

Real Time Investor Alphas and the Survival of the Anomaly Zoo

Andrei S. Gonçalves^{*} Johnathan A. Loudis[†] Richard E. Ogden[‡]

This Version: July 2026[§]

Abstract

Anomaly strategies generate positive and significant CAPM alphas even after becoming public information. Common explanations emphasize non-market risks, trading costs, and investment frictions. This paper introduces a complementary channel: allocation uncertainty. Investors who learn about an anomaly remain uncertain about the optimal weight for combining it with the market portfolio, making its future factor regression alpha unattainable. We introduce the real time investor alpha, which measures the Sharpe ratio improvement from adding an anomaly to the market portfolio using weights estimated in real time. Empirically, anomalies retain positive factor regression alphas after publication, but their average real time investor alpha is close to zero. Investors can profitably combine multiple anomalies, but only with shrinkage. We show theoretically that this uncertainty-induced shrinkage makes investors trade less aggressively than full information investors, allowing CAPM alphas to survive in equilibrium. This model explains 15% to 30% of the cross-sectional variation in anomaly alphas.

JEL Classification: G10, G11, G12, G14

Keywords: Anomaly Zoo, Real Time Investor Alphas, CAPM Alphas

^{*}Fisher College of Business, The Ohio State University, Columbus, OH. E-mail: goncalves.11@osu.edu.

[†]Mendoza College of Business, University of Notre Dame, Notre Dame, IN. E-mail: jloudis@nd.edu.

[‡]Rohrer College of Business, Rowan University, Glassboro, NJ. E-mail: ogdenr@rowan.edu.

[§]This paper benefited from helpful comments from Aditya Chaudhry, Alessandro Melone, Carter Davis, Dejanir Silva, Julien Penasse, Leland Bybee, Lukas Kremens, Mirela Sandulescu, Paymon Khorami, Ricardo Delao, Sean Myers, Theis Jensen, Thummim Cho, Xinwei Li, as well as from seminar participants at the 2026 Annual Valuation Workshop, 2025 ASU Sonoran Winter Finance Conference, 2025 International Behavioral Finance Conference at Booth, Notre Dame University, Purdue University, and The Ohio State University. This paper was used as chapter 3 of Richard Ogden's dissertation (publicly available [here](#) since 2023) and previously circulated under the title of "Out-of-Sample Alphas Post-Publication."

Introduction

A central pursuit of asset pricing is to understand the performance of anomaly strategies (i.e., systematic strategies that generate CAPM alpha). The number of papers documenting such strategies has grown significantly over time (Harvey, Liu, and Zhu (2015)). Yet anomaly returns and CAPM alphas remain economically large and statistically significant even after publication, when the anomalies become public information (McLean and Pontiff (2016) and Jensen, Kelly, and Pedersen (2023)). This fact challenges the core idea underlying asset pricing models that smart investors should eliminate alphas in equilibrium, at least after they learn about the existence of such alphas.

Three common explanations for this “survival of the anomaly zoo” are that (i) anomaly strategies are exposed to risks beyond market risk (e.g., Fama and French (1996)), (ii) little alpha remains after trading costs (e.g., Chen and Velikov (2023)), and (iii) frictions prevent smart investors from eliminating alphas (e.g., Shleifer and Vishny (1997)). This paper introduces a complementary channel: even unconstrained rational investors who avoid trading costs and focus only on market risk face an allocation uncertainty challenge that limits their ability to eliminate alphas. When they learn about a new anomaly strategy, they do not know the optimal weight to combine it with the market index, and estimating it in real time yields an expected alpha close to zero. Investors can profit by combining multiple anomalies, but doing so requires shrinkage or regularization, making them less aggressive in eliminating individual alphas. Consequently, the anomaly zoo survives in equilibrium.

Our starting point is as follows. The CAPM alpha of an anomaly strategy reflects the Sharpe ratio gain from optimally combining it with the market index (Gibbons, Ross, and Shanken (1989)). Standard factor regressions implicitly estimate this alpha using ex-post optimal weights (Cederburg et al. (2020)). So, investors learning about a new anomaly in real time would need to know the distribution of its future returns to choose the weight required to earn its future factor regression alpha. Since this is infeasible, we define and estimate real time investor alphas. They measure the future Sharpe ratio gain from combining an anomaly with the market using optimal weights estimated from past information.

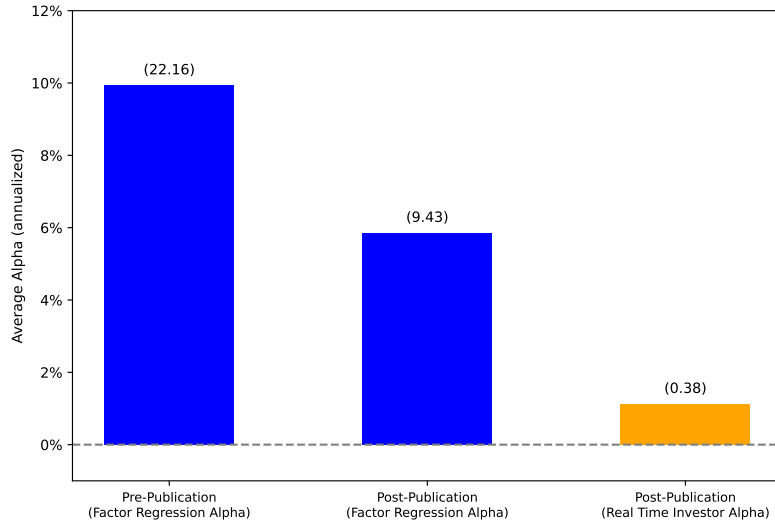


Figure 1
Average CAPM Alphas with their t-stats

The bars plot the average factor regression alpha across anomalies pre- and post-publication as well as the average real time investor alpha post-publication. The bootstrap t-statistic for each average alpha is written on the top of its respective bar. The sample goes from 1926 to 2023 and is based on 177 published equity anomalies from Chen and Zimmermann (2022), with the pre-publication period ending in December of the publication year. Factor regression alphas are based on standard CAPM factor regressions whereas real time investor alphas combine the Gibbons, Ross, and Shanken (1989) expression for alphas with maximum Sharpe ratio weights from the pre-publication period and realized returns from the post-publication period (Equation 8).

Our baseline analysis treats the publication date as the time investors learn about a new anomaly and our main empirical result can be seen in Figure 1. It is based on 177 anomaly long-short decile portfolios obtained from the Chen and Zimmermann (2022) dataset. Consistent with the literature, the first two bars show that anomaly strategies have a substantial average factor regression alpha pre-publication, which declines but remains positive and statistically significant over the post-publication period. The third bar highlights our main result: the average real time investor alpha post-publication is relatively close to zero and statistically insignificant. As such, allocating capital in real time after learning about a new anomaly strategy has not been as rewarding as factor regression alphas suggest.

The key patterns in Figure 1 are highly robust to empirical implementation choices. The average real time investor alpha remains relatively close to zero under alternative approaches for selecting anomalies, different ways for defining the pre- and post-publication periods, and alternative methods for constructing the anomaly strategies. Results are also similar if the

out-of-sample optimal weights account for the expected anomaly decay post-publication or incorporate shrinkage or regularization through Bayesian and machine learning methods.

Importantly, while using the publication date to infer the time investors learn about a new anomaly prevents contaminating our real time investor alphas with publication bias, our argument does not depend on investors learning about anomalies when they are published. Consistent with this point, our empirical results are similar if we use pseudo-publication dates that are based entirely on whether the anomaly past alpha is statistically significant and allow the same anomaly to “be published” multiple times. So, our point remains similar even if investors learn about each anomaly before or after its publication date, and even if investors continue to evaluate the anomaly each period after learning about it.

The above results show that, because of allocation uncertainty, a real time mean-variance investor who currently holds the market does not have an incentive to add an anomaly to his portfolio, and thus would not eliminate CAPM alphas. However, this does not imply that anomalies are collectively unattractive. Previous research shows that investors can achieve Sharpe ratios above the market Sharpe ratio by combining multiple anomalies in real time (e.g., Kozak, Nagel, and Santosh (2020) and Bryzgalova, Pelger, and Zhu (2024)). Similar conclusions hold after accounting for trading costs, publication considerations, and high-dimensional settings in which the number of assets exceeds the number of time observations (Jensen et al. (2025), Lassance and Martín-Utrera (2025), and Hellum et al. (2026)). A common feature of these approaches is their explicit or implicit use of shrinkage, consistent with the well-known instability of unregularized mean-variance optimization (e.g., DeMiguel, Garlappi, and Uppal (2009)). Thus, jointly investing in multiple anomalies can generate positive real time investor alphas, but doing so requires investors to shrink their allocations away from the ex-post optimal weights that would eliminate individual CAPM alphas.

We formalize this idea in a theoretical model in which the anomaly zoo survives entirely because of allocation uncertainty. Specifically, we model a representative Bayesian mean-variance investor who is uncertain about anomaly expected returns. We show that, because this investor endogenously applies shrinkage to his optimal allocations, he responds less

aggressively to variation in expected returns than a mean-variance investor with Full Information Rational Expectations (FIRE) would, in line with Martin and Nagel (2022) and Davis (2025). Consequently, CAPM alphas are not zero in equilibrium.

We next assess the model’s quantitative relevance by estimating its cross-sectional implications for anomaly alphas. The model links factor regression alphas to anomaly betas and volatilities and, under empirically relevant conditions, predicts lower alphas for higher-beta anomalies, providing an alternative mechanism for the low-beta anomaly (Frazzini and Pedersen (2014)). The estimated coefficients have the signs predicted by the model and are statistically significant, and imposing the model’s structural restrictions has little effect on fit. Allocation uncertainty explains around 15% of the variation in alphas across anomaly strategies and a little over 30% across the orthogonal principal component factors that most closely match the model structure. The model achieves this fit with reasonable combinations of risk aversion, expected-return uncertainty, and prior beliefs. However, it tends to imply an excessively high market risk premium, suggesting that it better explains relative differences in anomaly alphas than the overall level of expected returns.

In principle, allocation uncertainty can sustain the anomaly zoo even in a representative-investor economy in which all other CAPM assumptions hold. Empirically, however, it explains only 15% to 30% of the cross-sectional variation in alphas, suggesting that other mechanisms also matter. Introducing additional investors naturally generates such complementary channels. For example, in the presence of noise traders, rational mean-variance investors bear non-fundamental risk, creating a limit to arbitrage that further prevents them from eliminating alphas in equilibrium. Allocation uncertainty is therefore best viewed as a complement to existing explanations for persistent CAPM alphas in the presence of smart investors, rather than as a complete explanation for the survival of the anomaly zoo.

In summary, allocation uncertainty gives investors little incentive to trade individual anomalies in real time and makes jointly trading multiple anomalies profitable only with shrinkage. Yet investors who apply shrinkage are not aggressive enough to eliminate alphas. We argue that this mechanism helps explain why the anomaly zoo survives in equilibrium.

Contribution to the Literature

Our key empirical contribution is to show that the alphas earned by real time investors are substantially lower than the factor regression alphas typically studied in the literature. Central to this contribution is our definition of real time investor alpha, which is based on the Sharpe ratio increase from a portfolio that combines the anomaly and market index using weights estimated in real time. This definition builds on the established insight that the performance of maximum Sharpe ratio portfolios should be evaluated using weights estimated ex ante. For instance, Cederburg et al. (2020) apply this principle to volatility-managed portfolios, Kan, Wang, and Zheng (2024) emphasize that the ex post maximum Sharpe ratio portfolio implied by a factor model is unattainable to investors, and Chabi-Yo, Gonçalves, and Loudis (2025) use this idea to compare their intertemporal factor model with alternative models. However, none of these papers translates this insight into a measure of real time investor alpha or applies it broadly to explain why the anomaly zoo survives in equilibrium. These extensions constitute our contribution relative to this literature.

Our real time investor alpha definition builds on the idea that investors face uncertainty when deciding portfolio weights in real time. This allocation uncertainty concept is heavily studied in the portfolio allocation literature (e.g., Jorion (1986), MacKinlay and Pástor (2000), Kan and Zhou (2007), Garlappi, Uppal, and Wang (2007), DeMiguel, Garlappi, and Uppal (2009), Ledoit and Wolf (2017), Kan, Wang, and Zhou (2022), Kan and Wang (2024), and Kan and Lassance (2025)). Our contribution in this regard is to develop a real time investor alpha measure that accounts for this allocation uncertainty and to apply it broadly to the anomaly zoo. So, these papers study portfolio allocation decisions under uncertainty while we study the equilibrium asset pricing implications of investors incorporating uncertainty in their portfolio allocations.

Relatedly, our real time investor alphas rely on the logic that the intercept estimate from a factor regression reflects a concept of alpha that is achievable only to hypothetical investors with FIRE. Due to weight uncertainty, real investors rely on statistical methods to choose portfolio weights so that they achieve real time investor alphas that are weaker than

the respective factor regression alphas studied in the literature.¹ This connects our work to recent studies on statistical limits to arbitrage (e.g., Martin and Nagel (2022), Da, Nagel, and Xiu (2024), Davis (2025), and Baba Yara, Boyer, and Davis (2024)). Unlike standard limits to arbitrage, which arise from frictions in financial markets (Gromb and Vayanos (2010)), statistical limits to arbitrage arise from statistical difficulties investors without FIRE face when making investment decisions in real time.

Our paper has a particularly close connection to two papers in this statistical limits to arbitrage literature. First, Martin and Nagel (2022) provide a theoretical model where risk-neutral Bayesian investors face a real time high dimensional portfolio allocation problem. They show that if the number of expected return signals is comparable to the number of assets available to investors, equilibrium returns are predictable in-sample even though they are not predictable out-of-sample. Second, Davis (2025) argues that hypothetical statistical arbitrageurs face a similar real time high dimensional portfolio allocation problem, which leads them to have inelastic demand that is quantitatively consistent with demand estimates in the literature (e.g., Koijen and Yogo (2019)). The results in these papers arise from the fact that optimal portfolios are not aggressive in trading anomalies given the high dimensional portfolio allocation problem. We add to these papers by showing that statistical limits to arbitrage are important even in the low dimensional problem of adding a single anomaly to the market portfolio. In particular, we propose a novel real time investor alpha measure and apply it to the anomaly zoo to show that an investor who holds the market index has no incentive to add an anomaly to his portfolio, only to invest in multiple anomalies simultaneously and with shrinkage. Moreover, we incorporate this statistical limit to arbitrage in an equilibrium model to derive an equation mapping factor regression alphas to betas and volatilities, and estimate it to show that this mechanism explains between 15% and 30% of the empirical variation in anomaly alphas.

¹Note that the concept of real time investor alpha we use is different from real time forecasts of future factor regression alphas, which are sometimes studied in the literature (e.g., Jensen, Kelly, and Pedersen (2023) and Marrow and Nagel (2024)). In particular, even if an investor (accurately) forecasts an anomaly to have a positive future CAPM factor regression alpha, this does not imply that the investor knows how much to allocate to the anomaly strategy and market portfolio to earn the anomaly's alpha.

Given our empirical design, our work is also connected to the literature on anomaly performance pre- and post-publication. McLean and Pontiff (2016) show a strong decline in anomaly performance post-publication (see also Chen and Zimmermann (2023), Chen, Lopez-Lira, and Zimmermann (2023), and Pénasse (2022)). However, anomalies continue to provide a positive and statistically significant average factor regression alpha post-publication (Jensen, Kelly, and Pedersen (2023)). In contrast, we find a post-publication average real time investor alpha that is close to zero economically and statistically. So, our evidence complements these papers as investors who trade in real time cannot achieve factor regression alphas, but they can achieve real time investor alphas.

Finally, our work can be understood as a cross-sectional analogue to the time-series point made by Goyal and Welch (2008) (see also Campbell and Thompson (2008), Goyal, Welch, and Zafirov (2024), and Guo and Wachter (2025)). In particular, Goyal and Welch (2008) show that most variables that predict time variation in the equity premium in-sample do not do so out-of-sample, which is a consequence of uncertainty in the predictability parameters estimated in real time. Likewise, we show that anomalies provide a strong average factor regression alpha (which can be thought of as an in-sample alpha), but a close to zero average real time investor alpha (which can be thought of as an out-of-sample alpha). Moreover, in line with the parameter uncertainty point of Goyal and Welch (2008), our result is a consequence of uncertainty in estimating the optimal weights needed to combine anomalies and with the market index in order to harvest their alphas.

The rest of this paper is organized as follows. Section 1 introduces the different alpha concepts we study. Section 2 details our data and alpha estimation strategy. Section 3 empirically compares factor regression alphas with real time investor alphas. Section 4 argues that investors can only achieve a positive real time investor alpha by applying shrinkage while investing in multiple anomalies simultaneously, and demonstrates that investors who do so act less aggressively, not eliminating alphas in equilibrium. Section 5 concludes. The Internet Appendix provides derivations as well as supplementary empirical details and results.

1 Introducing the Different Alpha Concepts

This section introduces the different alpha concepts we build on in this paper. Subsection 1.1 defines the factor regression alpha and the investor alpha from the prior literature, which are mathematically equivalent but capture distinct economic concepts. Subsection 1.2 then introduces our real time investor alpha, which generalizes the investor alpha to reflect real time portfolio allocation decisions by investors (and differs from the real time forecast for the future factor regression alpha). Subsection 1.3 builds on these concepts to explain our key empirical finding that equity anomaly strategies have real time investor alphas that are close to zero on average despite their large factor regression alphas.

Throughout the main text, we rely on the CAPM for simplicity. So, all references to “alpha” refer to the CAPM alpha. However, Internet Appendix A shows that factor regression alphas are also mathematically equivalent to investor alphas in multifactor models, which allows one to generalize our real time investor alpha definition to multifactor models.

1.1 Defining the Factor Regression Alpha and the Investor Alpha

Let $r_{m,t}$ be the return on the market portfolio in excess of the risk-free rate and $r_{a,t}$ be the return on a long-short portfolio of interest (e.g., an active investment strategy or a high minus low anomaly strategy). Also, assume that, for $t > T_0$, they have the joint distribution $(r_{m,t}, r_{a,t}) \sim \mathcal{D}ist(\mu, \Sigma)$, with the relation

$$r_{a,t} = \alpha + \beta \cdot r_{m,t} + \epsilon_t. \quad (1)$$

From Gibbons, Ross, and Shanken (1989), we know that the above projection intercept (α) can be alternatively written as²

$$\alpha = \text{sign}[w] \cdot \sigma[\epsilon] \cdot \sqrt{|\text{SR}[r_p]^2 - \text{SR}[r_m]^2|} \quad (2)$$

²Equation 2 is a univariate version of the result in Gibbons, Ross, and Shanken (1989) linking pricing errors to mean-variance efficiency (see also the textbook treatment in Chapter 6.6 of Campbell, Lo, and MacKinlay (1997)). Internet Appendix A proves a generalization of Equation 2 to multifactor models.

where $\text{SR}[x] = \mathbb{E}[x]/\sigma[x]$ is the Sharpe ratio function and $r_{p,t} = (1 - w) \cdot r_{m,t} + w \cdot r_{a,t}$ is the maximum Sharpe ratio portfolio combining r_m and r_a so that³

$$\omega = \begin{pmatrix} 1 - w \\ w \end{pmatrix} = \frac{\Sigma^{-1}\mu}{|1'\Sigma^{-1}\mu|}. \quad (3)$$

We refer to the intercept from the projection in Equation 1 as the “factor regression alpha” since it is obtained from a factor regression. Its economic content comes from the fact that it reflects the expected return from a beta-neutral portfolio that shorts β dollars of r_m for each one dollar invested in r_a (with the CAPM predicting it to be zero).

However, this beta-neutral portfolio is not the optimal risky portfolio the mean-variance investor underlying the CAPM would want to hold. Rather, that investor would want to hold the maximum Sharpe ratio portfolio, r_p . Since the mean-variance utility gain from doing so is proportional to $\text{SR}[r_p]^2 - \text{SR}[r_m]^2$, the alpha from Equation 2 reflects the (normalized) economic benefit that such mean-variance investor earns by moving from investing in r_m to investing in r_p (i.e., combining r_a with r_m).⁴ So, we refer to it as the investor alpha.

Gibbons, Ross, and Shanken (1989) show that the investor alpha is mathematically equivalent to the factor regression alpha in population terms and in any given sample. In the next subsection, we highlight that this equivalence no longer holds when w needs to be estimated in real time, and we use this difference to define the real time investor alpha.

³There are infinitely many maximum Sharpe ratio portfolios, all satisfying $\omega = \theta \cdot \Sigma^{-1}\mu$ for $\theta > 0$. The choice of θ does not affect the calculated alphas, as $\text{SR}[r_p]$ and $\text{sign}(w)$ remain constant for any $\theta > 0$ (see Equation 2). We use the normalization $\theta = 1/|1'\Sigma^{-1}\mu|$, ensuring the absolute sum of ω elements equals one. Normalizing by $\theta = 1/(1'\Sigma^{-1}\mu)$, which forces ω elements to sum to one, would be incorrect if $1'\Sigma^{-1}\mu < 0$, as it would yield the minimum Sharpe ratio portfolio. Although rare in our empirical analysis, this is possible if the global minimum variance portfolio has a negative expected excess return. In this case, the tangency portfolio lies in the inefficient frontier, and the maximum Sharpe ratio portfolio involves a negative position in the tangency portfolio and a positive position in the risk-free asset.

⁴In our terminology, expressions such as “to earn” or “to achieve” the alpha refer to forming a portfolio that combines the market with the anomaly strategy to achieve the Sharpe ratio increase embedded in the given alpha expression (e.g., Equation 2). This is more appropriate for our analysis than to define “to achieve the alpha” as forming the beta-neutral strategy $r_{a,t} - \beta \cdot r_{m,t}$ (which has expected return equal to the factor regression alpha) since the CAPM mean-variance investor aims at forming the maximum Sharpe ratio portfolio (to combine with the risk-free asset). So, an investor who knows the future β would still need to estimate the optimal weights to combine $r_{a,t} - \beta \cdot r_{m,t}$ with $r_{m,t}$ in order to maximize Sharpe ratio. We return to this issue in Subsection 3.3.

