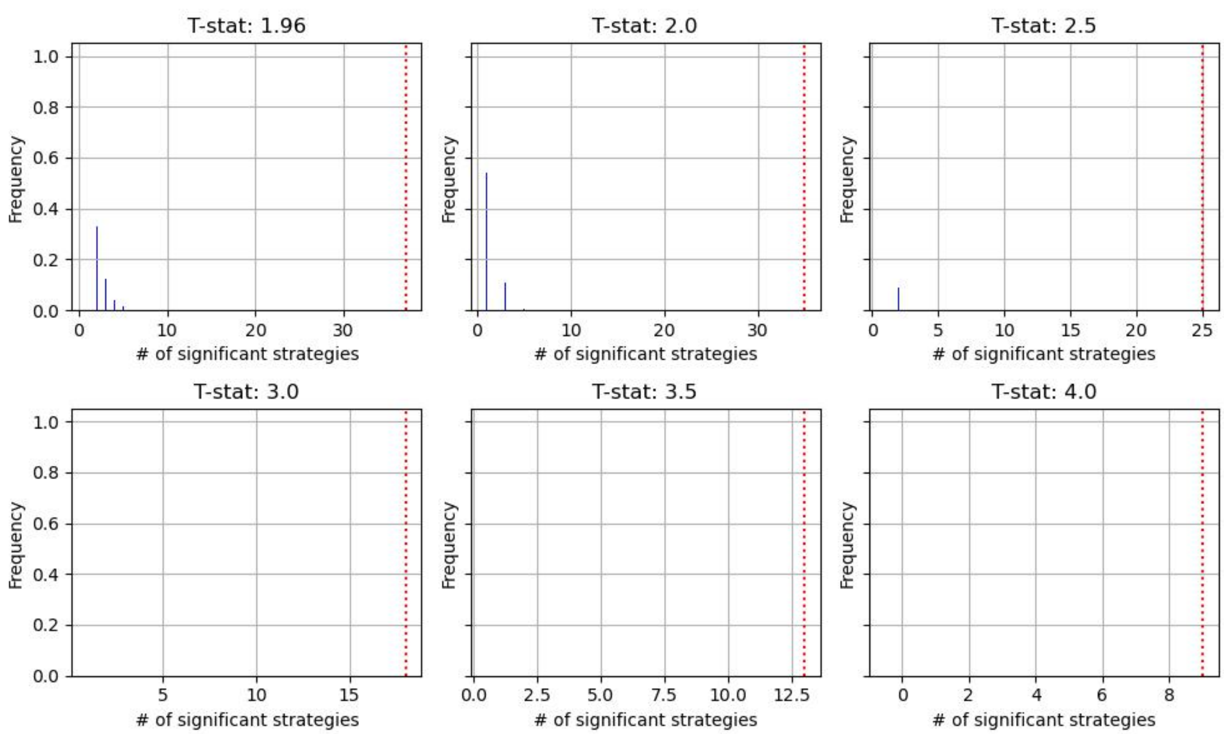
When the optimization strategy uses the simulated returns from DGP_{as}, I run regression Eq. 21 to examine whether the optimization strategy outperforms H1. And when the optimization strategy uses the simulated returns from DGP_a, I run regression Eq. 23 to examine whether the optimization strategy outperforms H1.

A strategy outperforms H1 if a_s or a is statistically significant and positive when the cutoff of t-statistics ranges from 1.96 to 4.00. Then I obtain the counts of characteristics for which the optimization strategy outperforms the benchmark strategy from the regressions under different t-statistic cutoffs. The simulation is repeated 2,000 times so that I obtain a distribution of the counts under different cutoffs.

Results are shown in Figure D1. The red vertical line is the number of characteristics for which the optimization strategy outperforms H1 with the actual data. In subfigure (a), a is the performance metric. And in subfigure (b), a_s is the performance metric. Under t-statistic cutoffs of 1.96, 2.00, 2.50, 3.00, 3.50, and 4.00, the number of characteristics for which the optimization strategy outperforms H1 is always beyond the maximum both when a_s is considered and when a is considered. Therefore, the null hypothesis that H1 is optimal can be rejected.