1.2 Defining the Real Time Investor Alpha

To earn the investor alpha, one needs to know μ and Σ in order to calculate w . However, actual investors do not know μ and Σ in real time. So, we define the real time investor alpha to be analogous to the investor alpha, except that it does not require the investor to know w before investing. Specifically, letting $\mathcal{R}$ reflect an information set containing data only up to time T_0 , we define the real time investor alpha as

$$\alpha_{\mathcal{R}} = \text{sign}[\Delta] \cdot \sigma[\epsilon] \cdot \sqrt{|\text{SR}[r_{\mathcal{R},p}]^2 - \text{SR}[r_m]^2|}, \quad (4)$$

where $r_{\mathcal{R},p,t} = (1 - w_{\mathcal{R}}) \cdot r_{m,t} + w_{\mathcal{R}} \cdot r_{a,t}$ and $\Delta = \text{SR}[r_{\mathcal{R},p}] - \text{SR}[r_m]$, with $w_{\mathcal{R}} = \mathbb{E}[w|\mathcal{R}]$.

Note that the only term investors need to know in real time in order to earn the alpha is w . So, the real time investor alpha replaces the true unknown w with $w_{\mathcal{R}} = \mathbb{E}[w|\mathcal{R}]$, which reflects the best estimate investors can have (under $\mathcal{R}$) for how much they should allocate to r_a in order to maximize Sharpe ratio. The real time investor alpha does not replace other terms (e.g., $\text{SR}[\cdot]$) since investors do not need to know them in advance in order to form their optimal portfolios.

Consequently, $\alpha_{\mathcal{R}} \neq \mathbb{E}[\alpha|\mathcal{R}]$ so that the real time investor alpha is not the real time forecast for the future factor regression alpha sometimes studied in the literature (e.g., Jensen, Kelly, and Pedersen (2023) and Marrow and Nagel (2024)). Rather, it is the alpha rational investors would be expected to earn when making portfolio allocation decisions in real time based on $\mathcal{R}$. So, papers that predict factor regression alphas in real time are not estimating the real time investor alpha, $\alpha_{\mathcal{R}}$, but rather the expected future factor regression alpha, $\mathbb{E}[\alpha|\mathcal{R}]$. Moreover, knowing $\mathbb{E}[\alpha|\mathcal{R}]$ is not sufficient to obtain $w_{\mathcal{R}} = \mathbb{E}[w|\mathcal{R}]$, which is what investors need to make portfolio allocation decisions in order to earn $\alpha_{\mathcal{R}}$.

It is instructive to consider a parallel to the subjective beliefs literature. Under full information rational expectations (FIRE), investors know all parameters of the return data generating process. So, the FIRE $\mathcal{R}$ includes μ and Σ so that $w_{\mathcal{R}} = \mathbb{E}[w|\mathcal{R}] = w$. Consequently, the investor alpha in Equation 2 (which matches the factor regression alpha) can be thought of as the FIRE investor alpha. From this perspective, we can think of our real

time investor alpha definition as recognizing the economic reality that actual investors do not have FIRE.

This FIRE interpretation also clarifies the other change from Equation 2 to Equation 4: the replacement of $sign[w]$ with $sign[\Delta]$. We have $sign[w] = sign[\alpha]$ for the FIRE investor because a positive (negative) factor regression alpha necessarily leads to a positive (negative) FIRE allocation in the maximum Sharpe ratio portfolio. However, this is no longer true when the investor does not have FIRE. In particular, since $\mathbb{E}[w|\mathcal{R}]$ is not necessarily optimal from a Sharpe ratio perspective, the investor may end up allocating too much or too little to r_a to the point that he earns a Sharpe ratio that is lower than what he could have earned by investing in the market alone. So, we replace $sign[w]$ with $sign[\Delta]$ so that the real time investor alpha is positive if and only if r_a helps the investor to achieve a higher Sharpe ratio than what is provided by the market.⁵

1.3 Understanding our Key Empirical Finding

Our key empirical finding is that the real time investor alphas of equity anomaly strategies are much lower than their FIRE investor alphas (i.e., factor regression alphas). Figure 2 provides a mean-variance diagram in excess return space that captures the essence of this empirical finding from a conceptual point of view (assuming $w > 0$ for simplicity). The red dot reflects the market portfolio (r_m) while the green dot reflects the equity anomaly strategy (r_a). The blue dot captures the maximum Sharpe ratio portfolio that can be formed by combining the market portfolio with the anomaly strategy, $r_p = (1 - w) \cdot r_{m,t} + w \cdot r_{a,t}$. Using the FIRE weight (w) maximizes the Sharpe ratio (i.e., line slope) increase from the red dot to the blue dot. As such, using the real time weight ($w_{\mathcal{R}}$) leads to a lower Sharpe ratio increase whether $w_{\mathcal{R}} > w$ or $w_{\mathcal{R}} < w$, as can be seen by comparing the slopes of the blue and orange lines. So, the empirical question we tackle is not whether alpha decreases as we move from α to $\alpha_{\mathcal{R}}$, but rather by how much it decreases and how high $\alpha_{\mathcal{R}}$ is. Next section details the data

⁵Note that our empirical analysis uses a normalization so that each r_a has a positive α over $t \leq T_0$. This normalization simplifies our exposition because it implies a “successful” r_a has a positive $\alpha_{\mathcal{R}}$ over $t > T_0$.

and estimation procedure we use to measure α and $\alpha_{\mathcal{R}}$ in our baseline empirical analysis.

2 The Alphas of Equity Anomaly Strategies: Data and Estimation

This section details the data and estimation for the alphas of equity anomaly strategies we rely on in our analysis. Subsection 2.1 focuses on the estimation while Subsection 2.2 describes the data and provides basic summary statistics.

2.1 The Estimation Procedure

Suppose we observe $r_{m,t}$ and $r_{a,t}$ for $t = 1, 2, \dots, T_0, \dots, T$. We can consistently estimate the factor regression alpha over $t > T_0$ from an Ordinary Least Squares regression on Equation 1, which is equivalent to

$$\hat{\alpha} = \text{sign}[w_{IS}] \cdot \hat{\sigma}[\epsilon] \cdot \sqrt{|\widehat{\text{SR}}[r_p]^2 - \widehat{\text{SR}}[r_m]^2|}, \quad (5)$$

where $r_{p,t} = (1 - w_{IS}) \cdot r_{m,t} + w_{IS} \cdot r_{a,t}$ and

$$\omega_{IS} = \begin{pmatrix} 1 - w_{IS} \\ w_{IS} \end{pmatrix} = \frac{\Sigma_{IS}^{-1} \mu_{IS}}{|1' \Sigma_{IS}^{-1} \mu_{IS}|}, \quad (6)$$

with μ_{IS} and Σ_{IS} obtained using data over $t > T_0$:

$$\mu_{IS} = \begin{bmatrix} \frac{\sum_{t=T_0+1}^T r_{m,t}}{T-T_0} \\ \frac{\sum_{t=T_0+1}^T r_{a,t}}{T-T_0} \end{bmatrix} \text{ and } \Sigma_{IS} = \begin{bmatrix} \frac{\sum_{t=T_0+1}^T (r_{m,t} - \bar{r}_m)^2}{T-T_0-1} & \frac{\sum_{t=T_0+1}^T (r_{m,t} - \bar{r}_m) \cdot (r_{a,t} - \bar{r}_a)}{T-T_0-1} \\ \frac{\sum_{t=T_0+1}^T (r_{m,t} - \bar{r}_m) \cdot (r_{a,t} - \bar{r}_a)}{T-T_0-1} & \frac{\sum_{t=T_0+1}^T (r_{a,t} - \bar{r}_a)^2}{T-T_0-1} \end{bmatrix} \quad (7)$$

Our notation attempts to distinguish between two aspects of this estimation process. First, Sharpe ratios ($\text{SR}[\cdot]$) and idiosyncratic volatility ($\sigma[\epsilon]$) need to be estimated and we use hats to denote the respective ex-post moment estimates, which use data over $t > T_0$. Second, the weight (w) the investor needs to allocate to the anomaly strategy in order to achieve the given alpha also needs to be estimated, and we use the subscript IS (denoting in-sample) to identify quantities needed to construct the weight estimate (w_{IS}). The reason for this distinction is that for investors to achieve $\hat{\alpha}$ they need to know w_{IS} at $t = T_0$, but not the quantities denoted with hats, which they can observe ex-post at $t = T$.

Consequently, the use of data over $t > T_0$ to obtain w_{IS} makes $\hat{\alpha}$ not achievable to investors making decisions in real time. However, $\hat{\alpha}$ is still an interesting economic object as it provides a consistent estimate for the factor regression alpha, which is also the FIRE investor alpha (as discussed in the previous section).

In contrast to $\hat{\alpha}$, the estimation of the real time investor alpha requires an out-of-sample (OS) estimate for w (i.e., a w estimate that uses data over $t \leq T_0$). For this purpose, we use

$$\hat{\alpha}_{\mathcal{R}} = \text{sign}[\hat{\Delta}] \cdot \hat{\sigma}[\epsilon] \cdot \sqrt{\left| \widehat{\text{SR}}[r_{\mathcal{R},p}]^2 - \widehat{\text{SR}}[r_m]^2 \right|} \quad (8)$$

where $\hat{\Delta} = \widehat{\text{SR}}[r_{\mathcal{R},p}] - \widehat{\text{SR}}[r_m]$, $r_{\mathcal{R},p,t} = (1 - w_{OS}) \cdot r_{m,t} + w_{OS} \cdot r_{a,t}$, and

$$\omega_{OS} = \begin{pmatrix} 1 - w_{OS} \\ w_{OS} \end{pmatrix} = \frac{\Sigma_{OS}^{-1} \mu_{OS}}{|1' \Sigma_{OS}^{-1} \mu_{OS}|}, \quad (9)$$

with μ_{OS} and Σ_{OS} obtained using data over $t \leq T_0$:

$$\mu_{OS} = \begin{bmatrix} \frac{\sum_{t=1}^{T_0} r_{m,t}}{T_0} \\ \frac{\sum_{t=1}^{T_0} r_{a,t}}{T_0} \end{bmatrix} \text{ and } \Sigma_{OS} = \begin{bmatrix} \frac{\sum_{t=1}^{T_0} (r_{m,t} - \bar{r}_m)^2}{T_0 - 1} & \frac{\sum_{t=1}^{T_0} (r_{m,t} - \bar{r}_m) \cdot (r_{a,t} - \bar{r}_a)}{T_0 - 1} \\ \frac{\sum_{t=1}^{T_0} (r_{m,t} - \bar{r}_m) \cdot (r_{a,t} - \bar{r}_a)}{T_0 - 1} & \frac{\sum_{t=1}^{T_0} (r_{a,t} - \bar{r}_a)^2}{T_0 - 1} \end{bmatrix} \quad (10)$$

so that the resulting w_{OS} is known to investors at $t = T_0$. Consequently, α_{OS} is achievable to investors who make their portfolio allocation decision at $t = T_0$.

2.2 Data and Summary Statistics

Obtaining $\hat{\alpha}$ and $\hat{\alpha}_{\mathcal{R}}$ requires us to decide:

1. How to select anomaly signals for the analysis
2. How to form anomaly strategies (long-short portfolios) for the anomaly signals selected
3. How to specify the weight estimation period (i.e., $t \leq T_0$) and the alpha evaluation period (i.e., $t > T_0$)

For all three decisions, we rely on the Open Source Asset Pricing (OSAP) dataset from Chen and Zimmermann (2022).⁶ In terms of decision (1), we select all anomaly signals in

⁶We use the August 2024 version of the OSAP dataset from <https://www.openassetpricing.com/>.

OSAP that are labeled as “predictors” by its signal documentation. In regards to decision (2), we form each anomaly strategy as the long-short decile portfolio with weights as per the original publication. The decile portfolio return data come directly from the “PredictorAlt-Ports_Deciles” file of OSAP, which uses stock weighting schemes from the original papers (typically equal- or value-weighting). Finally, for decision (3), our sample is monthly and $t = T_0$ reflects December of the publication year for the given anomaly. Also, $t = 1$ reflects the first month in the original publication sample or the first month of anomaly strategy return available in the OSAP dataset (whichever is later) and $t = T$ reflects the latest month with anomaly strategy returns available in the OSAP dataset (December 2023 for the vast majority of anomaly strategies).⁷ Additionally, we require anomaly strategies to have at least 5 years of monthly returns available both before and after the publication year and normalize the factor regression alpha ($\hat{\alpha}$) over the pre-publication period to be positive (so that we have $w_{OS} > 0$). This normalization is achieved by flipping the long and short positions underlying $r_{a,t}$ whenever the pre-publication $\hat{\alpha}$ of the anomaly strategy is negative.

It is important to highlight that we use the publication date to define T_0 only because this allows us to have an alpha evaluation period that is free of issues associated with publication bias (and to further link our empirical findings to the literature on the post-publication decline in alphas). However, none of our conceptual arguments depends on the publication date of anomalies. Once an investor learns about an anomaly (whether before, at, or after its publication date), that investor faces the challenge of whether to add that anomaly to his portfolio. In fact, the investor continues to face the same challenge every subsequent period. So, our point applies more broadly. One of our robustness checks in Section 3.2 explicitly accounts for that by considering pseudo publication dates that depend only on whether the

⁷Note that we start at the first month in the original publication sample even when anomaly strategy returns are available in the OSAP dataset for prior months (which is the case for the vast majority of anomalies). We do so to reflect the data accessible to investors at the time of publication. While data that starts at the first month of the original publication sample was available to investors, data used in the OSAP dataset today prior to the start of the original publication sample may not have been available to investors that easily at the time of publication. Nevertheless, Subsection 3.2 shows that our results are similar if we always use the first month of anomaly strategy return available in the OSAP dataset.

anomaly has a statistically significant historical factor regression alpha (and allow the same anomaly to be “published” multiple times). The results from that analysis are similar to the ones we present for our baseline analysis that uses actual publication dates.

The OSAP dataset includes 212 anomaly signals labeled as “predictors” in their respective original samples, with 179 of these signals having monthly anomaly strategy returns available.⁸ Our final baseline sample contains the subset of 177 anomalies that also meet our criteria of five years with anomaly strategy return data available both before and after the publication year. The summary statistics for these 177 anomaly strategies are reported in Table 1.⁹ The average publication year is 2004, but the sample includes anomalies with publication year as early as 1973 and as late as 2016. The average first year in the publication sample is 1970 while the average first year of available anomaly strategy return is 1955, highlighting that the OSAP dataset typically contains anomaly strategy returns starting even earlier than the original anomaly publication sample. The average anomaly strategy has returns for 638 months in the sample we use in our main analysis, and even the anomaly with the worst coverage has return data for 240 months in this baseline sample. Most of the return data is from the pre-publication period, with an average of 412 pre-publication months versus 226 post-publication months.

Anomaly portfolios earn positive average returns in all subperiods, but average returns are generally lower after publication, in line with the results in McLean and Pontiff (2016). In contrast, the CAPM market risk premium ($\beta \cdot \mathbb{E}[r_m]$) is slightly negative on average and shows little change from the pre-publication to the post-publication period. Consequently, factor regression alphas strongly decline after publication. Nevertheless, factor regression alphas are still positive on average and often economically and statistically significant even after publication. Our main argument, however, is that these factor regression alphas are not accessible to real time investors, as they would need prior knowledge of the optimal

⁸The drop from 207 to 179 anomaly signals is because our anomaly strategy returns are based on decile portfolios, which are not available for all anomalies. For instance, some anomaly signals are discrete variables taking on less than ten possible values.

⁹Throughout the paper, we “annualize” performance metrics when reporting statistics by multiplying average returns and alphas by twelve.

weights required to combine each anomaly with the market portfolio. Since these weights are unknown, the remainder of this paper also explores real time investor alphas, calculated using weights estimated from the pre-publication period, making them feasible for investors in real time.

Note that Table 1 reports alpha t-statistics based on Newey and West (1987, 1994), which are common in the literature, as well as bootstrap t-statistics, with the bootstrap procedure described in Internet Appendix B. The two procedures yield similar distributions for the t-statistics of post-publication factor regression alphas in Table 1. They also tend to be similar for individual anomaly strategies, as Figure IA.1 in the Internet Appendix demonstrates. So, in the following sections, we use only the bootstrap procedure for statistical inference (since Newey and West (1987, 1994) cannot be applied to obtain standard errors for our real time investor alphas).

3 The Alphas of Equity Anomaly Strategies: Results

This section presents the results from our estimation of alphas. Subsection 3.1 focuses on the main results. Subsection 3.2 presents a robustness analysis with respect to our core empirical decisions. Subsection 3.3 considers beta-neutral strategies. The key result of this section is that while factor regression alphas are, on average, large and positive over the post-publication period, real time investor alphas are, on average, relatively close to zero. Moreover, this result is driven by a decline in the (ex-post) optimal anomaly weight from the pre-publication period to the post-publication period.

3.1 Main Results

Table 2 provides our main results. The first three rows show the distribution of pre-publication factor regression alphas (α), with their respective t-statistics and optimal weights (i.e., maximum Sharpe ratio weights). The most interesting observation from these rows is that the median optimal weight underlying the pre-publication α is 0.63 so that an in-

vestor would typically need to allocate more than 50% of his portfolio to a given anomaly to achieve its pre-publication α . The next three rows show the distribution of α estimates post-publication, again with their respective t-statistics and optimal weights. In stark contrast to the pre-publication period, the median optimal weight is 0.36 so that there is a large decline in how much investors should allocate to a given anomaly in comparison to the pre-publication period. While not reported in the table, we also find that the rank correlation (known as Spearman correlation) between the pre-publication optimal weights and post-publication optimal weights is only 0.39.¹⁰ So, an anomaly receiving a relatively high optimal weight in the pre-publication period is not a strong indicator that it should receive a relatively high weight over the post publication period.¹¹

Our core point is that, at the time of publication (or at any given date), investors do not have a reliable way to know the optimal weight to hold on a given anomaly strategy going forward. As a consequence, the real time investor alphas ($\alpha_{\mathcal{R}}$), which apply the pre-publication optimal weights to generate each strategy over the post-publication period, are relatively close to zero on average (see the next two rows of Table 2). $\alpha_{\mathcal{R}}$ estimates can be large for individual anomalies, but they are of relatively similar magnitude on both sides of the $\alpha_{\mathcal{R}}$ distribution. For instance, the 10% and 90% quantiles of the $\alpha_{\mathcal{R}}$ distribution are -8.8% and 12.8%, respectively. Moreover, for most anomalies, these $\alpha_{\mathcal{R}}$ estimates are statistically insignificant.¹²

Figures 3(a) and 3(b) plot the density functions for the post-publication α and $\alpha_{\mathcal{R}}$ es-

¹⁰The usual linear correlation (i.e., the Pearson correlation) is even lower at -0.04. However, this low correlation is due to extreme weights (typically in the post-publication weights). If we bound weights to be between 0 and 1, then the linear correlation becomes 0.41, which is similar to the rank correlation.

¹¹One may conjecture that an investor applying shrinkage to the pre-publication optimal weight would reach a weight much closer to the ex-post optimal weight post-publication. Internet Appendix C shows that this is not the case using empirical Bayesian and machine learning regularization methods. As that section explains, shrinkage is only helpful when jointly investing in multiple anomalies. Importantly, Section 4 also shows that mean-variance investors who apply shrinkage when jointly investing in multiple anomalies are not aggressive enough to eliminate the CAPM alphas of anomalies in equilibrium.

¹²We find that some anomalies have economically and statistically significant $\alpha_{\mathcal{R}}$ estimates. However, investors need to identify these anomalies ex-ante to harvest their alphas. Exploring whether investors can do that in real time is an interesting avenue for future research.

timates, as well as their respective t-statistics. While only a small fraction of the anomaly strategies have a negative α (12.4%), almost half of the anomaly strategies have a negative $\alpha_{\mathcal{R}}$ (46.9%). Similarly, while 48.0% of the α estimates have a t-stat above 1.64, only 10.7% of the $\alpha_{\mathcal{R}}$ estimates do. In terms of average alphas across all anomalies, their t-stats are 9.43 for α and 0.38 for $\alpha_{\mathcal{R}}$ (as shown in Figure 1). So, the average α is statistically significant while the average $\alpha_{\mathcal{R}}$ is not.

Figures 3(c) and 3(d) (as well as the last two rows of Table 2) provide information on the distribution of $\alpha - \alpha_{\mathcal{R}}$ estimates and their respective t-statistics. By construction, all $\alpha - \alpha_{\mathcal{R}}$ estimates are positive so that the risk-adjusted return investors face in reality is worse than the α estimates suggest for all anomaly strategies (as highlighted in the diagram of Figure 2 and related commentary in Section 1.3). Given the level of uncertainty in alphas, only 20.3% of the $\alpha - \alpha_{\mathcal{R}}$ estimates have a t-stat above 1.64. So, at a 5% significance level, we can only reject the null of a zero $\alpha - \alpha_{\mathcal{R}}$ in favor of a positive $\alpha - \alpha_{\mathcal{R}}$ for 20.3% of the anomaly strategies we study. However, the p-value for the average $\alpha - \alpha_{\mathcal{R}}$ is 6.09%, so we can reject the hypothesis that the average $\alpha - \alpha_{\mathcal{R}}$ is zero at a 10% level and are close to reject the same hypothesis at a 5% level.¹³

The results outlined above summarize our key findings, which imply that a real time mean-variance investor who holds the market portfolio would not benefit from adding an anomaly to his portfolio. The next subsection examines the robustness of these results across various empirical choices. We focus on $\alpha_{\mathcal{R}}$ rather than $\alpha - \alpha_{\mathcal{R}}$, as $\alpha_{\mathcal{R}}$ values are more relevant from an economic standpoint. Investors do not choose between the ex-post optimal strategy and the ex-ante optimal strategy (with tradeoff captured by $\alpha - \alpha_{\mathcal{R}}$). Instead, they face the decision to invest solely in the market or in the ex-ante optimal strategy that combines the anomaly with the market (with tradeoff captured by $\alpha_{\mathcal{R}}$). Therefore, from an asset pricing perspective, the distribution of $\alpha_{\mathcal{R}}$ is more relevant than that of $\alpha - \alpha_{\mathcal{R}}$.

¹³Note that, by construction, $\alpha - \alpha_{\mathcal{R}} \geq 0$ for each anomaly strategy, and thus the alternative hypothesis in the statistical test is “average $\alpha - \alpha_{\mathcal{R}} > 0$ ” and not “average $\alpha - \alpha_{\mathcal{R}} \neq 0$ ”.

3.2 Robustness Analysis

We consider sixteen alternative specifications relative to our baseline empirical procedure. The barplot in Figure 4 reports the resulting average alphas as well as the respective interquartile ranges of the distributions of alphas across anomalies, with the lower and upper bounds indicating the 25th and 75th percentiles. We sort specifications based on the magnitude of their average pre-publication factor regression alpha and color each bar based on whether the respective average alpha is statistically significant at the 5% level (blue) or not (orange). As is clear from the figure, our key finding (summarized in Figure 1) is robust to empirical decisions. Specifically, the average post-publication factor regression alpha (α) is economically and statistically significant in all specifications, albeit much lower than its respective average pre-publication α . In contrast, the average post-publication real time investor alpha ($\alpha_{\mathcal{R}}$) is relatively close to zero and statistically insignificant in all specifications. Moreover, the interquartile range of alphas only contains positive alpha values for almost all specifications when looking at α , but always contains both positive and negative alphas when looking at $\alpha_{\mathcal{R}}$. Below, we detail the specifications considered in Figure 4.

The first group of specifications modifies how we select anomalies for the analysis. The objective is to deal with the concern that we may be relying on relatively weak anomalies either because anomalies are weak to start with or because the sample available for those anomalies is short. The specification “Clear Predictors” keeps only the 138 anomalies classified as clear predictors in the OSAP dataset (i.e., anomalies which are clearly treated as predictors in their original publications). The specification “Alpha t-stat ≥ 3 ” keeps only the 120 anomalies with an α t-stat above 3.0 in the pre-publication period (based on Newey and West (1987, 1994)). Specification “10 Years Pre and Post Publication” keeps only the 162 anomalies that have at least 10 years in both the pre-publication and the post-publication periods. And specification “Returns Available since 1963” keeps only the 115 anomalies that have anomaly strategy returns available in the OSAP dataset over the entire period from July 1963 to December 2023 (and uses returns starting in July 1963 for these 115 anomalies even in cases in which the original publication sample starts after July 1963).

The second group of specifications modifies how we form the estimation period (used to estimate the weight implicit in $\alpha_{\mathcal{R}}$) and the evaluation period (used to calculate each $\alpha_{\mathcal{R}}$ estimate). Specification “Pre-Publication Period ≥ 1992 ” starts the pre-publication period in January 1992 to avoid estimating pre-publication weights with data that is too stale relative to the post-publication period (in this case, we keep only the 156 anomalies published at or after 1996). Specification “All Returns Available in OSAP” starts the pre-publication period in the first month of anomaly strategy return available in the OSAP dataset, which is typically earlier than the first month in the original publication sample used in our baseline analysis. Specification “ T_0 from Last Year of Publication Sample” defines the estimation and evaluation periods based on the last year of the publication sample as opposed to the publication year. Specification “Account for Alpha Shock in Transition” ends the pre-publication period one year before the end of the publication sample and starts the post-publication period three years after the publication date (all months in between are dropped from the analysis). This analysis accounts for the evidence in Pénasse (2022) that realized returns are abnormally high in the period around the end of the publication sample as well as right after the publication date (this happens because true alpha declines, inducing an unexpected negative shock in discount rates). And specification “Anomalies Even Before Publication” ignores actual publication dates. Specifically, for each anomaly and each month T_0 (from five years after the anomaly’s first month in OSAP to five years before the anomaly’s last month in OSAP), we estimate the anomaly alpha over $t = 1, \dots, T_0$ and treat cases with a t-statistic above 2.0 as new anomalies “published” at T_0 , with their “post-publication” period covering $t = T_0 + 1, \dots, T$.¹⁴ As such, the same anomaly can “be published” at multiple months in this analysis and these months can precede the true publication date of the anomaly. The $\alpha_{\mathcal{R}}$ estimates in this specification are subject to publication bias since the “post-publication period” here can cover the original publication sample (which would lead to anomaly returns that

¹⁴In this “Anomalies Even Before Publication” specification, $t = 1$ reflects the first month of anomaly strategy return available in the OSAP dataset (not the first month in the original publication sample). The reason is that this specification implicitly assumes investors evaluate anomaly signals even before their publication date. So, the original publication sample is irrelevant in the context of this specification.

are artificially inflated by the selection effect implicit in the publication process). However, this analysis is useful in illustrating that our core point that real time investor alphas are substantially lower than factor regression alphas is not restricted to the true post-publication period of anomaly strategies, which we use in our baseline results to avoid the publication bias issue described above.

The third group of specifications modifies how we form the long-short portfolio returns used to define anomaly strategies. The “Equal-Weighted Deciles” and “Value-Weighted-Weighted Deciles” specifications use equal-weighted and value-weighted deciles, respectively, instead of relying on the same weights used in the original publication (the value-weighted deciles are available in the OSAP dataset whereas we construct our own equal-weighted deciles based on OSAP signals because OSAP does not provide these portfolios). The “Original Anomaly Portfolios” specification constructs each long-short portfolio using the exact same procedure as in the original paper instead of using decile portfolios (these anomaly strategy returns are also available directly in the OSAP dataset). And specifications “Dynamic Portfolio Weights, Expanding” and “Dynamic Portfolio Weights, Rolling” follow our baseline weight estimation procedure at $t = T_0$, but update the weights each month from $t = T_0 + 1, \dots, T - 1$ to construct $r_{p,t+1}$ in the subsequent month. We consider both an expanding window as well as a 5-year rolling window for the weight estimation.

The final group of specifications allows investors to incorporate anomaly performance decay in their μ estimates. In particular, they replace $\bar{r}_a$ (the second element of μ_{OS} in Equation 10) with $\bar{r}_a^{decay}$. In specification “Account for Mu Decay”, we use $\bar{r}_a^{decay} = \theta \cdot \bar{r}_a$, where $0 \leq \theta \leq 1$ so that the θ parameter reflects investors’ estimate for the decay in average returns after publication (the baseline analysis effectively imposes $\theta = 1$). The value used for θ varies by anomaly based on the information investors have on all other anomalies that have been published as of the time of the given anomaly publication. Specifically, we estimate θ for a given anomaly (let’s call it anomaly a) as follows. We start by identifying all anomalies that have been published (with at least five years of returns post-publication) by the end of the pre-publication period of anomaly a . Then, for each of these anomalies, we

calculate the ratio between their respective average returns over their own post-publication and pre-publication periods (all within the pre-publication period of anomaly a). Finally, we estimate the θ for anomaly a using the average of these ratios across all of these other anomalies while bounding it to be between zero and one. In specification “Account for Alpha Decay”, we use $\bar{r}_a^{decay} = \theta \cdot \alpha_a + \beta_a \cdot \bar{r}_m$, where again $0 \leq \theta \leq 1$ so that the θ parameter reflects investors’ estimate for the decay in α after publication (with α_a and β_a estimated over the pre-publication period for anomaly a). To obtain θ , we again start by identifying all anomalies that have been published by the end of the pre-publication period of anomaly a (with at least five years of returns post-publication). Then, for each of these anomalies, we calculate $\theta = \alpha_{Post}/\alpha_{Pre}$, where α_{Post} and α_{Pre} are estimated, respectively, over the post-publication and pre-publication periods for the given anomaly (again, within the pre-publication period of anomaly a). Finally, as before, we estimate the θ for anomaly a using the average of these ratios across all of these other anomalies while bounding it to be between zero and one.

In addition to the specifications discussed above, Internet Appendix C further explores applying shrinkage when estimating w_{OS} . It shows that with the “right level of shrinkage” the distribution of post-publication $\alpha_{\mathcal{R}}$ estimates resembles that of α estimates. However, when the level of shrinkage is chosen using past data (i.e., through empirical Bayesian or machine learning regularization methods), the distribution of $\alpha_{\mathcal{R}}$ estimates is similar to the one we report in the main text. Since a real time investor would not know the “right level of shrinkage” ex ante, this result indicates that shrinkage does not solve the issue that anomalies provide low real time investor alphas on average.

3.3 Beta-Neutral Portfolios

The previous two subsections obtain real time investor alphas from portfolios that attempt to maximize Sharpe ratio in real time. This is the relevant economic objective for real time mean-variance investors, who are the relevant investors for the purpose of understanding why CAPM alphas survive in equilibrium.

However, an alternative way to invest in anomalies is to consider a beta-neutral portfolio

given by $r_{a,t} - \beta \cdot r_{m,t}$. This portfolio is economically relevant because its expected return exactly matches the anomaly’s factor regression alpha if the portfolio is constructed using the in-sample beta. This result also holds approximately even if the beta-neutral portfolio is constructed out-of-sample. Specifically, Figure 5(a) shows that post-publication factor regression alphas are very similar to the post-publication average returns of beta-neutral portfolios constructed in real time using only pre-publication betas. In addition, Figure 5(b) shows that this result continues to hold under a simple $\beta = 0$ approximation, in which case the beta-neutral portfolio is represented by the anomaly return itself. So, investors can approximately “earn the factor regression alpha” as the average return of a beta-neutral portfolio constructed in real time.

This fact does not change our message that deviating from the market portfolio to invest in an anomaly is not attractive to real time mean-variance investors. The reason is that a mean-variance investor would not want to hold a beta-neutral portfolio since such a portfolio does not attempt to maximize Sharpe ratio (which is equivalent to maximizing investor alpha). In fact, Figure 5(c) shows that the post-publication investor alphas of beta-neutral portfolios constructed in real time tend to be even lower than the real time investor alphas we study in Subsection 3.1 (because of the lower post-publication Sharpe ratios of beta-neutral portfolios). Figure 5(d) shows a similar result under the $\beta = 0$ approximation, in which case the beta-neutral portfolio is the anomaly itself. So, investing directly in an anomaly (without a market position) or in its real time beta-neutral portfolio is unattractive to mean-variance investors despite the fact that such a portfolio delivers an average return close to the anomaly’s factor regression alpha.

To understand this empirical result, note that the Sharpe ratio of a beta neutral-portfolio is the same as the Sharpe ratio of $r_{p,t} = (1 - w_\beta) \cdot r_{m,t} + w_\beta \cdot r_{a,t}$, where $w_\beta = 1/(1 - \beta)$. Since the median pre-publication β is -0.098, the typical w_β is around 0.91, whereas the median optimal weight post-publication is 0.36 (see Table 2). So, beta-neutral strategies allocate substantially more to anomalies than what is optimal from a Sharpe ratio perspective.

4 The Need for Shrinkage and the Survival of the Anomaly Zoo

The previous section shows that allocation uncertainty implies that a real time mean-variance investor who currently holds the market portfolio does not have an incentive to deviate by adding an anomaly to his portfolio, which suggests this investor would not drive CAPM alphas to zero. However, previous papers show that jointly investing in multiple anomalies in real time can lead to Sharpe ratios above the market Sharpe ratio (e.g., Kozak, Nagel, and Santosh (2020) and Bryzgalova, Pelger, and Zhu (2024)). Moreover, recent work find similar results after considering tradings costs (Jensen et al. (2025)), publication considerations (Lassance and Martín-Utrera (2025)), and situations where the number of assets is larger than the number of time periods (Hellum et al. (2026)). Importantly, all these papers apply some form of shrinkage (explicitly or implicitly). In fact, many papers highlight the difficulties of applying mean-variance optimization without some form of shrinkage (e.g., Jorion (1986), Kan and Zhou (2007), Garlappi, Uppal, and Wang (2007), DeMiguel, Garlappi, and Uppal (2009), Kan, Wang, and Zhou (2022), and Kan and Wang (2024)). So, simultaneously investing in multiple anomalies in real time can lead to positive real time investor alphas, but it requires using shrinkage during the allocation process.

This section shows that investors who take this shrinkage approach are not aggressive enough to eliminate CAPM alphas in equilibrium. So, allocation uncertainty (which leads to the need for shrinkage) implies the anomaly zoo survives in equilibrium. Subsection 4.1 provides the model. Subsection 4.2 estimates the model to understand much of the anomaly alphas it explains. Subsection 4.3 argues that the allocation uncertainty mechanism we emphasize can also be incorporated into models that include other complementary channels and demonstrates this explicitly in a model with noise traders (in the spirit of Long et al. (1990)). Internet Appendix D provides all model derivations.

4.1 A Model for the Survival of the Anomaly Zoo

We consider a simple model that deviates from the CAPM only because investors face some uncertainty when forming expected returns. In the main text, investors take returns as given, but Internet Appendix E generalizes the model to an endowment economy so that investors take cash flows as given, with returns being fully endogenous.

There is a representative investor who can allocate capital to the risk-free asset with known gross return R_f (and zero net supply), to the market portfolio with excess return $r_{m,t} = R_{m,t} - R_f$ (and unit net supply), and to a set of anomaly portfolios with long-short return vector r_t (and zero net supply). The investor has mean-variance preferences with risk aversion $\gamma > 0$, and thus forms his wealth portfolio (with gross return $R_{w,t}$) by solving

$$[\omega_{m,t}, \omega_t] = \operatorname{argmax} \mathbb{E}_t[R_{w,t+1}] - 0.5 \cdot \gamma \cdot \mathbb{V}ar_t[R_{w,t+1}] \quad (11)$$

$$s.t. \quad R_{w,t+1} = R_f + \omega_{m,t} \cdot r_{m,t+1} + \omega_t' r_{t+1}$$

The first order conditions of this problem yield the mean-variance portfolio allocation solution:

$$\omega_{m,t} = (1/\gamma) \cdot \mathbb{V}ar_t[r_{m,t+1}]^{-1} \left(\mathbb{E}_t[r_{m,t+1}] - \gamma \cdot \mathbb{C}ov_t[\omega_t' r_{t+1}, r_{m,t+1}] \right) \quad (12)$$

$$\omega_t = (1/\gamma) \cdot \mathbb{V}ar_t[r_{t+1}]^{-1} \left(\mathbb{E}_t[r_{t+1}] - \gamma \cdot \mathbb{C}ov_t[r_{t+1}, \omega_{m,t} \cdot r_{m,t+1}] \right). \quad (13)$$

Each return is normally distributed with conditional mean $\mu_{n,t}$, where $n = m, a$ with a representing an anomaly index. The investor observes the market variance ($\mathbb{V}ar_t[r_{m,t+1}] = \sigma_{m,t}^2$) and the covariance matrix of anomaly returns ($\mathbb{V}ar_t[r_{t+1}] = \Sigma_t$), which we assume to be diagonal without loss of generality: $\Sigma_t[a, a] = \sigma_{a,t}^2$.¹⁵ However, in line with the arguments in Merton (1980), the investor faces uncertainty about $\mu_{n,t}$. Specifically, using all publicly available information (including current prices), the investor can only obtain the $\mu_{n,t}$ maximum likelihood estimate (MLE), which has sampling distribution $\hat{\mu}_{n,t} \sim \mathcal{N}(\mu_{n,t}, \hat{v} \cdot \sigma_{n,t}^2)$.

¹⁵If the original N anomaly long-short returns are not orthogonal, we can always construct their N principal components (PCs), which are orthogonal by construction. These PCs span the same space as the original anomalies and remain at zero net supply. So, assuming a diagonal Σ_t is without loss of generality.

To form $\mathbb{E}_t[r_{n,t+1}] = \mathbb{E}_t[\mu_{n,t}]$, the investor combines the $\hat{\mu}_{n,t}$ distribution with the prior belief that $\mu_{n,t} \sim \mathcal{N}(-\lambda \cdot \sigma_{n,t}^2, \nu \cdot \sigma_{n,t}^4)$, with $\lambda \geq 0$. This prior matches the one in Bryzgalova, Pelger, and Zhu (2024) (and in Kozak, Nagel, and Santosh (2020) if $\lambda = 0$), assigning a higher probability of high Sharpe ratio to principal component factors with larger eigenvalues. As emphasized by Bryzgalova, Pelger, and Zhu (2024), this prior leads to a robust mean-variance representation that is isomorphic to a regularization method for the recovery of the maximum Sharpe ratio portfolio.

The resulting expected returns are given by the posterior expectation

$$\mathbb{E}_t[r_{n,t+1}] = \mathbb{E}_t[\mu_{n,t}] = \frac{\sigma_{n,t}^2}{\sigma_{n,t}^2 + \psi} \cdot (\hat{\mu}_{n,t} - \lambda \cdot \psi), \quad (14)$$

and thus the investor's optimal allocation in Equations 12 and 13 can be written as

$$\omega_{m,t} = \frac{1}{\gamma \cdot \sigma_{m,t}^2} \cdot \left(\frac{\sigma_{m,t}^2}{\sigma_{m,t}^2 + \psi} \cdot (\hat{\mu}_{m,t} - \lambda \cdot \psi) - \gamma \cdot \text{Cov}_t[\omega'_t r_{t+1}, r_{m,t+1}] \right) \quad (15)$$

$$\omega_{a,t} = \frac{1}{\gamma \cdot \sigma_{a,t}^2} \cdot \left(\frac{\sigma_{a,t}^2}{\sigma_{a,t}^2 + \psi} \cdot (\hat{\mu}_{a,t} - \lambda \cdot \psi) - \gamma \cdot \text{Cov}_t[r_{a,t+1}, \omega_{m,t} \cdot r_{m,t+1}] \right), \quad (16)$$

where $\psi = \hat{\nu}/\nu > 0$.

Importantly, the optimal allocation in Equations 15 and 16 implies investors are not as aggressive as standard mean-variance investors since

$$\partial_{\hat{\mu}_{n,t}} \omega_{n,t} = \frac{1}{\gamma \cdot (\sigma_{n,t}^2 + \psi)} < \frac{1}{\gamma \cdot \sigma_{n,t}^2}. \quad (17)$$

Consequently, CAPM alphas survive in equilibrium. To see that, impose market clearing ($\omega_{m,t} = 1$ and $\omega_{a,t} = 0$) and solve for the equilibrium MLE of expected returns to get

$$\hat{\mu}_{m,t} = \gamma \cdot (\sigma_{m,t}^2 + \psi) + \lambda \cdot \psi \quad (18)$$

$$\hat{\mu}_{a,t} = (1 + \psi/\sigma_{a,t}^2) \cdot \gamma \cdot \sigma_{m,t}^2 \cdot \beta_{a,t} + \lambda \cdot \psi \quad (19)$$

so that the MLE for the CAPM alpha is given by

$$\begin{aligned} \hat{\alpha}_{a,t} &= \hat{\mu}_{a,t} - \beta_{a,t} \cdot \hat{\mu}_{m,t} \\ &= \psi \cdot (\lambda + \gamma \cdot (\sigma_{m,t}^2/\sigma_{a,t}^2) \cdot \beta_{a,t} - (\gamma + \lambda) \cdot \beta_{a,t}) \end{aligned} \quad (20)$$

where $\beta_{a,t} = \text{Cov}_t[r_{a,t+1}, r_{m,t+1}] / \sigma_{m,t}^2$.¹⁶

So, the CAPM alphas differ from zero. Yet, the investor does not eliminate these alphas in equilibrium. The reason is that the investor does not perceive non-zero alphas since $\mathbb{E}_t[r_{a,t+1}] = \beta_{a,t} \cdot \mathbb{E}_t[r_{m,t+1}]$ holds in equilibrium. Only the econometrician measures non-zero alphas. Importantly, this difference between the econometrician and the investor arises entirely from the μ uncertainty the investor faces since we have $\hat{\alpha}_{a,t} = 0$ if $\psi = \hat{\nu}/\nu = 0$, in which case the investor treats the MLE $\hat{\mu}_{n,t}$ as if it had no sampling uncertainty and the model reduces to the CAPM. So, endogenous shrinkage is responsible for the non-zero alphas in this model.

However, the factor regression alphas we study in prior sections are based on realized alphas over a period (the post-publication period in our empirical analysis) as opposed to $\hat{\alpha}_t$, which depends on all information available to investors at time t . To obtain an expression for realized alphas in the model, we add the auxiliary assumption that the return distribution remains fixed after time t . In this case, the realized alpha of anomaly a from $t+1$ to $t+H$ is

$$\hat{\alpha}_a = \hat{\mu}_a - \hat{\beta}_a \cdot \hat{\mu}_m \quad (21)$$

where $\hat{\mu}_n = (1/H) \cdot \sum_{h=1}^H r_{n,t+h}$ and $\hat{\beta}_a = (\sum_{h=1}^H (r_{a,t+h} - \hat{\mu}_a) \cdot (r_{m,t+h} - \hat{\mu}_m)) / (\sum_{h=1}^H (r_{m,t+h} - \hat{\mu}_m)^2)$.

Moreover, the investor can forecast these realized alphas (through $\hat{\alpha}_t$) and still does not eliminate them in equilibrium. Specifically, we have

$$\hat{\alpha}_a = \hat{\alpha}_{a,t} + \varepsilon_a \quad (22)$$

where $\mathbb{E}_t[\varepsilon_a] = 0$ so that $\hat{\alpha}_{a,t}$ is an unbiased estimate for $\hat{\alpha}_a$.