(a) Peformance metric is a



(b) Peformance metric is a_s

Figure D1: Simulated distributions for the number of strategies that outperforms the benchmark strategy

E Appendix: Additional empirical evidence

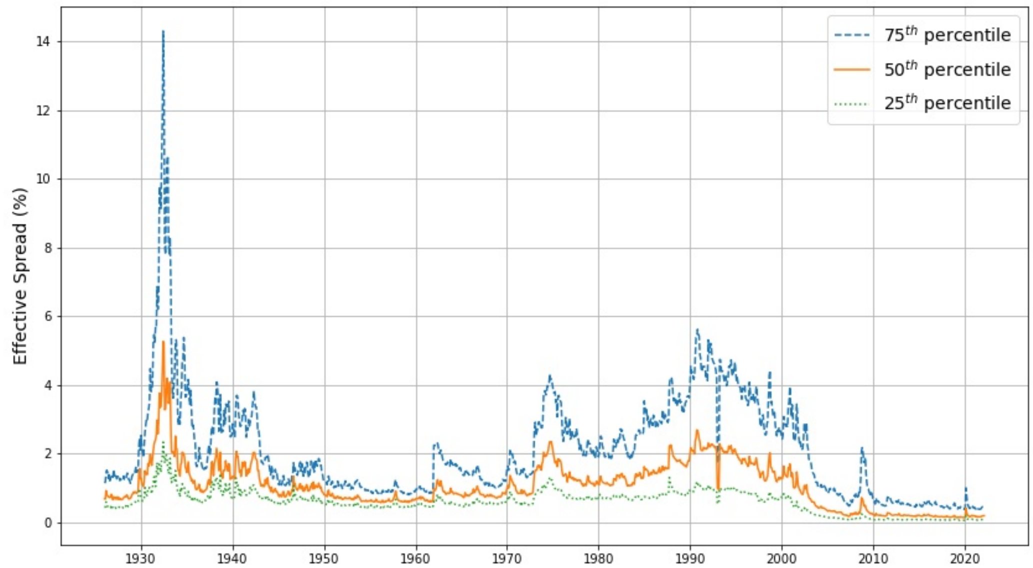


Figure E1: Monthly stock effective spreads over time

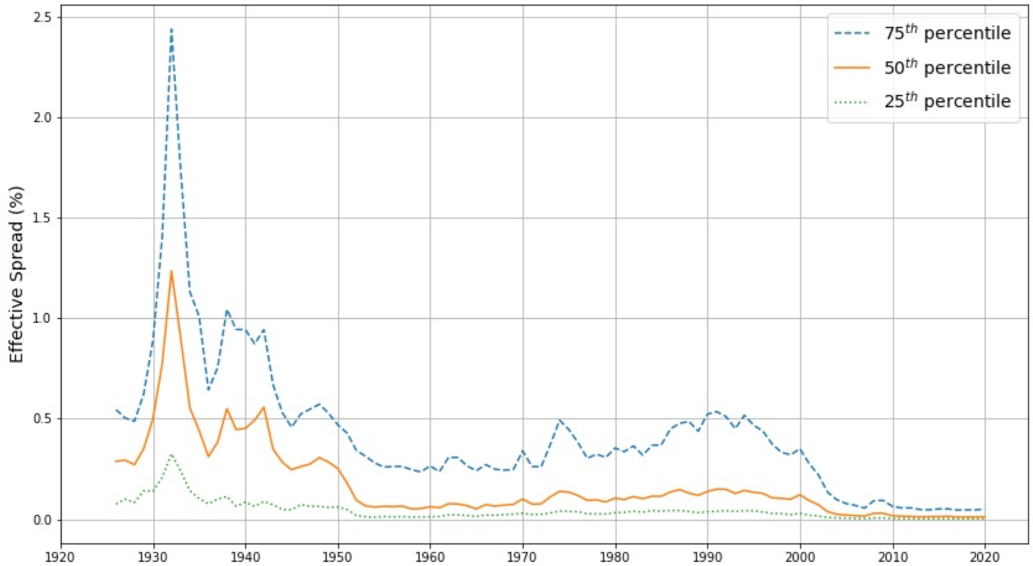


Figure E2: Monthly portfolio rebalancing costs over time

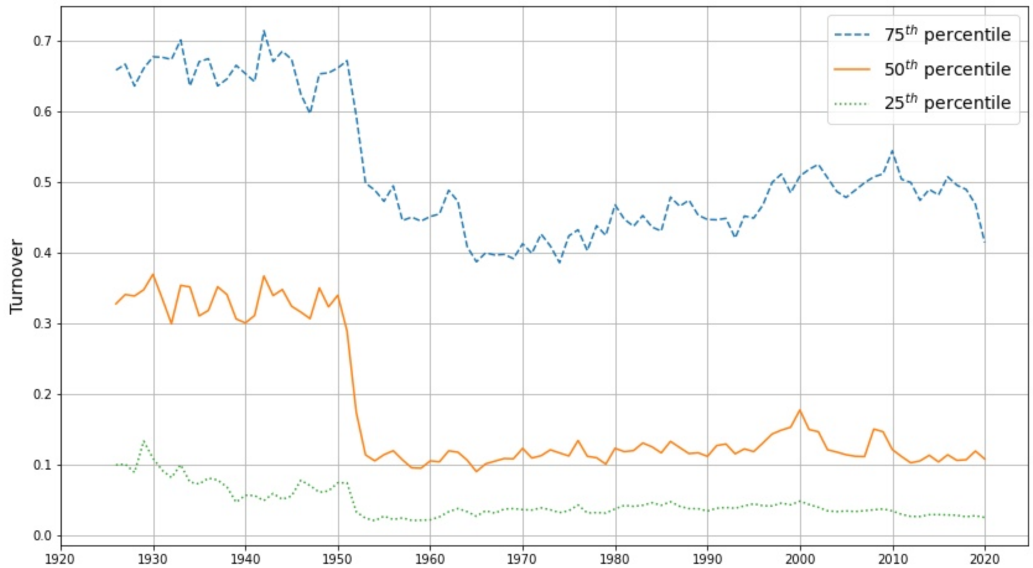


Figure E3: Monthly portfolio turnover rate over time

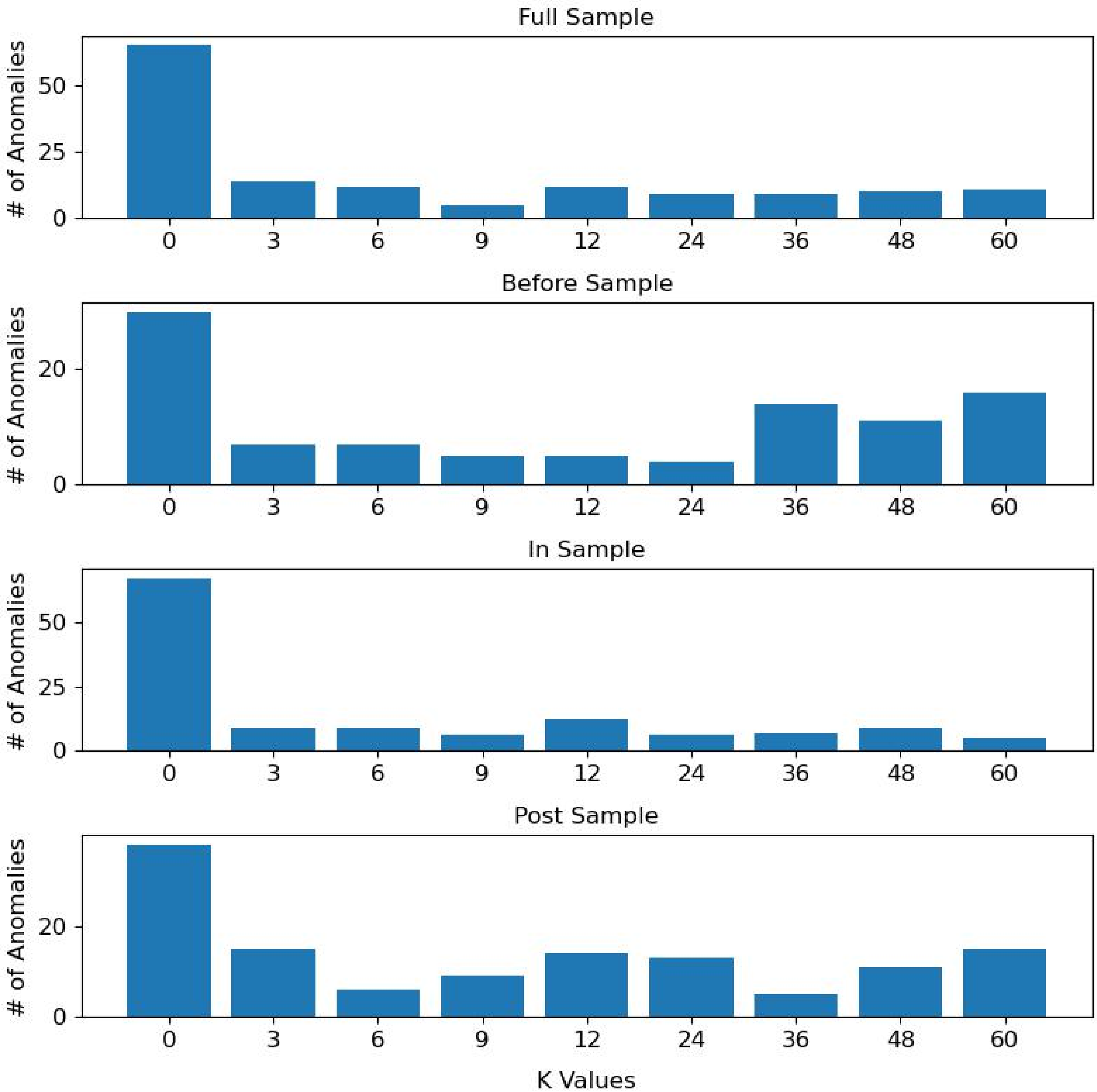


Figure E4: Number of anomalies with the highest alpha at different k values

The figure plots the number of anomalies with the highest CAPM alpha at different k values under an unconditional analysis. I study 205 published anomalies. k is the number of skipped months after a portfolio sort. The holding period is