Thus, alphas persist even in an economy in which the CAPM assumption that all in-

¹⁶If our model was based on investors who have FIRE, the analogue of Equation 18 would provide an expression for μ_t instead of its maximum likelihood estimate, $\hat{\mu}_t$. It is easy to see how demand affects μ_t in equilibrium since high (low) demand induces high (low) prices, which leads to low (high) μ_t . A similar logic applies to $\hat{\mu}_t$ since it reflects the maximum likelihood estimate based on all publicly available information, including prices. So, high (low) demand induces high (low) prices, which leads the maximum likelihood estimator to indicate a low expected return going forward (i.e., a low $\hat{\mu}_t$). Internet Appendix E provides an endowment economy version of the model that makes this endogenous price mechanism clear.

vestors have mean-variance preferences and homogeneous beliefs holds. The reason is that investors endogenously apply shrinkage to their portfolio weights to deal with the uncertainty present in the estimation of anomaly portfolio weights. As a consequence, their demand is not aggressive enough to eliminate the alphas of anomaly strategies in equilibrium.

Importantly, we simplify the model in the main text by assuming mean-variance preferences, specifying the data generating process in terms of returns, and treating the MLE $\hat{\mu}_t$ as given. However, Internet Appendix E provides an analogous model in which the representative investor has power utility, operates in an endowment economy, and derives $\hat{\mu}_t$ from current prices and a noisy signal about the endowment growth. The implications from this endowment economy version of the model are very similar to those presented in this section.

4.2 How Much of the CAPM Alphas of Anomalies Can the Model Explain?

Combining Equations 22 and 20, we have

$$\hat{\alpha}_a = A + B \cdot \beta_a + C \cdot \beta_a \cdot \frac{\sigma_m^2}{\sigma_a^2} + \varepsilon_a \quad (23)$$

where $A = \psi \cdot \lambda \geq 0$, $C = \psi \cdot \gamma \geq 0$, and $B = -(A + C) \leq 0$ since $\gamma > 0$, $\psi \geq 0$, and $\lambda \geq 0$.

Equation 23 shows that the equilibrium CAPM alphas are not zero despite the model featuring a representative mean-variance investor. Intuitively, even though the market portfolio is mean-variance efficient from the perspective of the Bayesian investor, it is not mean-variance efficient from the perspective of the econometrician. So, the market portfolio does not perfectly price assets under the econometrician's measure, which induces a failure of the CAPM. The only exception is if $\psi = 0$ (i.e., the investor perceives no uncertainty in $\hat{\mu}$).

The failure of the CAPM can be summarized by two effects. First, CAPM alphas are positive on average so long as $A > 0$ (which requires $\psi > 0$ and $\lambda > 0$). This effect arises from the fact that the investor has the pessimistic prior that $\mathbb{E}_t^{prior}[\mu_{n,t}] = -\lambda \cdot \sigma_{n,t}^2 \leq 0$, and thus views anomalies as bad strategies on average. So, his overall demand for anomalies is low, which induces anomalies to have high expected returns in equilibrium (on average). Second, holding σ_m^2/σ_a^2 fixed, alphas decrease with beta if $\sigma_m^2/\sigma_a^2 < 1 + \lambda/\gamma$ (which tends to be the case

empirically). This effect is a joint consequence of $B < 0$ and $C > 0$. Intuitively, the investor perceives no alpha under his posterior expected returns, which are compressed relative to the econometrician’s expected returns. However, the CAPM alpha is measured using the econometrician’s expected returns, so we need to “unshrink” expected returns to measure the CAPM alpha. The necessary adjustment is larger for the market if $\sigma_m^2/\sigma_a^2 < 1 + \lambda/\gamma$, which induces lower alphas for higher beta anomalies. This mechanism provides an explanation for the low-beta anomaly based on uncertainty-induced endogenous shrinkage, rather than leverage constraints as in Frazzini and Pedersen (2014).

Table 3 provides results from estimating Equation 23. The first two rows use post-publication factor regression alphas for the left hand side of the equation and pre-publication betas and variances for the right hand side. The first four columns consider a reduced form implementation that estimates A , B , and C by Ordinary Least Squares (OLS), imposing no structural restriction. Interestingly, the reduced-form parameters are of the correct sign and statistically significant. Moreover, the $R^2 = 17.1\%$ suggests that the allocation uncertainty mechanism can explain a non-trivial fraction of the variation in anomaly alphas. The last four columns impose the structural restrictions by directly estimating the $\gamma > 0$, $\psi \geq 0$, and $\lambda \geq 0$ parameters by Non-Linear Least Squares (NLS) and recovering the reduced form parameters based on $A = \psi \cdot \lambda$, $C = \psi \cdot \gamma$, and $B = -(A + C)$. Remarkably, the implied reduced form parameter estimates are similar to the ones obtained in the reduced form implementation. Moreover, the $R^2 = 16.0\%$ value is only slightly lower than in the specification that imposes no structural restriction.

The next two rows of Table 3 repeat the analysis for the 115 anomalies with return data available throughout the common July 1963 to December 2023 sample. In this specification, the alphas, betas, and variances are all estimated over the same period, rather than using pre-publication betas and variances to explain post-publication alphas. The results remain similar. The reduced form parameters still have the signs predicted by the model and are statistically significant (albeit C is only statistically significant at the 10% level in a one-sided test). The R^2 is 15.0%, which is also close to the 17.1% value obtained in the previous

specification. Imposing the structural restrictions again produces similar estimates for the reduced form parameters and an R^2 of 14.5%, only slightly lower than in the reduced form implementation. These results show that our findings do not depend on the use of pre- vs post-publication data.

The final two rows of Table 3 replace the 115 anomaly strategies with their 115 principal component (PC) factors, using the same July 1963 to December 2023 sample.¹⁷ This specification more directly corresponds to the model because PC factors are orthogonal by construction, satisfying the diagonal covariance matrix representation used in Subsection 4.1, while spanning the same return space as the original anomaly strategies. The model performs substantially better in this orthogonal basis. The reduced form parameters have the signs predicted by the model and produce an R^2 of 32.5%. Moreover, imposing the structural restrictions leads to similar parameter estimates, with an R^2 of 31.6%, which is comparable to the R^2 from the reduced form implementation.

Beyond the model fit, we can also inspect the structural parameter estimates to check whether the model requires unreasonable structural parameters to explain anomaly alphas. However, Equation 23 does not separately identify the three structural parameters.¹⁸ So, Figure 6 performs two exercises that fix one structural parameter to obtain estimates for the other two based on the reduced form estimates. For these exercises, we rely on the model estimation in Table 3 that uses 115 PC factors (last two rows) and imposes the model structural restrictions (last four columns). This is the specification that best corresponds to the model in Subsection 4.1. However, Internet Appendix Figure IA.5 further provides the analogous figure based on the model estimation that uses 177 anomalies while incorporating the pre- vs post-publication structure (last four columns of first two rows in Table 3).

Figure 6(a) fixes ψ (in the x-axis) to recover estimates for γ and λ (in the y-axes). From Equation 14 (and the ψ definition), we know that reasonable values for ψ are of the same order of magnitude as market variance. So, we set ψ to values within the interval $[\bar{\psi}/5, \bar{\psi} \cdot 5]$

¹⁷We normalize each PC factor to have a positive CAPM alpha from July 1963 to December 2023.

¹⁸To see that the structural parameters are not separately identifiable from Equation 23, note that multiplying ψ by an arbitrary $\theta > 0$ and dividing λ and γ by the same θ leaves A , B , and C unchanged.

where $\bar{\psi} = 0.0025$, which corresponds to a 5% monthly volatility. When $\psi = \bar{\psi}$, we have the estimates $\gamma = 2.96$ and $\lambda = 1.49$, which are reasonable (note that the λ definition implies it is of the same order of magnitude as γ). As we increase ψ , we get lower γ and λ values. However, we only see unreasonable γ values (below 0.1 or higher than 10) when ψ largely deviates from $\bar{\psi}$.

Figure 6(b) fixes γ (in the x-axis) to recover estimates for ψ and λ (in the y-axes). We set γ to values within the interval $[\bar{\gamma}/5, \bar{\gamma} \cdot 5]$ where $\bar{\gamma} = 1$, which corresponds to a representative investor who seeks to maximize Sharpe ratio. When $\gamma = \bar{\gamma}$, we have the estimates $\psi = 0.0074$ and $\lambda = 0.51$, which are also reasonable.¹⁹ As we increase γ , we get lower ψ values and higher λ values. However, we again only see unreasonable ψ values (that deviate from $\bar{\psi}$ by a factor of 10) when γ largely deviates from $\bar{\gamma}$.

Figures 6(c) and 6(d) show that while the model explains anomaly alphas well, it tends to produce market risk premia that are too high. Specifically, these figures take the structural parameters from Figures 6(a) and 6(b), respectively, and report the implied market risk premia based on Equation 18. While the average market risk premium over our sample is below 10%, the model always imply values above 10% (within the range of reasonable structural parameters we considered). So, the model is better at explaining the relative risk premia of anomalies in comparison to the market rather than the level of risk premia in the market.

Overall, the results indicate that allocation uncertainty can explain a substantial fraction of the cross-sectional variation in CAPM alphas. Across specifications, the reduced form parameters generally have the signs predicted by the model, imposing the structural restrictions has little effect on model fit, and the R^2 ranges from approximately 15% for the original anomaly strategies to more than 30% for the orthogonal PC factors that most closely match the model structure. Moreover, the model can achieve this explanatory power with

¹⁹The λ value is statistically insignificant in this case. This result would favor the Kozak, Nagel, and Santosh (2020) prior, which is the same as the Bryzgalova, Pelger, and Zhu (2024) prior with the restriction $\lambda = 0$. The parameter that governs the level of uncertainty (ψ) is statistically significance. Moreover, even λ is statistically significant in the analyses that fix ψ instead of γ .

reasonable combinations of the structural parameters γ , ψ , and λ . Thus, uncertainty-induced endogenous shrinkage is not only capable of generating non-zero CAPM alphas in equilibrium, but is also quantitatively relevant for understanding the variation in CAPM alphas across anomalies (albeit it tends to produce a counterfactually high market risk premium).

4.3 Uncertainty Mechanism as a Complement to other Mechanisms

To keep the message simple, the model in the previous subsections shuts down all potential channels for why CAPM alphas survive in equilibrium other than allocation uncertainty (through μ_t uncertainty). For instance, the model does not include non-market risk factors, trading costs, or limits to arbitrage. However, our message is not that allocation uncertainty is the only reason for the survival of the anomaly zoo. Rather, our point is that allocation uncertainty is an important channel that can complement other mechanisms previously proposed in the literature.

To demonstrate our point more explicitly, consider adding noise traders to the model, with the representative investor from the previous section reflecting the rational trader in this heterogeneous agent model. Specifically, in the spirit of Long et al. (1990), we assume there is a mass $1 - \theta$ of rational traders and θ of noise traders, who are identical to the rational traders except that they mistakenly think $\hat{\mu}_t + \tilde{\mu}_t$ is the MLE for μ_t so that their demand has an extra noise component given by $\tilde{\mu}_t$. In this case, the equilibrium alpha becomes

$$\hat{\alpha}_{a,t} = \underbrace{\psi \cdot (\lambda + \gamma \cdot (\sigma_{m,t}^2 / \sigma_{a,t}^2) \cdot \beta_{a,t} - (\gamma + \lambda) \cdot \beta_{a,t})}_{\text{Effect of } \mu \text{ Uncertainty}} \underbrace{- \theta \cdot \tilde{\alpha}_{a,t}}_{\text{Effect of Noise Traders}} \quad (24)$$

where $\tilde{\alpha}_{a,t} = \tilde{\mu}_{a,t} - \beta_{a,t} \cdot \tilde{\mu}_{m,t}$.

Hence, even if the rational traders applies no shrinkage (i.e., $\psi = 0$), the noise trader's belief distortion still moves equilibrium expected returns away from the CAPM implication. The reason is that even though the rational trader absorbs the noise trader's demand, he requires compensation for the extra risk that doing so entails. In particular, the rational trader overweights assets with negative $\tilde{\alpha}$ (because the noise trader underweights them), and thus these assets have higher risk premia and CAPM alphas in equilibrium. This is the

mechanism in Long et al. (1990).

The above discussion makes it clear that our allocation uncertainty mechanism is complementary to other mechanisms in the literature. In particular, while allocation uncertainty by itself can generate non-zero CAPM alphas in equilibrium, our Subsection 4.2 shows that it explains between 15% and 30% of the variation in anomaly alphas, not all of it. So, other mechanisms (e.g., the presence of noise traders) are likely also present in the data and are crucial to fully explain the survival of the anomaly zoo.

5 Conclusion

This paper provides a new mechanism for why the anomaly zoo can survive in equilibrium even after anomalies become public information. The mechanism is allocation uncertainty: investors may know that an anomaly has historically generated a positive CAPM alpha, yet remain uncertain about how much capital to allocate to it when combining it with the market portfolio in real time. To quantify the investment opportunities available under this uncertainty, we introduce the real time investor alpha, which measures the Sharpe ratio improvement obtained using portfolio weights that can be estimated when the investment decision is made. Unlike the usual factor regression alpha, which implicitly relies on ex-post optimal weights, the real time investor alpha captures the incentive mean-variance investors actually have to trade an anomaly away. It therefore provides the link between uncertainty about portfolio weights and the persistence of CAPM alphas in equilibrium.

Empirically, we find that anomaly strategies continue to generate positive and statistically significant factor regression alphas after publication, but their average real time investor alpha is relatively close to zero. This result arises because the optimal weights estimated before publication provide limited guidance about the weights that would have been optimal afterward. Moreover, it is robust to a wide range of empirical choices, including alternative anomaly samples, portfolio constructions, estimation periods, adjustments for post-publication decay, and Bayesian and machine-learning methods for estimating weights.

Thus, the persistence of factor regression alphas does not imply that investors could have profitably traded individual anomalies away in real time.

Combining these empirical findings with an equilibrium model yields a mechanism for the survival of the anomaly zoo. The low average real time investor alpha of individual anomalies gives investors little incentive to add them one at a time to a passive position in the market portfolio. Investors may instead combine multiple anomalies using shrinkage (or regularization) to deal with uncertainty, as is common in quantitative investing. However, investors who account for expected return uncertainty through shrinkage respond less aggressively to estimated return differences than investors with Full Information Rational Expectations. Their demand therefore does not eliminate CAPM alphas in equilibrium. In this sense, the same allocation uncertainty that creates the need for shrinkage also prevents investors from trading anomalies aggressively enough to eliminate their factor regression alphas.

The model shows that this channel can generate nonzero CAPM alphas even in an economy with a representative mean-variance investor, homogeneous beliefs, and no behavioral biases, trading costs or other limits to arbitrage. It also explains a meaningful fraction of the cross-sectional variation in anomaly alphas and generates empirical patterns such as positive average alphas and a negative relation between alphas and betas. These results suggest that uncertainty-induced endogenous shrinkage is not only a qualitative explanation for the survival of the anomaly zoo, but also a quantitatively relevant one.

Our findings do not imply that allocation uncertainty is the only reason for the survival of the anomaly zoo. Non-market risks, trading costs, behavioral biases, and other limits to arbitrage may all contribute to the persistence of CAPM alphas. Rather, allocation uncertainty provides a complementary statistical limit to arbitrage: even smart and unconstrained investors cannot trade as if they know the optimal portfolio weights ex ante. More broadly, the results suggest that understanding whether an alpha should disappear in equilibrium requires considering not only whether investors know that the alpha exists, but also whether they can determine how aggressively to trade it in real time.

References

- Baba Yara, F., B. H. Boyer, and C. Davis (2024). “Messy Asset Pricing: Can AI Models Lead to a Consensus?” Working Paper.
- Bryzgalova, S., M. Pelger, and J. Zhu (2024). “Forest Through the Trees: Building Cross-Sections of Stock Returns”. In: *Journal of Finance* Forthcoming.
- Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997). *The Econometrics of Financial Markets*. Princeton University Press.
- Campbell, J. Y. and S. B. Thompson (2008). “Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?” In: *Review of Financial Studies* 21.4, pp. 1509–1531.
- Cederburg, S., M. S. O’Doherty, F. Wang, and X. S. Yan (2020). “On the Performance of Volatility-Managed Portfolios”. In: *Journal of Financial Economics* 138.1, pp. 95–117.
- Chabi-Yo, F., A. S. Gonçalves, and J. Loudis (2025). “An Intertemporal Risk Factor Model”. In: *Management Science*. Forthcoming.
- Chen, A. Y. and T. Zimmermann (2023). “Publication Bias and the Cross-Section of Stock Returns”. In: *Oxford Research Encyclopedia of Economics and Finance*. Forthcoming.
- Chen, A. Y., A. Lopez-Lira, and T. Zimmermann (2023). “Does peer-reviewed theory help predict the cross-section of stock returns?” Working Paper.
- Chen, A. Y. and M. Velikov (2023). “Zeroing in on the Expected Returns of Anomalies”. In: *Journal of Financial and Quantitative Analysis* 58.3, pp. 968–1004.
- Chen, A. Y. and T. Zimmermann (2022). “Open Source Cross-Sectional Asset Pricing”. In: *Critical Finance Review* 27.2, pp. 207–264.
- Da, R., S. Nagel, and D. Xiu (2024). “The Statistical Limit of Arbitrage”. Working Paper.
- Davis, C. (2025). “The Elasticity of Quantitative Investment”. In: *Review of Financial Studies* 38.10, pp. 2845–2886.

- DeMiguel, V., L. Garlappi, and R. Uppal (2009). “Optimal Versus Naive Diversification: How Inefficient is the $1/N$ Portfolio Strategy?” In: *Review of Financial Studies* 22.5, pp. 1915–1953.
- Fama, E. F. and K. R. French (1996). “Multifactor Explanations of Asset Pricing Anomalies”. In: *Journal of Finance* 55.1, pp. 55–84.
- Frazzini, A. and L. H. Pedersen (2014). “Betting Against Beta”. In: *Journal of Financial Economics* 111.1, pp. 1–25.
- Garlappi, L., R. Uppal, and T. Wang (2007). “Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach”. In: *Review of Financial Studies* 22.5, pp. 1915–1953.
- Gibbons, M. R., S. A. Ross, and J. Shanken (1989). “A Test of the Efficiency of a Given Portfolio”. In: *Econometrica* 57.5, pp. 1121–1152.
- Goyal, A. and I. Welch (2008). “A Comprehensive Look at The Empirical Performance of Equity Premium Prediction”. In: *Review of Financial Studies* 21.4, pp. 1455–1508.
- Goyal, A., I. Welch, and A. Zafirov (2024). “A Comprehensive 2022 Look at the Empirical Performance of Equity Premium Prediction”. In: *Review of Financial Studies* Forthcoming.
- Gromb, D. and D. Vayanos (2010). “Limits of Arbitrage”. In: *Annual Review of Financial Economics* 2, pp. 251–275.
- Guo, H. and J. A. Wachter (2025). “Forecast-agnostic portfolios”. Working Paper.
- Harvey, C. R., Y. Liu, and H. Zhu (2015). “... and the Cross-Section of Expected Returns”. In: *Review of Financial Studies* 29.1, pp. 5–68.
- Hellum, O., T. I. Jensen, B. Kelly, and S. Malamud (2026). “Complex Modern Portfolio Theory”. Working Paper.
- Jensen, T. I., B. T. Kelly, S. Malamud, and L. H. Pedersen (2025). “Machine Learning and the Implementable Efficient Frontier”. In: *Review of Financial Studies* Forthcoming.
- Jensen, T. I., B. T. Kelly, and L. H. Pedersen (2023). “Is There a Replication Crisis in Finance?” In: *Journal of Finance* 78.5. Forthcoming.

- Jorion, P. (1986). “Bayes-Stein Estimation for Portfolio Analysis”. In: *Journal of Financial and Quantitative Analysis* 21.3, pp. 279–292.
- Kan, R. and N. Lassance (2025). “Optimal Portfolio Choice with Fat Tails and Parameter Uncertainty”. In: *Journal of Financial and Quantitative Analysis* Forthcoming.
- Kan, R. and X. Wang (2024). “Optimal Portfolio Choice with Unknown Benchmark Efficiency”. In: *Management Science* 70.9, pp. 6117–6138.
- Kan, R., X. Wang, and X. Zheng (2024). “In-sample and out-of-sample Sharpe ratios of multi-factor asset pricing model”. In: *Journal of Financial Economics* 155, p. 103837.
- Kan, R., X. Wang, and G. Zhou (2022). “Optimal Portfolio Choice with Estimation Risk: No Risk-free Asset Case”. In: *Management Science* 63.3, pp. 2047–2068.
- Kan, R. and G. Zhou (2007). “Optimal Portfolio Choice with Parameter Uncertainty”. In: *Journal of Financial and Quantitative Analysis* 42.3, pp. 621–656.
- Koijen, R. S. J. and M. Yogo (2019). “A Demand System Approach to Asset Pricing”. In: *Journal of Political Economy* 127.4, pp. 1475–1515.
- Kozak, S., S. Nagel, and S. Santosh (2020). “Shrinking the Cross-Section”. In: *Journal of Financial Economics* 135.2, pp. 271–292.
- Lassance, N. and A. Martín-Utrera (2025). “Does the Factor Zoo Pay Off? A Portfolio View on Mispricing and the Limited Gains from New Anomalies”. Working Paper.
- Ledoit, O. and M. Wolf (2017). “Nonlinear Shrinkage of the Covariance Matrix for Portfolio Selection: Markowitz Meets Goldilocks”. In: *Review of Financial Studies* 30.12, pp. 4349–4388.
- Long, J. B. D., A. Shleifer, L. H. Summers, and R. J. Waldmann (Aug. 1990). “Noise Trader Risk in Financial Markets”. In: *Journal of Political Economy* 98.4, pp. 703–738.
- MacKinlay, A. C. and L. Pástor (2000). “Asset Pricing Models: Implications for Expected Returns and Portfolio Selection”. In: *Review of Financial Studies* 13.4, pp. 883–916.
- Marrow, B. and S. Nagel (2024). “Real-Time Discovery and Tracking of Return-Based Anomalies”. Working Paper.

- Martin, I. W. and S. Nagel (2022). “Market Efficiency in the Age of Big Data”. In: *Journal of Financial Economics* 145.1, pp. 154–177.
- McLean, R. D. and J. Pontiff (2016). “Does Academic Research Destroy Stock Return Predictability?” In: *Journal of Finance* 71.1, pp. 5–32.
- Merton, R. C. (1980). “On Estimating the Expected Return on the Market: An Exploratory Investigation”. In: *Journal of Financial Economics* 8, pp. 323–361.
- Newey, W. K. and K. D. West (1987). “A Simple, Positive-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix”. In: *Econometrica* 55.3, pp. 703–708.
- Newey, W. K. and K. D. West (1994). “Automatic Lag Selection in Covariance Matrix Estimation”. In: *Review of Economic Studies* 61.4, pp. 631–653.
- Pénasse, J. (2022). “Understanding Alpha Decay”. In: *Management Science* 68.5, pp. 3966–3973.
- Shleifer, A. and R. W. Vishny (1997). “The Limits of Arbitrage”. In: *Journal of Finance* 52.1, pp. 35–55.

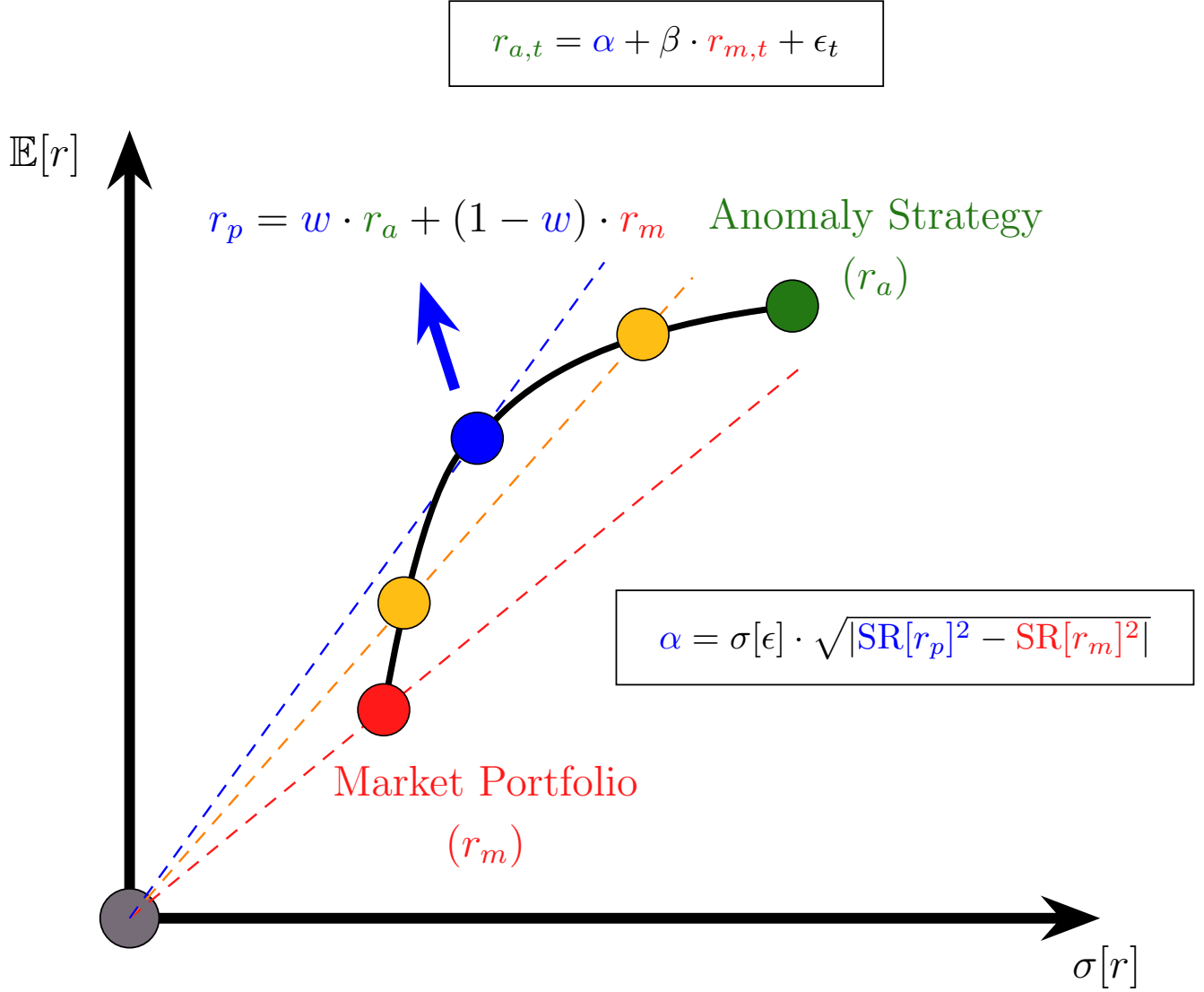
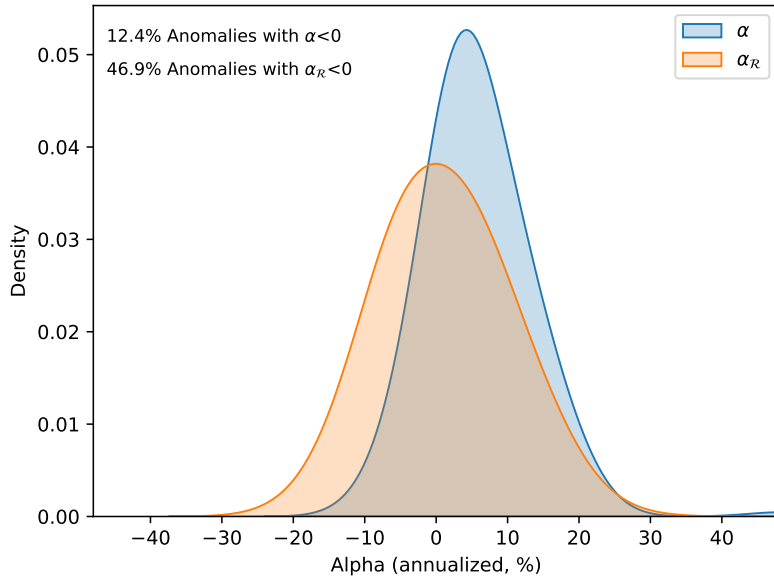
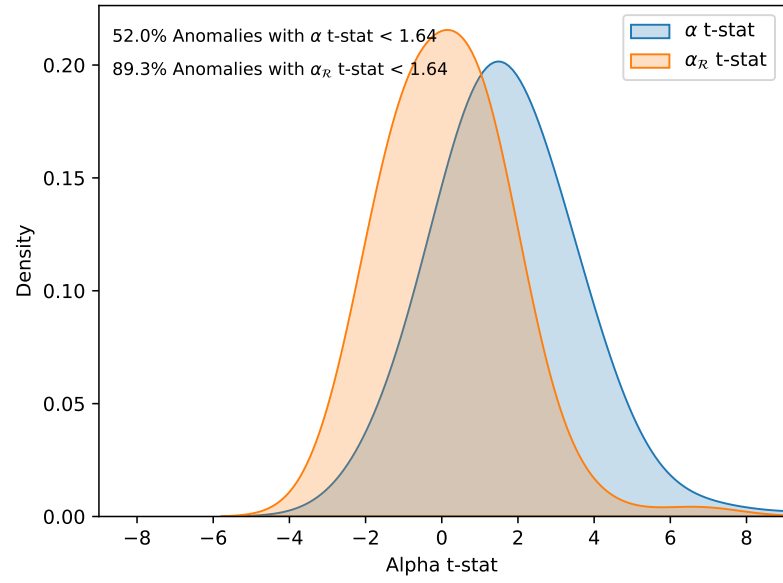
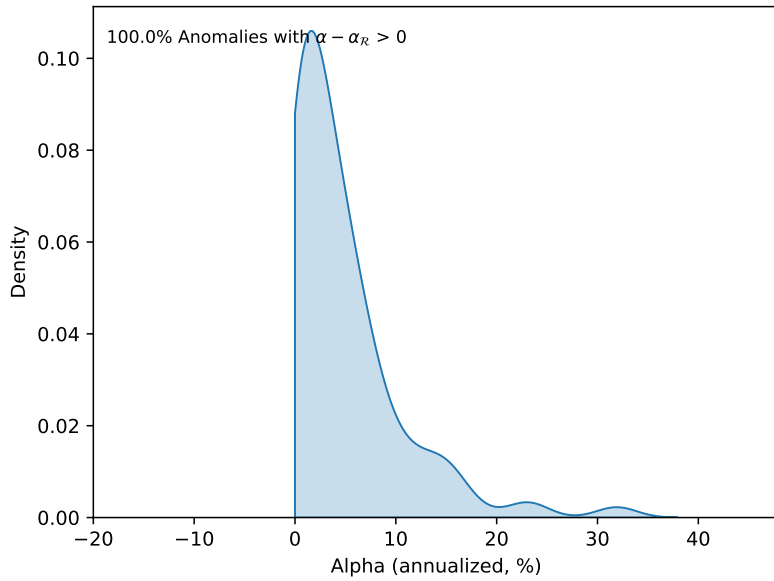


Figure 2
Conceptual Relation Between Alphas and Sharpe Ratios

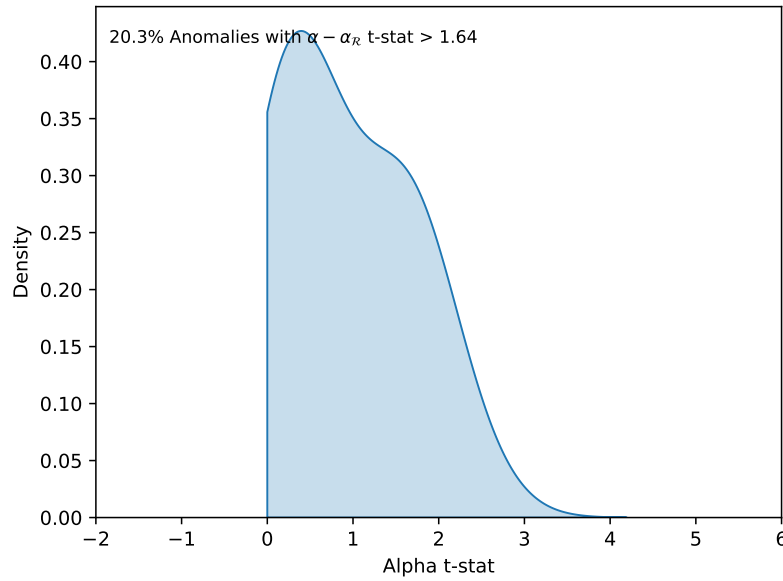
This figure provides a mean-variance diagram in excess return space that reflects the conceptual relation between alphas and Sharpe ratios (for a case with $w > 0$). The red dot reflects the market portfolio (r_m) while the green dot reflects the anomaly strategy (r_a). The blue dot captures the maximum Sharpe ratio portfolio that can be formed by combining the market portfolio with the anomaly strategy, $r_p = (1 - w) \cdot r_{m,t} + w \cdot r_{a,t}$. Using w maximizes the Sharpe ratio (i.e., line slope) increase from the red dot to the blue dot by construction. As such, replacing w with $w_{\mathcal{R}}$ (the optimal weight given the real time information set $\mathcal{R}$) leads to a lower Sharpe ratio increase whether $w_{\mathcal{R}} > w$ or $w_{\mathcal{R}} < w$, as can be seen by comparing the slopes of the blue and orange lines. Section 1 provides more details on the relation between alphas and Sharpe ratios, including an analysis of this mean-variance diagram.

(a) Alphas (α and $\alpha_{\mathcal{R}}$)

(b) t-stats for Alphas

(c) Alpha Differences ($\alpha - \alpha_{\mathcal{R}}$)

(d) t-stats for Alpha Differences

**Figure 3****Distributions of Alphas and their t-stats Across Anomalies (Main Results)**

This figure contains density plots for the distributions of CAPM alphas and their t-statistics across 177 anomaly strategies. We form each anomaly strategy as the long-short decile portfolio with weights as per the original publication. We consider factor regression alphas (α) and real time investor alphas ($\alpha_{\mathcal{R}}$) as well as their differences ($\alpha - \alpha_{\mathcal{R}}$). These alphas differ based on the weights used to combine the anomaly strategy with the market portfolio. Factor regression alphas (real time investor alphas) use weights calculated from the post-publication (pre-publication) period to build a portfolio over the post-publication period, with the pre-publication period ending in December of the publication year. Consequently, factor regression alphas are equivalent to intercepts from factor regressions over the post-publication period, but real time investor alphas are not. The t-statistics are based on the bootstrap procedure described in Internet Appendix B. Results are based on monthly returns and annualized (approximately) by multiplying by 12. Section 1 presents the alpha definitions, Section 2 provides more details about the estimation of alphas, and Subsection 3.1 discusses the results from this figure.

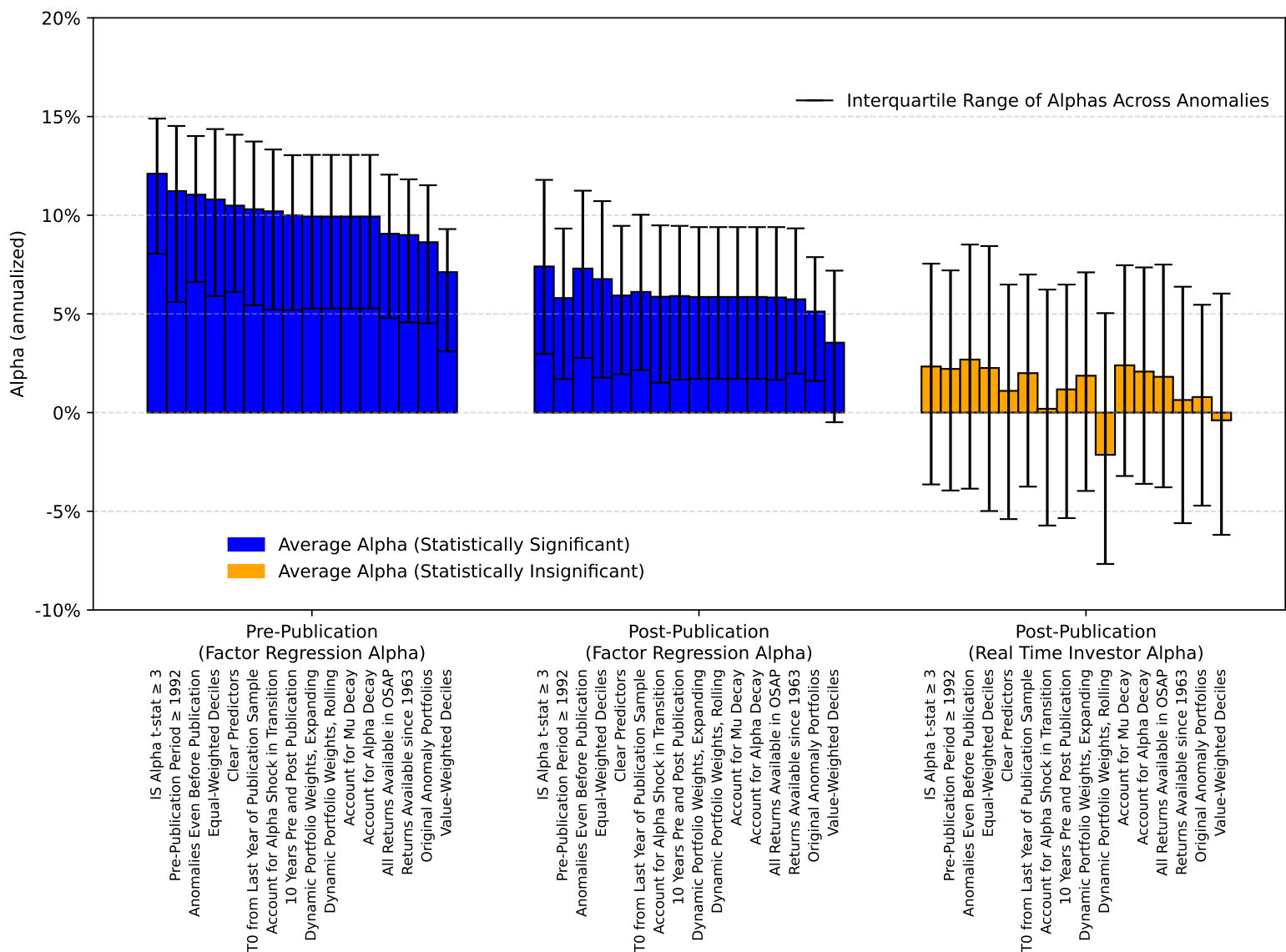


Figure 4

Average and Interquartile Range of Alpha Distribution Across Anomalies (Robustness Analysis)

This figure contains a barplot that summarizes the results from the different specifications we study in our robustness analysis (effectively extending Figure 1 to the different specifications we explore). There are three blocks of bars, with each bar reporting an average alpha. The first block focuses on factor regression alphas pre-publication, the second block focuses on factor regression alphas post-publication, and the third block focuses on real time investor alphas post-publication. The intervals around each bar correspond to the interquartile range of the alpha distribution across anomalies, with the lower and upper bounds indicating the 25th and 75th percentiles, respectively. We sort specifications based on the magnitude of their average factor regression alphas pre-publication and color bars based on whether the average alpha is statistically significant at the 5% level (blue) or not (orange). The t-statistics are based on the bootstrap procedure described in Internet Appendix B. Results are based on monthly returns and annualized (approximately) by multiplying by 12. Section 1 presents the alpha definitions, Section 2 provides more details about the estimation of alphas, and Subsection 3.2 discusses the results from this figure.