one month. The y axis is the number of anomalies, while the x axis are the k values. From top to bottom, I plot the distributions in the full-sample, before-sample, in-sample, and post-sample periods.

Table E1: Does k have an impact on after-cost profitability?

This table presents the number of anomalies for which the optimization strategy outperforms the default strategies. I compare results for two optimization strategies. The first searches for h only and the second searches for both k and h. The benchmark strategies take either $k = 0$ and $h = 12$ (H12) or $k = 0$ and $h = 1$ (H1) based on after-cost alphas. k is the number of skipped months and h is the holding period. I only consider anomalies that pass either the t test or alpha-dynamic tests in each sample period. In total, I examine 132, 65, 128, and 92 anomalies in the full-sample, before-sample, in-sample, and post-sample periods. Panel A shows results when k is restricted to zero and only optimal h is searched for. Panel B shows results when both optimal k and h are searched for. The performance metrics are a and a_s as in Equation 14 and 15. a examines whether the optimization strategy generates a higher after-cost alpha than a default strategy. And a_s examines whether the returns of the optimization strategy can be completely explained by a default strategy and the market factor.

Panel A: searching for h only

	Default Strategy			
	H12		H1	
Period	a	a_s	a	a_s
Full-sample	30	26	16	33
Before-sample	6	5	4	8
In-sample	30	28	19	29
Post-sample	12	10	6	8

Panel B: searching for both k and h

	Default Strategy			
	H12		H1	
Period	a	a_s	a	a_s
Full-sample	24	24	15	29
Before-sample	3	5	4	9
In-sample	27	30	19	24
Post-sample	20	14	7	9