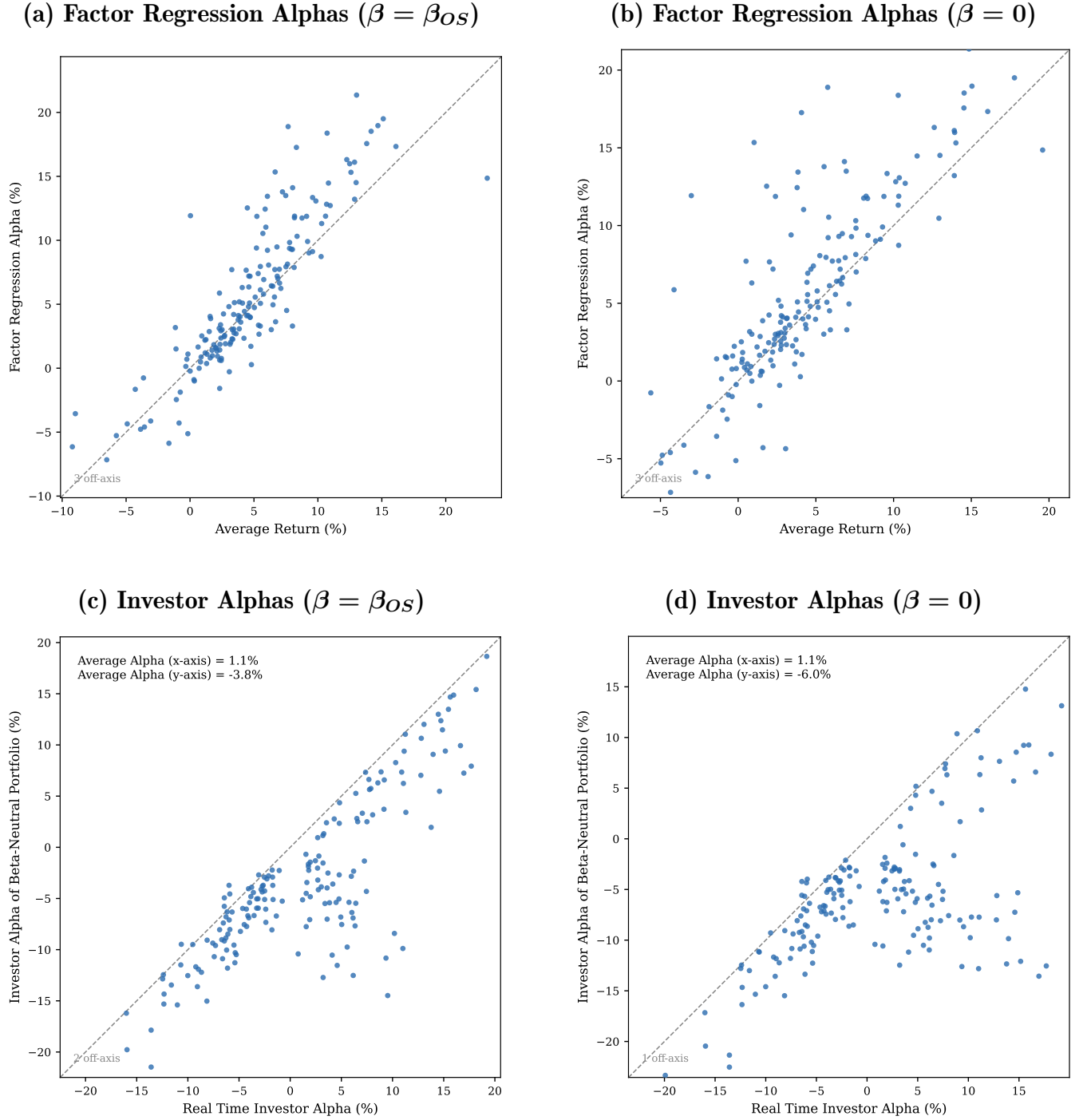
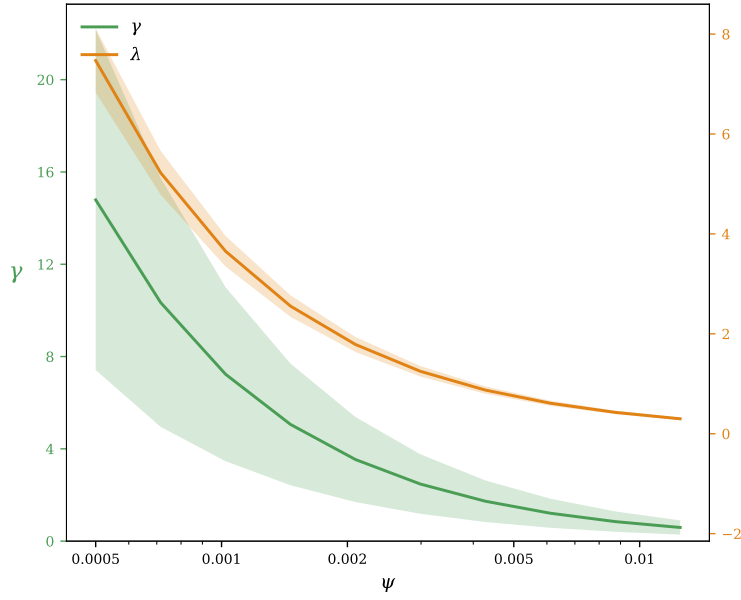
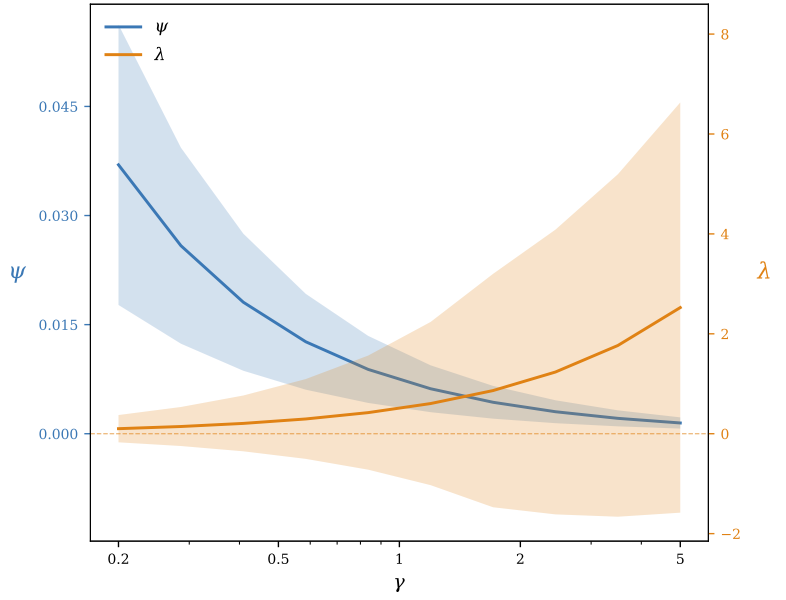
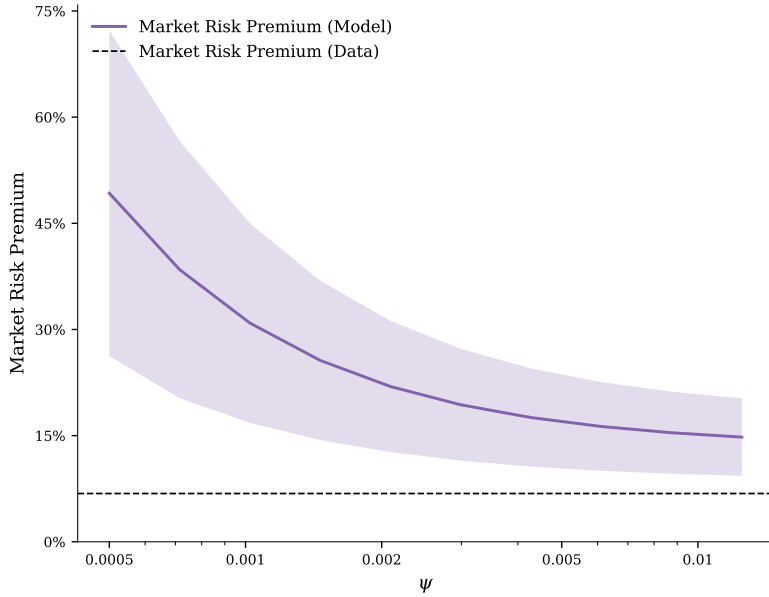
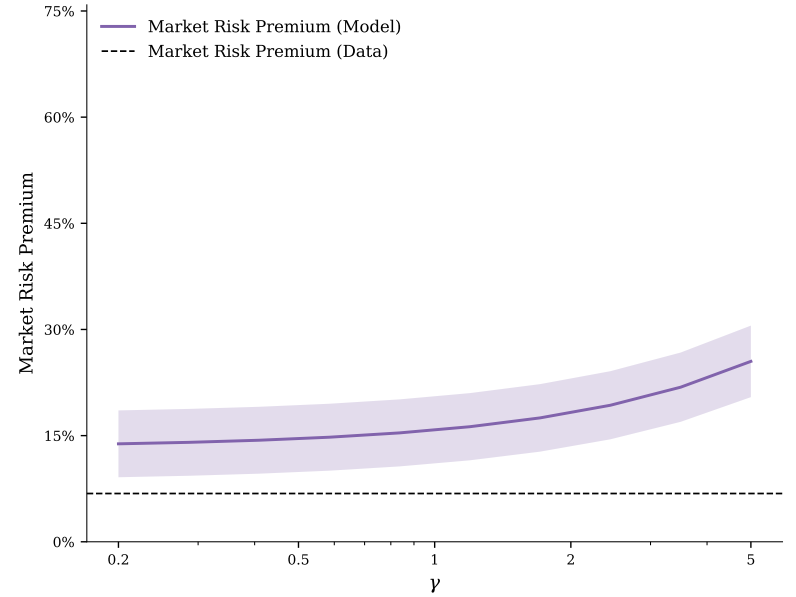


Figure 5
Alphas of Beta-Neutral Portfolios

This figure contains scatterplots studying the alphas of beta-neutral portfolios post-publication for the 177 anomaly strategies in our baseline sample. Panels (a) and (b) compare the average return of a beta-neutral portfolio with the anomaly's factor regression alpha. In Panel (a), the beta-neutral portfolio is constructed in real time as $r_{a,t} - \beta_{OS} \cdot r_{m,t}$, where β_{OS} is estimated over the pre-publication period. Panel (b) imposes the approximation $\beta = 0$, in which case the beta-neutral portfolio is represented by the anomaly return itself. Panels (c) and (d) compare the real time investor alpha obtained by combining the anomaly with the market portfolio using the pre-publication optimal weight with the investor alpha of the corresponding beta-neutral portfolio. Panel (c) constructs the beta-neutral portfolio using β_{OS} whereas Panel (d) imposes $\beta = 0$. The dashed lines are 45-degree lines. Results are based on monthly returns and annualized approximately by multiplying by 12. Section 1 presents the alpha definitions, Section 2 provides more details about the estimation of alphas, and Subsection 3.3 discusses the results from this figure.

(a) Structural Parameter Estimates (Fixing ψ)**(b) Structural Parameter Estimates (Fixing γ)****(c) Implied Market Risk Premium (Fixing ψ)****(d) Implied Market Risk Premium (Fixing γ)****Figure 6****Model Estimation: Structural Parameters and Implied Market Risk Premium**

This figure reports structural parameter estimates and the market risk premium implied by the model in Subsection 4.1. The results are based on the structural estimation in the last four columns of the final two rows of Table 3, which uses the 115 principal component factors constructed from anomaly returns available throughout the common July 1963 to December 2023 sample. Because Equation 23 identifies the products $A = \psi \cdot \lambda$ and $C = \psi \cdot \gamma$, but does not separately identify ψ , γ , and λ , the figure fixes one structural parameter and recovers the other two. Panel (a) fixes ψ over the interval $[\bar{\psi}/5, \bar{\psi} \cdot 5]$ where $\bar{\psi} = 0.0025$, and reports the corresponding estimates of γ and λ . Panel (b) fixes γ over the interval $[\bar{\gamma}/5, \bar{\gamma} \cdot 5]$ where $\bar{\gamma} = 1$, and reports the corresponding estimates of ψ and λ . Panels (c) and (d) report the market risk premium implied by Equation 18 using the structural parameters from Panels (a) and (b), respectively. The solid lines report point estimates, while the shaded areas report 95% confidence intervals. Implied market risk premia are annualized approximately by multiplying monthly values by 12. Subsection 4.2 provides more details about the estimation and discusses the results from this figure.

Table 1
Summary Statistics Across our Baseline 177 Anomaly Strategies

This table reports summary statistics for the baseline sample of 177 anomaly strategies we study, with each anomaly strategy reflecting a long-short decile portfolio with weights as per the original anomaly publication. The table includes information on publication years, the first and last year of the publication sample, the first and last year of anomaly strategy returns available in the OSAP dataset, the number of months in the pre- and post-publication periods, average returns, CAPM market risk premia ($\beta \cdot \mathbb{E}[r_m]$), and factor regression CAPM alphas. The table also provides t-statistics for these factor regression alphas, based on both Newey and West (1987, 1994) and the bootstrap procedure described in Internet Appendix B. These statistics highlight the performance of anomaly portfolios both before and after their publication, emphasizing changes in returns and alphas over time. Results are based on monthly returns and annualized (approximately) by multiplying by 12. Section 1 presents the alpha definitions, Section 2 provides more details about the estimation of alphas, and Subsection 2.2 discussing the results from this table.

		Mean	Median	Min	Q10%	Q25%	Q75%	Q90%	Max
Publication Year		2004	2006	1973	1993	2001	2009	2012	2016
First Year (of Publication Sample)		1970	1970	1926	1962	1963	1979	1986	2002
Last Year (of Publication Sample)		1999	2001	1968	1989	1995	2004	2009	2014
First Year (of Returns Available)		1955	1953	1926	1927	1931	1971	1983	2001
Last Year (of Returns Available)		2023	2023	2013	2023	2023	2023	2023	2023
Full Sample		638	642	240	445	534	726	738	1170
# of Months	Pre-Publication	412	420	106	210	294	522	582	1038
	Post-Publication	226	204	84	121	157	264	360	600
Full Sample		7.4%	6.1%	-4.3%	2.0%	4.0%	9.5%	13.4%	46.3%
Average Return	Pre-Publication	9.1%	7.6%	-4.5%	2.5%	4.6%	11.4%	16.2%	49.4%
	Post-Publication	4.5%	3.7%	-5.6%	-0.7%	1.4%	6.7%	10.3%	41.9%
Full Sample		-1.0%	-0.7%	-9.9%	-3.9%	-1.8%	0.0%	0.8%	5.7%
$\beta \cdot \mathbb{E}[r_m]$	Pre-Publication	-0.9%	-0.5%	-9.5%	-3.1%	-1.4%	0.0%	0.7%	9.8%
	Post-Publication	-1.4%	-0.7%	-14.9%	-5.8%	-2.5%	0.4%	2.2%	7.4%
Full Sample		8.4%	7.3%	-1.2%	1.9%	4.1%	11.5%	15.9%	49.4%
CAPM Alpha	Pre-Publication	9.9%	8.9%	0.7%	2.9%	5.3%	13.1%	17.4%	50.4%
	Post-Publication	5.9%	4.4%	-9.7%	-0.8%	1.7%	9.4%	14.3%	48.5%
Full Sample		(4.22)	(3.84)	(-0.58)	(1.38)	(2.58)	(5.73)	(7.30)	(14.94)
CAPM Alpha t-stat (Newey-West)	Pre-Publication	(4.29)	(3.64)	(0.23)	(1.53)	(2.48)	(5.69)	(7.44)	(14.22)
	Post-Publication	(1.67)	(1.62)	(-1.97)	(-0.29)	(0.56)	(2.65)	(3.64)	(9.42)
Full Sample		(4.65)	(4.19)	(-0.94)	(1.64)	(2.66)	(6.52)	(8.12)	(16.14)
CAPM Alpha t-stat (Bootstrap)	Pre-Publication	(4.67)	(4.20)	(0.37)	(1.76)	(2.63)	(6.15)	(8.15)	(15.42)
	Post-Publication	(1.70)	(1.54)	(-1.82)	(-0.30)	(0.58)	(2.67)	(3.72)	(10.36)

Table 2
Distributions of Alphas and their t-stats Across Anomalies (Main Results)

This table reports our main results for the distributions of CAPM alphas and their t-statistics across the 177 anomaly strategies we study, with each anomaly strategy reflecting a long-short decile portfolio with weights as per the original anomaly publication. We consider factor regression alphas (α) and real time investor alphas ($\alpha_{\mathcal{R}}$) as well as their differences ($\alpha - \alpha_{\mathcal{R}}$). These alphas differ based on the weights used to combine the anomaly strategy with the market portfolio. Factor regression alphas (real time investor alphas) use weights calculated from the post-publication (pre-publication) period to build a portfolio over the post-publication period, with the pre-publication period ending in December of the publication year. Consequently, factor regression alphas are equivalent to intercepts from factor regressions over the post-publication period, but real time investor are not. Results are based on monthly returns and annualized (approximately) by multiplying by 12. Section 1 presents the alpha definitions, Section 2 provides more details about the estimation of alphas, and Subsection 3.1 discusses the results from this table.

PANEL A: Pre-Publication

	Mean	Median	Min	Q10%	Q25%	Q75%	Q90%	Max
Factor Regression Alpha (α)	9.9%	8.9%	0.7%	2.9%	5.3%	13.1%	17.4%	50.4%
α t-stat	(4.67)	(4.20)	(0.37)	(1.76)	(2.63)	(6.15)	(8.15)	(15.42)
Optimal Weight on Anomaly	0.62	0.63	0.15	0.38	0.52	0.76	0.85	1.07

PANEL B: Post-Publication

	Mean	Median	Min	Q10%	Q25%	Q75%	Q90%	Max
Factor Regression Alpha (α)	5.9%	4.4%	-9.7%	-0.8%	1.7%	9.4%	14.3%	48.5%
α t-stat	(1.70)	(1.54)	(-1.82)	(-0.30)	(0.58)	(2.67)	(3.72)	(10.36)
Optimal Weight on Anomaly	-0.22	0.36	-40.74	-0.17	0.22	0.52	0.63	0.84
Real Time Investor Alpha ($\alpha_{\mathcal{R}}$)	1.1%	1.6%	-19.9%	-8.8%	-5.4%	6.4%	12.8%	25.3%
$\alpha_{\mathcal{R}}$ t-stat	(0.15)	(0.31)	(-2.57)	(-1.58)	(-1.00)	(1.06)	(1.65)	(6.97)
$\alpha - \alpha_{\mathcal{R}}$	4.7%	3.0%	0.0%	0.2%	0.8%	6.2%	12.1%	32.0%
$\alpha - \alpha_{\mathcal{R}}$ t-stat	(0.88)	(0.66)	(0.00)	(0.05)	(0.22)	(1.56)	(1.87)	(2.64)

Table 3
Model Estimation: Explaining Factor Regression Alphas

This table reports results from estimating Equation 23 (which follows from the model in Subsection 4.1):

$$\hat{\alpha}_a = A + B \cdot \beta_a + C \cdot \beta_a \cdot \frac{\sigma_m^2}{\sigma_a^2} + \varepsilon_a$$

where $\hat{\alpha}_a$ and β_a are the factor regression alpha and beta of anomaly a , σ_a is the volatility of anomaly a , and σ_m is the market volatility. The first four columns are based on an Ordinary Least Squares regression (OLS) estimation that does not impose restrictions on the reduced form parameters (A , B , and C) while the last four columns are based on a Non-Linear Least Squares estimation that imposes the model structural restrictions: $A = \psi \cdot \lambda \geq 0$, $C = \psi \cdot \gamma \geq 0$, and $B = -(A + C) \leq 0$ (through $\gamma > 0$, $\psi \geq 0$, and $\lambda \geq 0$). The first two rows use all 177 anomalies studied in Subsection 3.1, with $\hat{\alpha}_a$ values obtained from the post-publication period and β_a , σ_a , and σ_m values obtained from the pre-publication period. The other rows use the 115 anomalies with return data available throughout the common July 1963 to December 2023 sample, with α_a , β_a , σ_a , and σ_m estimated over this full period. The last two rows transform anomalies into principal component factors before the estimation. Section 1 presents the alpha definitions, Section 2 provides more details about the estimation of alphas, and Subsection 4.2 discusses the results from this table.

		Reduced Form (OLS)				Structural (NLS)			
		<i>A</i>	<i>B</i>	<i>C</i>	<i>R</i> ²	<i>A</i>	<i>B</i>	<i>C</i>	<i>R</i> ²
177 Anomalies	Estimate	0.0041	-0.0099	0.0032	17.1%	0.0046	-0.0086	0.0040	16.0%
(pre- vs post-pub)	(t-stat)	(12.02)	(-5.84)	(3.21)		(18.49)	(-6.90)	(3.62)	
115 Anomalies	Estimate	0.0060	-0.0061	0.0015	15.0%	0.0058	-0.0063	0.0005	14.5%
(1963 to 2023)	(t-stat)	(29.78)	(-5.25)	(1.47)		(38.80)	(-6.86)	(0.65)	
115 PC Factors	Estimate	0.0037	-0.0113	0.0012	32.5%	0.0037	-0.0111	0.0074	31.6%
(1963 to 2023)	(t-stat)	(23.56)	(-4.95)	(0.68)		(23.32)	(-5.54)	(3.76)	

Internet Appendix

“Real Time Investor Alphas and the Survival of the Anomaly Zoo”

By Andrei S. Gonçalves, Johnathan A. Loudis, and Richard E. Ogden

Contents

A	Derivation of Alpha Equation	IA.2
B	Bootstrap Procedure for Standard Errors	IA.4
C	Real Time Investor Alphas with Shrinkage in the Weights	IA.6
C.1	Bayesian Weight Estimation	IA.6
C.2	Machine Learning Weight Estimation	IA.9
D	Derivations for the Model in Section 4	IA.12
D.1	Deriving the Optimal Portfolio Allocation	IA.12
D.2	Deriving Equilibrium Expected Returns and Alphas	IA.13
D.3	Deriving Equilibrium Expected Returns and Alphas with Noise Traders . . .	IA.14
E	An Endowment Economy Version of the Model in Section 4	IA.16

A Derivation of Alpha Equation

This section derives (a generalization to) Equation 2 in the main text.

Suppose we have a factor model such that

$$r_{a,t} = \alpha + \sum_{k=1}^K \beta_k \cdot f_{k,t} + \varepsilon_t \quad (\text{IA.1})$$

where $f_{k,t}$ reflects the long-short return on the zero-cost tradable factor k and $r_{a,t}$ reflects the long-short return on a zero-cost tradable strategy (i.e., an anomaly strategy).

Our objective is to derive an expression that connects α to Sharpe ratios. For that, it is useful to note that a Stochastic Discount Factor (SDF) that perfectly prices all $f_{k,t}$ can be written as $SDF_t = a + b \cdot \sum_{k=1}^K w_k \cdot f_{k,t} = a + b \cdot f_t^*$, where w_k are the maximum Sharpe ratio weights for a portfolio including only the factors (see Chapter 6 in Cochrane (2005)). As such, the α from Equation IA.1 is identical to the α from equation²⁰

$$r_{a,t} = \alpha + \beta \cdot f_t^* + \epsilon_t \quad (\text{IA.2})$$

Equation 6.6.17 in Campbell, Lo, and MacKinlay (1997) (originally derived in Gibbons, Ross, and Shanken (1989)) shows that the α from Equation IA.2 satisfies

$$\frac{\alpha^2}{\text{Var}[\epsilon]} = \text{SR}[r_p^*]^2 - \text{SR}[f^*]^2 \quad (\text{IA.3})$$

where $r_{p,t}^* = w \cdot r_{a,t} + (1 - w) \cdot f_t^*$ represents the ex-post maximum Sharpe ratio portfolio that can be formed with $r_{a,t}$ and f_t^* . Consequently, we have

$$|\alpha| = \sigma[\epsilon] \cdot \sqrt{\text{SR}[r_p^*]^2 - \text{SR}[f^*]^2} \quad (\text{IA.4})$$

Since $r_{p,t}^*$ is the ex-post maximum Sharpe ratio portfolio that can be formed with f_t^* and $r_{a,t}$, we have $\Delta = \text{SR}[r_p^*] - \text{SR}[f^*] \geq 0$, with this inequality holding with equality if and only if $\alpha = 0$. Moreover, $\alpha > 0$ ($\alpha < 0$) implies $w > 0$ ($w < 0$). Consequently, Equation IA.4 can

²⁰Note that the residual terms in Equations IA.1 and IA.2 differ so that the idiosyncratic volatility (and consequently R^2 values) differ between the two equations. For our purpose, it only matters that the alphas are the same.

be alternatively written as

$$\alpha = \text{sign}[w] \cdot \sigma[\epsilon] \cdot \sqrt{|\mathbb{SR}[r_p^*]^2 - \mathbb{SR}[f^*]^2|} \quad (\text{IA.5})$$

Note that if $r_{m,t}$ is the only factor (i.e., $f_t^* = r_{m,t}$), then Equation [IA.5](#) becomes

$$\alpha = \text{sign}[w] \cdot \sigma[\epsilon] \cdot \sqrt{|\mathbb{SR}[r_p]^2 - \mathbb{SR}[r_m]^2|} \quad (\text{IA.6})$$

where $r_{p,t} = w \cdot r_{a,t} + (1 - w) \cdot r_{m,t}$ so that we have Equation [2](#) in the main text.

Note also that the coefficient of determination of the factor model in Equation [IA.2](#) is given by $\mathbb{Cor}[r_a, f^*]^2 = 1 - \mathbb{Var}[\epsilon]/\mathbb{Var}[r_a]$, implying

$$\sigma[\epsilon] = \sigma[r_a] \cdot \sqrt{1 - \mathbb{Cor}[r_a, f^*]^2} \quad (\text{IA.7})$$

While we do not directly use Equation [IA.7](#) in the main text, it is useful in highlighting that $\sigma[\epsilon]$ is a simple function of the r_a volatility and the correlation between r_a and f^* .

B Bootstrap Procedure for Standard Errors

In the main text, we use bootstrap standard errors for alphas. This section describes our bootstrap procedure.

For the bootstrap standard errors associated with a given anomaly strategy, we proceed as follows. We sample (with replacement) months from the anomaly specific pre-publication and post-publication periods separately (this approach conserves the structure of returns present in the data). We then concatenate the sampled pre-publication and post-publication periods to form a full bootstrap sample for the given anomaly. Using this sample, we calculate full sample factor regression alphas, pre-publication factor regression alphas, post-publication factor regression alphas, and post-publication real time investor alphas, as well as the difference between post-publication factor regression alphas and real time investor alphas (which we refer to as $\alpha - \alpha_{\mathcal{R}}$). We repeat this process 10,000 times and obtain standard errors from the cross-simulation standard deviation of each of these alphas. In the case of alphas computed using Bayesian optimal weights, we hold fixed the prior distribution of alphas used. In the case of the machine learning methods (KNS and BPZ), we repeat the 4-fold cross validation design to select the hyperparameters in each simulation. Figure [IA.1](#) contrasts (for post-publication factor regression alphas) our bootstrap t-statistics with t-statistics obtained from Newey and West ([1987](#), [1994](#)), which are common in the literature. The two types of t-statistics tend to be very similar.

For the bootstrap standard errors associated with average alphas across anomalies, we cannot sample pre-publication and post-publication periods separately (as they differ across anomalies), and thus proceed as follows. We start by splitting the sample between early (1963 to 2002) and late (2003 to 2023) periods. Then, we sample (with replacement) months from each of these periods separately and obtain anomaly returns for the sampled months accordingly (we sample $12 \times 40 = 480$ months from the early period and $12 \times 21 = 252$ months from the late period). For each anomaly, we classify each month as a pre-publication month or a post-publication month based on whether it comes from the given anomaly’s pre-publication

or post-publication period. Finally, we calculate full sample factor regression alphas, pre-publication factor regression alphas, post-publication factor regression alphas, and post-publication real time investor alphas, as well as the difference between post-publication factor regression alphas and real time investor alphas (which we refer to as $\alpha - \alpha_{\mathcal{R}}$). We then average these values across anomalies. Finally, we repeat this process 10,000 times and obtain standard errors for average alphas from the cross-simulation standard deviation of each of these average alphas. The only exception is the specification “Anomalies Even Before Publication” (in Figure 4) as we repeat the process 1,000 times in that case (due to the long computing time for that specification).

C Real Time Investor Alphas with Shrinkage in the Weights

Section 3 indicates that while factor regression alphas remain high in the post-publication period, real time investor alphas are relatively close to zero, on average. This outcome stems from the fact that the ex-post optimal anomaly weights in the pre-publication period are significantly higher than those in the post-publication period. So, it is reasonable to expect that applying some shrinkage in the weight estimation might achieve better real time investor alphas. Accordingly, this section examines real time investor alphas obtained using Bayesian shrinkage (Subsection C.1) and machine learning shrinkage through regularization methods (Subsection C.2). The key result is that shrinkage applied in real time to estimate optimal weights does not change our main conclusion that real time investor alphas are close to zero, on average.

C.1 Bayesian Weight Estimation

To apply Bayesian shrinkage when estimating optimal anomaly weights, we consider a simple 1-factor index model, $r_{a,t} = \alpha + \beta \cdot r_{m,t} + \epsilon_t$, which implies the μ and Σ can be written as

$$\mu = \begin{bmatrix} \mathbb{E}[r_m] \\ \alpha + \beta \cdot \mathbb{E}[r_m] \end{bmatrix} \quad \text{and} \quad \Sigma = \begin{bmatrix} \text{Var}[r_m] & \beta \cdot \text{Var}[r_m] \\ \beta \cdot \text{Var}[r_m] & \text{Var}[r_a] \end{bmatrix} \quad (\text{IA.8})$$

Our Bayesian weight estimation (over the pre-publication period) uses Ordinary Least Squares (OLS) to estimate β but shrinks α towards zero (i.e., it uses the prior that anomalies have no alpha), effectively shrinking the anomaly weight towards zero.²¹ Specifically, we use

$$\alpha_{BAY} = \left(\frac{1/\sigma_{OLS,\alpha}^2}{1/\sigma_\alpha^2 + 1/\sigma_{OLS,\alpha}^2} \right) \cdot \alpha_{OLS} \quad (\text{IA.9})$$

²¹We also consider a simple strategy that allocates 50% of the portfolio to the market index and 50% to the anomaly strategy. This approach can be thought of as a generic Bayesian method designed to combat estimation error and often performs better than more complex methods in mean-variance applications with limited time series (see DeMiguel, Garlappi, and Uppal (2009)). However, in the context of our analysis, it leads to real time investor alpha results that are similar to our baseline real time investor alpha results (reported in Section 3). So, to conserve space, we do not tabulate the results from this analysis.

where $\sigma_{OLS,\alpha}$ reflects the standard error of α_{OLS} (estimated by Newey and West (1987, 1994)) whereas σ_α reflects the prior uncertainty about alpha. In turn, we obtain

$$\mu_{BAY} = \begin{bmatrix} \bar{r}_m \\ \alpha_{BAY} + \beta_{OLS} \cdot \bar{r}_m \end{bmatrix} \quad \text{and} \quad \Sigma_{BAY} = \begin{bmatrix} \bar{\sigma}_m^2 & \beta_{OLS} \cdot \bar{\sigma}_m^2 \\ \beta_{OLS} \cdot \bar{\sigma}_m^2 & \bar{\sigma}_a^2 \end{bmatrix} \quad (\text{IA.10})$$

and get the (Bayesian) OS anomaly weight by using μ_{BAY} and Σ_{BAY} as a replacement for μ_{OS} and Σ_{OS} in Equation 9.

Our empirical analysis explores three arbitrary priors (i.e., three arbitrary values of σ_α) to illustrate the quantitative effect of Bayesian shrinkage. However, an investor would need to select σ_α for each anomaly using its respective pre-publication sample. We account for that by using an empirical Bayesian estimator as a fourth specification for σ_α . Consider an anomaly with pre-publication sample ending in month T_0 . We start by collecting all anomalies that were published up to T_0 . We then estimate α_{OLS} for each of these anomalies using the sample ending in T_0 . Finally, we obtain σ_α as the standard deviation of α_{OLS} across these anomalies. We repeat this process for each anomaly so that σ_α varies across anomalies based on the sequence of published anomalies.²²

Table IA.1 and Figure IA.2 summarize our Bayesian OS alpha results. Panel A and Figure IA.2(a) consider an extremely informative prior ($\sigma_\alpha = 0.1\%$).²³ In this case, the Bayesian

²²In untabulated results (that are similar to the reported results), we also consider a Bayesian estimator that shrinks both α and β toward zero (since each anomaly portfolio is a long-short equity portfolio, it is sensible to use $\beta = 0$ as the benchmark). Specifically, write the regression in Equation 1 as $y = X\theta + \epsilon$, where $\theta' = [\alpha \ \beta]$. Then, the Bayesian estimator for θ is given by

$$\hat{\theta}_{BAY} = (V_o^{-1} + V_{OLS}^{-1})^{-1} \cdot [V_o^{-1}\theta_0 + V_{OLS}^{-1}\theta_{OLS}]$$

where $\theta'_0 = [\alpha_0 \ \beta_0] = [0 \ 0]$, $\theta_{OLS} = (X'X)^{-1}(X'y)$ is the OLS estimate, and V_{OLS} is the covariance matrix of θ_{OLS} (estimated by Newey and West (1987, 1994)). V_o reflects the prior uncertainty about θ , which is a symmetric matrix with diagonal elements σ_α^2 and σ_β^2 , and off-diagonal element $\sigma_{\alpha,\beta}$. Similar to our main Bayesian analysis, we use an empirical Bayesian approach to estimate σ_α^2 , σ_β^2 , and $\sigma_{\alpha,\beta}$. Consider an anomaly with pre-publication sample ending in month T_0 . We start by collecting all anomalies that were published up to T_0 . We then estimate θ_{OLS} for each of these anomalies using the sample ending in T_0 . Finally, we obtain σ_α and σ_β as the standard deviations of α_{OLS} and β_{OLS} across these anomalies, and $\sigma_{\alpha,\beta}$ as the sample covariance between α_{OLS} and β_{OLS} across these anomalies. We repeat this process for each anomaly so that σ_α^2 , σ_β^2 , and $\sigma_{\alpha,\beta}$ vary across anomalies based on the sequence of published anomalies.

²³As with alpha values, we state σ_α in annualized alpha units. For instance, our annual $\sigma_\alpha = 0.1\%$

optimal weights are very close to zero. The mean weight is 0.01 and even the 90% quantile of the weight distribution is only 0.04. Consequently, the investor alphas are very close to zero. In particular, the mean alpha is 0.6% and almost no alpha is statistically different from zero. Panel B and Figure [IA.2\(b\)](#) consider a prior that is still highly informative, but a little less extreme ($\sigma_\alpha = 1.0\%$). In this case, the Bayesian optimal weights are close to the optimal post-publication in-sample weights. In particular, the mean weight is 0.33 and the median weight is 0.29. Consequently, the distribution of investor alphas gets closer to the distribution of factor regression alphas. Panel C and Figure [IA.2\(c\)](#) consider a moderately informative prior ($\sigma_\alpha = 5.0\%$). In this case, the overall results are very similar to the baseline real time investor alpha results. In particular, the Bayesian optimal average and median weights are similar to the pre-publication optimal weights. Consequently, the average and median investor alphas (2.0% and 2.3%, respectively) are relatively close to the average and median real time investor alphas obtained in our baseline analysis (1.1% and 1.6%, respectively).

Clearly, whether a Bayesian investor would experience a distribution of OS alphas that is similar to that of factor regression alphas or not depends on the prior used. In reality, a Bayesian investor would not know the correct prior, but could attempt to infer it from the data on published anomalies. Our empirical Bayesian approach described above captures this idea, with results provided in Panel D of Table [IA.1](#) and Figure [IA.2\(d\)](#). The key finding is that the empirical Bayesian approach leads to results that are similar to the moderately informative prior from Panel C. Specifically, the Bayesian optimal average and median weights are similar to the pre-publication optimal weights. As such, the average and median of real time investor alphas (1.8% and 1.9%, respectively) are relatively close to the average and median investor alphas obtained with the moderately informative prior (2.0% and 2.3%, respectively). The reason is that the average σ_α from the empirical Bayesian approach is 6.2%, which is similar to the $\sigma_\alpha = 5.0\%$ used in the moderately informative prior.²⁴

corresponds to a monthly $\sigma_\alpha = 0.1\%/12$.

²⁴One may conjecture that the high average σ_α in the empirical Bayesian method is driven by early anomalies, when the distribution of anomalies was not wide enough for the empirical Bayesian procedure to reasonably estimate σ_α . However, this is not the case. In particular, Figure [IA.3](#) shows that the empirical Bayesian estimates for σ_α are actually lower in the early periods, when less anomaly strategies were present.

Consequently, we conclude that the distribution of real time investor alphas available to an empirical Bayesian investor is similar to the distribution of real time investor alphas available to a frequentist investor.

C.2 Machine Learning Weight Estimation

Following Davis (2025), we explore two machine learning methods to compute portfolio weights (the KNS method from Kozak, Nagel, and Santosh (2020) and the BPZ method from Bryzgalova, Pelger, and Zhu (2024)).²⁵

These methods are cast in the language of Stochastic Discount Factors (SDFs) since estimating the maximum Sharpe ratio weights of a portfolio is equivalent to estimating the coefficients of a linear SDF. As such, we introduce new notation. Let $f_t = [r_m, r_a]$ be the vector containing the excess returns on the market index and the given anomaly strategy. With this f definition, we have $\mathbb{E}[f] = \mu$ and $\mathbb{V}ar[f] = \Sigma$. Then, consider an SDF given by $M_t = 1 - b'(f_t - \mu)$ so that the asset pricing equation $\mathbb{E}[M_t \cdot f_t] = 0$ yields $b = \Sigma^{-1}\mu$. It is easy to see that the optimal weights in Equation 3 are proportional to b (i.e., $\omega = b/|1'b|$). Moreover, this b can be equivalently written as the solution to the following optimization problem:

$$b = \underset{\{b\}}{\operatorname{argmin}} \quad (\mu - \Sigma b)' \Sigma^{-1} (\mu - \Sigma b) \quad (\text{IA.11})$$

The two machine learning methods we explore rely on generalizations of Equation IA.11.

In the more recent years, they are stable at around $\sigma_\alpha = 7.0\%$, which is even higher than the average σ_α in the empirical Bayesian method.

²⁵Although it may be possible for one to design a machine learning method that performs better than the ones we explore (e.g., by tailoring the machine learning method to recognize the nature of pre-publication vs post-publication data), it is not obvious that such a more advanced method would help us better understand the alphas available to investors in real time. The reason is that typical investors are unlikely to have had knowledge of (or access to) sophisticated machine learning methods historically. Given that our goal is to measure the alphas that anomalies provide to investors in real time, the weights used in our real time investor alphas should be out-of-sample not only in terms of the data used, but also in terms of the technology used. While the KNS and BPZ methods are also subject to this limitation, they can be written as a standard LASSO regression (Bryzgalova, Pelger, and Zhu (2024)), which partially alleviates this concern since LASSO was introduced in Tibshirani (1996), ten years before the median publication year of anomalies in our sample.

The first is the KNS method, which uses $\omega_{KNS} = b_{KNS}/|1'b_{KNS}|$ with

$$b_{KNS} = \underset{\{b\}}{\operatorname{argmin}} \quad (\mu - \Sigma b)' \Sigma^{-1} (\mu - \Sigma b) + \lambda_1 \cdot \|b\|_1 + \lambda_2 \cdot \|b\|_2 \quad (\text{IA.12})$$

while the second is the BPZ method, which uses $\omega_{BPZ} = b_{BPZ}/|1'b_{BPZ}|$ with

$$b_{BPZ} = \underset{\{b\}}{\operatorname{argmin}} \quad ((\mu + \lambda_0 \cdot 1) - \Sigma b)' \Sigma^{-1} ((\mu + \lambda_0 \cdot 1) - \Sigma b) + \lambda_1 \cdot \|b\|_1 + \lambda_2 \cdot \|b\|_2 \quad (\text{IA.13})$$

These methods are regularization techniques and can be equivalently written as a LASSO or LARS regression for computational efficiency (see the simplified discussion in Davis (2025)). For both methods and for each anomaly strategy, we obtain μ and Σ over the respective pre-publication period using the μ_{OS} and Σ_{OS} estimates from Subsection 2.1. We then follow the prior asset pricing literature (e.g., Gu, Kelly, and Xiu (2020), Bryzgalova, Pelger, and Zhu (2024), and Davis (2025)) in using a 4-fold cross validation design to select the hyperparameters (λ_1 and λ_2 in the case of b_{KNS} and λ_0, λ_1 , and λ_2 in the case of b_{BPZ}).²⁶ Finally, we combine the μ_{OS} and Σ_{OS} values and the selected hyperparameters to obtain $\omega_{KNS} = b_{KNS}/|1'b_{KNS}|$ and $\omega_{BPZ} = b_{BPZ}/|1'b_{BPZ}|$ (from Equations IA.12 and IA.13) and form the respective portfolios that are evaluated over the post-publication period.

To ensure we have enough data for the 4-fold design of the prior paragraph, we restrict our machine learning analysis to the 149 anomalies that have portfolio return data for at least 20 years in the pre-publication period (so that each fold has a minimum of 5 years of data). Below, we provides results for these 149 anomalies using the two machine learning weight estimation methods we consider as well as our baseline weight estimation method for comparison.

²⁶Consider the b_{KNS} estimation for the momentum anomaly. We start by creating a grid for the hyperparameters (λ_1 and λ_2) and splitting the momentum pre-publication period into four contiguous samples of equal size (known as folds). Then, for each hyperparameter gridpoint, we select one fold, obtain $\omega_{KNS} = b_{KNS}/|1'b_{KNS}|$ (from Equation IA.12) over the three remaining folds (using μ_{OS} and Σ_{OS} estimated over the respective set of three folds), and calculate the Sharpe ratio of the resulting portfolio over the selected fold. We repeat this process four times (one for each selection of fold) and obtain the average Sharpe ratio (across the four selected folds) for each hyperparameter gridpoint. Finally, we choose the hyperparameter gridpoint with the highest average Sharpe ratio. The process is analogous for b_{BPZ} and other anomalies.

Table [IA.2](#) provides the real time investor alpha ($\alpha_{\mathcal{R}}$) results. Panel A uses our regular weight estimation but applied to the 149 anomalies used in this machine learning analysis. The results are similar to what we find in our baseline 177 anomalies, except that $\alpha_{\mathcal{R}}$ estimates tend to be a little lower in this set of anomalies (e.g., the average alpha here is 0.8% in comparison to 1.1% in our baseline analysis). Panel B provides the results for the KNS method. The KNS optimal weights are similar to the optimal weights from the frequentist method, with a mean of 0.60 and a median of 0.62 (in comparison to a mean of 0.61 and a median of 0.63 from Panel A). Consequently, the KNS $\alpha_{\mathcal{R}}$ distribution is similar to the frequentist $\alpha_{\mathcal{R}}$ distribution, with an average $\alpha_{\mathcal{R}}$ of 0.2% and a median $\alpha_{\mathcal{R}}$ of 0.0%. Panel C provides the results for the BPZ method. The BPZ optimal weights are also similar to the frequentist optimal weights, with a mean of 0.61 and a median of 0.62. As such, the BPZ $\alpha_{\mathcal{R}}$ distribution also resembles the frequentist $\alpha_{\mathcal{R}}$ distribution. In particular, the average $\alpha_{\mathcal{R}}$ is 0.2% and the median $\alpha_{\mathcal{R}}$ is 0.0%.

The fact that KNS and BPZ do not improve upon our baseline frequentist weight estimation is perhaps not that surprising given our setting. Specifically, these methods are regularization techniques designed to create an SDF (or maximum Sharpe ratio portfolio) out of multiple tradable factors. In our setting, we just have two tradable factors (the market and the given anomaly strategy). Therefore, these methods provide little benefit. We consider investors who explore multiple anomaly strategies jointly in [Section 4](#).

D Derivations for the Model in Section 4

This section provides derivations for the model in Section 4.

D.1 Deriving the Optimal Portfolio Allocation

Combining the prior $\mu_{n,t} \sim \mathcal{N}(-\lambda \cdot \sigma_{n,t}^2, \nu \cdot \sigma_{n,t}^4)$ with the signal $\hat{\mu}_{n,t} \sim \mathcal{N}(\mu_{n,t}, \hat{\nu} \cdot \sigma_{n,t}^2)$, we have the Bayesian posterior expectation

$$\begin{aligned}
\mathbb{E}_t[\mu_{n,t}] &= \left(\frac{1}{\hat{\nu} \cdot \sigma_{n,t}^2} + \frac{1}{\nu \cdot \sigma_{n,t}^4} \right)^{-1} \left(\frac{1}{\hat{\nu} \cdot \sigma_{n,t}^2} \cdot \hat{\mu}_{n,t} - \frac{1}{\nu \cdot \sigma_{n,t}^4} \cdot \lambda \cdot \sigma_{n,t}^2 \right) \\
&= \left(\frac{1}{\hat{\nu} \cdot \sigma_{n,t}^2} + \frac{1}{\nu \cdot \sigma_{n,t}^4} \right)^{-1} \cdot \frac{1}{\sigma_{n,t}^2} \cdot \left(\frac{1}{\hat{\nu}} \cdot \hat{\mu}_{n,t} - \frac{1}{\nu} \cdot \lambda \right) \\
&= \left(\frac{1}{\hat{\nu}} + \frac{1}{\nu \cdot \sigma_{n,t}^2} \right)^{-1} \left(\frac{1}{\hat{\nu}} \cdot \hat{\mu}_{n,t} - \frac{1}{\nu} \cdot \lambda \right) \\
&= \sigma_{n,t}^2 \cdot \left(\frac{\sigma_{n,t}^2}{\hat{\nu}} + \frac{1}{\nu} \right)^{-1} \left(\frac{1}{\hat{\nu}} \cdot \hat{\mu}_{n,t} - \frac{1}{\nu} \cdot \lambda \right) \\
&= \sigma_{n,t}^2 \cdot (\sigma_{n,t}^2 + \psi)^{-1} (\hat{\mu}_{n,t} - \lambda \cdot \psi)
\end{aligned} \tag{IA.14}$$

where $\psi = \hat{\nu}/\nu > 0$. This is Equation 14 in the main text.

Substituting $\text{Var}_t[r_{m,t+1}] = \sigma_{m,t}^2$ and $\mathbb{E}_t[r_{m,t+1}] = \mathbb{E}_t[\mu_{m,t}]$ into the optimal market allocation (Equation 12) yields

$$\omega_{m,t} = (1/\gamma) \cdot \sigma_{m,t}^{-2} \cdot \left(\sigma_{m,t}^2 \cdot (\sigma_{m,t}^2 + \psi)^{-1} (\hat{\mu}_{m,t} - \lambda \cdot \psi) - \gamma \cdot \text{Cov}_t[\omega'_t r_{t+1}, r_{m,t+1}] \right) \tag{IA.15}$$

which is Equation 15 in the main text.

Similarly, substituting $\text{Var}_t[r_{t+1}] = \Sigma_t$ and $\mathbb{E}_t[r_{t+1}] = \mathbb{E}_t[\mu_t]$ into the optimal anomaly allocation (Equation 13) yields

$$\omega_t = (1/\gamma) \cdot \Sigma_t^{-1} \cdot \left(\Sigma_t \cdot (\Sigma_t + \psi \cdot \mathbf{I})^{-1} (\hat{\mu}_t - \lambda \cdot \psi \cdot \mathbf{1}) - \gamma \cdot \text{Cov}_t[r_{t+1}, \omega_{m,t} \cdot r_{m,t+1}] \right) \tag{IA.16}$$

which is Equation 16 in the main text after recognizing that Σ_t is diagonal with $\Sigma_t[a, a] = \sigma_{a,t}^2$.

D.2 Deriving Equilibrium Expected Returns and Alphas

Substituting market clearing ($\omega_{m,t} = 1$ and $\omega_t = 0$) into the optimal market allocation (Equation IA.15), we have

$$\begin{aligned}
1 &= \frac{1}{\gamma} \cdot \frac{1}{\sigma_{m,t}^2} \cdot \left(\frac{\sigma_{m,t}^2}{\sigma_{m,t}^2 + \psi} \cdot (\hat{\mu}_{m,t} - \lambda \cdot \psi) - \gamma \cdot \mathbb{Cov}_t[0' r_{t+1}, r_{m,t+1}] \right) \\
&\Downarrow \\
1 &= \frac{1}{\gamma} \cdot \left(\frac{1}{\sigma_{m,t}^2 + \psi} \cdot (\hat{\mu}_{m,t} - \lambda \cdot \psi) \right) \\
&\Downarrow \\
\gamma \cdot (\sigma_{m,t}^2 + \psi) &= \hat{\mu}_{m,t} - \lambda \cdot \psi \\
&\Downarrow \\
\hat{\mu}_{m,t} &= \gamma \cdot (\sigma_{m,t}^2 + \psi) + \lambda \cdot \psi,
\end{aligned} \tag{IA.17}$$

which is Equation 18 in the main text.

Similarly, substituting market clearing ($\omega_{m,t} = 1$ and $\omega_t = 0$) into the optimal anomaly allocation (Equation IA.16), we have

$$\begin{aligned}
0 &= (1/\gamma) \cdot \Sigma_t^{-1} (\Sigma_t (\Sigma_t + \psi \cdot \mathbf{I})^{-1} (\hat{\mu}_t - \lambda \cdot \psi \cdot 1) - \gamma \cdot \mathbb{Cov}_t[r_{t+1}, r_{m,t+1}]) \\
&\Downarrow \\
0 &= \Sigma_t (\Sigma_t + \psi \cdot \mathbf{I})^{-1} (\hat{\mu}_t - \lambda \cdot \psi \cdot 1) - \gamma \cdot \mathbb{Cov}_t[r_{t+1}, r_{m,t+1}] \\
&\Downarrow \\
\gamma \cdot \mathbb{Cov}_t[r_{t+1}, r_{m,t+1}] &= \Sigma_t (\Sigma_t + \psi \cdot \mathbf{I})^{-1} (\hat{\mu}_t - \lambda \cdot \psi \cdot 1) \\
&\Downarrow \\
\hat{\mu}_t &= (\mathbf{I} + \psi \cdot \Sigma_t^{-1}) \cdot \gamma \cdot \sigma_{m,t}^2 \cdot \beta_t + \lambda \cdot \psi \cdot 1,
\end{aligned} \tag{IA.18}$$

where $\beta_t = \mathbb{Cov}_t[r_{t+1}, r_{m,t+1}] / \sigma_{m,t}^2$. This is Equation 19 in the main text after recognizing that Σ_t is diagonal with $\Sigma_t[a, a] = \sigma_{a,t}^2$.

So, defining $\hat{\alpha}_{a,t} = \hat{\mu}_{a,t} - \beta_{a,t} \cdot \hat{\mu}_{m,t}$ yields

$$\begin{aligned}
\hat{\alpha}_{a,t} &= ((1 + \psi/\sigma_{a,t}^2) \cdot \gamma \cdot \sigma_{m,t}^2 \cdot \beta_{a,t} + \lambda \cdot \psi) - \beta_{a,t} \cdot (\gamma \cdot (\sigma_{m,t}^2 + \psi) + \lambda \cdot \psi) \\
&= ((\sigma_{m,t}^2 + \psi \cdot \sigma_{m,t}^2/\sigma_{a,t}^2) \cdot \gamma \cdot \beta_{a,t} + \lambda \cdot \psi) - \gamma \cdot \beta_{a,t} \cdot \sigma_{m,t}^2 - \psi \cdot \beta_{a,t} \cdot (\gamma + \lambda) \\
&= \psi \cdot (\lambda + \gamma \cdot (\sigma_{m,t}^2/\sigma_{a,t}^2)) \cdot \beta_{a,t} - (\gamma + \lambda) \cdot \beta_{a,t}, \tag{IA.19}
\end{aligned}$$

which is Equation 20 in the main text.

The vector of realized alphas from $t + 1$ to $t + H$ is

$$\hat{\alpha}_a = \hat{\mu}_a - \hat{\beta}_a \cdot \hat{\mu}_m \tag{IA.20}$$

where $\hat{\mu}_n = (1/H) \cdot \sum_{h=1}^H r_{n,t+h}$ and $\hat{\beta}_a = (\sum_{h=1}^H (r_{a,t+h} - \hat{\mu}_a) \cdot (r_{m,t+h} - \hat{\mu}_m)) / (\sum_{h=1}^H (r_{m,t+h} - \hat{\mu}_m)^2)$.

To derive Equation 22 in the main text ($\hat{\alpha}_a = \hat{\alpha}_{a,t} + \varepsilon_a$), start by noting that if $\hat{\theta}$ is the MLE for θ , then $f(\hat{\theta})$ is the MLE for $f(\theta)$. In our context, this implies $\hat{\alpha}_t = \hat{\mu}_t - \beta_{a,t} \cdot \hat{\mu}_{m,t}$ is the MLE for $\alpha_t = \mu_t - \beta_{a,t} \cdot \mu_{m,t}$, and thus we can write $\hat{\alpha}_{a,t} = \alpha_{a,t} + u_a$, where $\mathbb{E}_t[u_a] = 0$. Then, since the return distribution remains fixed after time t , we have $\hat{\alpha}_a = \alpha_{a,t} + \epsilon_a$, where $\mathbb{E}[\epsilon_a | \mathcal{F}_t] = 0$ with $\mathcal{F}_t$ reflecting the information filtration process that includes all public information (which investors know) as well as all conditional return moments (with investors not knowing the conditional means). Then, the law of iterated expectations implies $\mathbb{E}_t[\epsilon_a] = \mathbb{E}_t[\mathbb{E}[\epsilon_a | \mathcal{F}_t]] = 0$, so that $\hat{\alpha} = \hat{\alpha}_t + \varepsilon_a$, where $\varepsilon_a = \epsilon_a - u_a$ satisfies $\mathbb{E}_t[\varepsilon_a] = 0$.

D.3 Deriving Equilibrium Expected Returns and Alphas with Noise Traders

We assume there is a mass $1 - \theta$ of rational traders (rt) and θ of noise traders (nt), who are identical to the rational traders except that they mistakenly think $\hat{\mu}_t + \tilde{\mu}_t$ is the MLE for μ_t so that their demand has an extra noise component given by $\tilde{\mu}_t$. In this case, market clearing is given by $1 = (1 - \theta) \cdot \omega_{m,t}^{(rt)} + \theta \cdot \omega_{m,t}^{(nt)}$ and $0 = (1 - \theta) \cdot \omega_{a,t}^{(rt)} + \theta \cdot \omega_{a,t}^{(noise)}$ with

$$\omega_{m,t}^{(rt)} = \frac{1}{\gamma \cdot \sigma_{m,t}^2} \cdot \left(\frac{\sigma_{m,t}^2}{\sigma_{m,t}^2 + \psi} \cdot (\hat{\mu}_{m,t} - \lambda \cdot \psi) - \gamma \cdot \text{Cov}_t[\omega_t^{(rt)'} r_{t+1}, r_{m,t+1}] \right), \tag{IA.21}$$

$$\omega_{a,t}^{(rt)} = \frac{1}{\gamma \cdot \sigma_{a,t}^2} \cdot \left(\frac{\sigma_{a,t}^2}{\sigma_{a,t}^2 + \psi} \cdot (\hat{\mu}_{a,t} - \lambda \cdot \psi) - \gamma \cdot \text{Cov}_t[r_{a,t+1}, \omega_{m,t}^{(rt)} \cdot r_{m,t+1}] \right), \quad (\text{IA.22})$$

$$\omega_{m,t}^{(nt)} = \frac{1}{\gamma \cdot \sigma_{m,t}^2} \cdot \left(\frac{\sigma_{m,t}^2}{\sigma_{m,t}^2 + \psi} \cdot (\hat{\mu}_{m,t} + \tilde{\mu}_{m,t} - \lambda \cdot \psi) - \gamma \cdot \text{Cov}_t[\omega_t^{(nt)'} r_{t+1}, r_{m,t+1}] \right), \quad (\text{IA.23})$$

$$\omega_{a,t}^{(nt)} = \frac{1}{\gamma \cdot \sigma_{a,t}^2} \cdot \left(\frac{\sigma_{a,t}^2}{\sigma_{a,t}^2 + \psi} \cdot (\hat{\mu}_{a,t} + \tilde{\mu}_{a,t} - \lambda \cdot \psi) - \gamma \cdot \text{Cov}_t[r_{a,t+1}, \omega_{m,t}^{(nt)} \cdot r_{m,t+1}] \right). \quad (\text{IA.24})$$

So, for the market portfolio, we have

$$\begin{aligned} 1 &= (1 - \theta) \cdot \omega_{m,t}^{(rt)} + \theta \cdot \omega_{m,t}^{(nt)} \\ &= \frac{1}{\gamma \cdot \sigma_{m,t}^2} \cdot \left(\frac{\sigma_{m,t}^2}{\sigma_{m,t}^2 + \psi} \cdot (\hat{\mu}_{m,t} - \lambda \cdot \psi + \theta \cdot \tilde{\mu}_{m,t}) - \gamma \cdot \text{Cov}_t[((1 - \theta) \cdot \omega_t^{(rt)} + \theta \cdot \omega_t^{(nt)})' r_{t+1}, r_{m,t+1}] \right) \\ &= \frac{1}{\gamma \cdot (\sigma_{m,t}^2 + \psi)} \cdot (\hat{\mu}_{m,t} - \lambda \cdot \psi + \theta \cdot \tilde{\mu}_{m,t}) \end{aligned}$$

$\Downarrow$

$$\hat{\mu}_{m,t} = \gamma \cdot (\sigma_{m,t}^2 + \psi) + \lambda \cdot \psi - \theta \cdot \tilde{\mu}_{m,t}. \quad (\text{IA.25})$$

Similarly, for each anomaly, we have

$$\begin{aligned} 0 &= (1 - \theta) \cdot \omega_{a,t}^{(rt)} + \theta \cdot \omega_{a,t}^{(nt)} \\ &= \frac{1}{\gamma \cdot \sigma_{a,t}^2} \cdot \left(\frac{\sigma_{a,t}^2}{\sigma_{a,t}^2 + \psi} \cdot (\hat{\mu}_{a,t} - \lambda \cdot \psi + \theta \cdot \tilde{\mu}_{a,t}) - \gamma \cdot \text{Cov}_t[r_{a,t+1}, ((1 - \theta) \cdot \omega_{m,t}^{(rt)} + \theta \cdot \omega_{m,t}^{(nt)}) \cdot r_{m,t+1}] \right) \\ &= \frac{\sigma_{a,t}^2}{\sigma_{a,t}^2 + \psi} \cdot (\hat{\mu}_{a,t} - \lambda \cdot \psi + \theta \cdot \tilde{\mu}_{a,t}) - \gamma \cdot \text{Cov}_t[r_{a,t+1}, r_{m,t+1}] \end{aligned}$$

$\Downarrow$

$$\hat{\mu}_{a,t} = (1 + \psi/\sigma_{a,t}^2) \cdot \gamma \cdot \text{Cov}_t[r_{a,t+1}, r_{m,t+1}] + \lambda \cdot \psi - \theta \cdot \tilde{\mu}_{a,t}. \quad (\text{IA.26})$$

So, defining $\hat{\alpha}_{a,t} = \hat{\mu}_{a,t} - \beta_{a,t} \cdot \hat{\mu}_{m,t}$, we have

$$\hat{\alpha}_{a,t} = \psi \cdot (\lambda + \gamma \cdot (\sigma_{m,t}^2/\sigma_{a,t}^2) \cdot \beta_{a,t} - (\gamma + \lambda) \cdot \beta_{a,t}) - \theta \cdot \tilde{\alpha}_{a,t}, \quad (\text{IA.27})$$

which is Equation 24 in the main text with $\tilde{\alpha}_{a,t} = \tilde{\mu}_{a,t} - \beta_{a,t} \cdot \tilde{\mu}_{m,t}$.

E An Endowment Economy Version of the Model in Section 4

There is a representative investor who can allocate capital to the risk-free asset with known gross return R_f (and zero net supply), to the market portfolio with excess return $r_{m,t} = R_{m,t} - R_f$ (and unit net supply), and to a set of anomaly portfolios with long-short return vector $R_{L,t} - R_{S,t}$ (and zero net supply). The investor has constant relative risk aversion (CRRA) preferences in the form of power utility so that he forms his wealth portfolio (with gross return $R_{w,t}$) by solving

$$[\omega_{m,t}, \omega_t] = \underset{\omega}{\operatorname{argmax}} \quad \frac{1}{1-\gamma} \mathbb{E}_t[W_{t+1}^{1-\gamma}] \quad \text{s.t.} \quad W_{t+1} = W_t \cdot R_{w,t+1} \quad (\text{IA.28})$$

$$R_{w,t+1} = R_f + \omega_{m,t} \cdot (R_{m,t+1} - R_f) + \omega'_t (R_{L,t+1} - R_{S,t+1})$$

Substituting the two constraints, we have

$$[\omega_{m,t}, \omega_t] = \underset{\omega}{\operatorname{argmax}} \quad \frac{1}{1-\gamma} \mathbb{E}_t \left[(R_{f,t+1} + \omega_{m,t} \cdot (R_{m,t+1} - R_f) + \omega'_t (R_{L,t+1} - R_{S,t+1}))^{1-\gamma} \right] \quad (\text{IA.29})$$

which yields the first order conditions

$$\mathbb{E}_t \left[e^{-\gamma \cdot \log(R_{w,t+1})} \cdot (R_{m,t+1} - R_f) \right] = 0 \quad (\text{IA.30})$$

$$\mathbb{E}_t \left[e^{-\gamma \cdot \log(R_{w,t+1})} \cdot (R_{aL,t+1} - R_{aS,t+1}) \right] = 0 \quad (\text{IA.31})$$

with a reflecting an anomaly index (and aL and aS the respective long and short portfolios).

Assuming log returns are normally distributed (which is a valid assumption given the endowment process described below), we can rewrite Equations [IA.30](#) and [IA.31](#) as

$$\mathbb{E}_t[r_{m,t+1}] + 0.5 \cdot \mathbb{V}ar_t[r_{m,t+1}] = \gamma \cdot \mathbb{C}ov_t[r_{m,t+1}, r_{w,t+1}] \quad (\text{IA.32})$$

$$\mathbb{E}_t[r_{a,t+1}] + 0.5 \cdot (\mathbb{V}ar_t[r_{aL,t+1}] - \mathbb{V}ar_t[r_{aS,t+1}]) = \gamma \cdot \mathbb{C}ov_t[r_{a,t+1}, r_{w,t+1}] \quad (\text{IA.33})$$

where $r_{n,t} = \log(R_{n,t}) - \log(R_f)$ (for $n = m, aL, aS$) and $r_{a,t} = r_{aL,t} - r_{aS,t}$, which represent a slight abuse of notation relative to the main text as r terms here reflect excess log returns (instead of excess returns as in the main text).

Finally, using the portfolio log return approximation from Campbell, Chan, and Viceira (2003) (which is exact in continuous time),

$$r_{w,t+1} = \text{const}_t + \omega_{m,t} \cdot r_{m,t+1} + \omega'_t r_{t+1}, \quad (\text{IA.34})$$

and imposing market clearing ($\omega_{m,t} = 1$ and $\omega_t = 0$), we can write Equations IA.32 and IA.33 as

$$\mathbb{E}_t[r_{m,t+1}] + 0.5 \cdot \text{Var}_t[r_{m,t+1}] = \gamma \cdot \text{Var}_t[r_{m,t+1}] \quad (\text{IA.35})$$

$$\mathbb{E}_t[r_{a,t+1}] + 0.5 \cdot (\text{Var}_t[r_{aL,t+1}] - \text{Var}_t[r_{aS,t+1}]) = \gamma \cdot \text{Cov}_t[r_{a,t+1}, r_{m,t+1}] \quad (\text{IA.36})$$

Now, we describe the endowment process for each asset $n = m, aL, aS$ and how investors form beliefs. The economy has two periods (t and $t + 1$). At time t , the investor observes the risk-free rate ($r_f = \log(R_f)$) as well as the current dividend ($D_{n,t}$) and equilibrium price ($P_{n,t}$) for each asset. At time $t + 1$, the next dividend realizes ($D_{n,t+1}$), which yields $R_{n,t+1} = D_{n,t+1}/P_{n,t}$. So, $r_{n,t+1} = \Delta d_{n,t+1} + dp_{n,t} - r_f$, which implies $\mathbb{E}_t[r_{n,t+1}] = \mathbb{E}_t[\Delta d_{n,t+1}] + dp_{n,t} - r_f$ and $\text{Var}_t[r_{n,t+1}] = \text{Var}_t[\Delta d_{n,t+1}]$. We also define $\Delta d_{a,t+1} = \Delta d_{aL,t+1} - \Delta d_{aS,t+1}$ so that $\text{Var}_t[r_{a,t+1}] = \text{Var}_t[\Delta d_{a,t+1}]$.

The endowment growth of each asset n is normally distributed. The investor observes the endowment growth variances ($\text{Var}_t[\Delta d_{n,t+1}] = \sigma_{n,t}^2$) and covariances ($\text{Cov}_t[\Delta d_{n1,t+1}, \Delta d_{n2,t+1}]$), which implicitly give the investor the variances and covariances of log returns since $r_{n,t+1} - \mathbb{E}_t[r_{n,t+1}] = \Delta d_{n,t+1} - \mathbb{E}_t[\Delta d_{n,t+1}]$. Moreover, analogous to the main text, $\Delta d_{a,t} = \Delta d_{aL,t} - \Delta d_{aS,t}$ are uncorrelated across anomalies so that $r_{a,t} = r_{aL,t} - r_{aS,t}$ are uncorrelated across anomalies (without loss of generality).

However, in line with the arguments in Merton (1980), the investor faces uncertainty about the conditional mean growth terms: $g_{m,t}$ and $g_{a,t} = g_{aL,t} - g_{aS,t}$. Specifically, the investor only observes the unbiased but noisy signal $\hat{g}_{n,t} \sim \mathcal{N}(g_{n,t}, \hat{\nu} \cdot \sigma_{n,t}^2)$ for $n = m, a$. Given this endowment growth information, the investor has the conditional mean return signals $\hat{\mu}_{m,t} = \hat{g}_{m,t} + dp_{m,t} - r_f + 0.5 \cdot \sigma_{m,t}^2$ and $\hat{\mu}_{a,t} = \hat{g}_{a,t} + dp_{aL,t} - dp_{aS,t} + 0.5 \cdot (\sigma_{aL,t}^2 - \sigma_{aS,t}^2)$. Moreover, given the $\hat{g}_{n,t}$ distribution, the conditional mean return signal has distribution

$\hat{\mu}_{n,t} \sim \mathcal{N}(\mu_{n,t}, \hat{\nu} \cdot \sigma_{n,t}^2)$, where the unobservable conditional mean returns are given by $\mu_{m,t} = g_{m,t} + dp_{m,t} - r_f + 0.5 \cdot \sigma_{m,t}^2$ and $\mu_{a,t} = g_{a,t} + dp_{aL,t} - dp_{aS,t} + 0.5 \cdot (\sigma_{aL,t}^2 - \sigma_{aS,t}^2)$.

Finally, the investor forms $\mathbb{E}_t[r_{n,t+1}] + 0.5 \cdot \sigma_{n,t}^2 = \mathbb{E}_t[\mu_{n,t}]$ by combining the $\hat{\mu}_{n,t}$ distribution with his prior belief that $\mu_{n,t} \sim \mathcal{N}(-\lambda \cdot \sigma_{n,t}^2, \nu \cdot \sigma_{n,t}^4)$, which yields

$$\mathbb{E}_t[\mu_{n,t}] = \frac{\sigma_{n,t}^2}{\sigma_{n,t}^2 + \psi} \cdot (\hat{\mu}_{n,t} - \lambda \cdot \psi), \quad (\text{IA.37})$$

where $\psi = \hat{\nu}/\nu$.

Substituting Equation [IA.37](#) into Equations [IA.35](#) and [IA.36](#) yields

$$\hat{\mu}_{m,t} = \gamma \cdot (\sigma_{m,t}^2 + \psi) + \lambda \cdot \psi \quad (\text{IA.38})$$

$$\hat{\mu}_{a,t} = (1 + \psi/\sigma_{a,t}^2) \cdot \gamma \cdot \sigma_{m,t}^2 \cdot \beta_{a,t} + \lambda \cdot \psi, \quad (\text{IA.39})$$

where $\beta_{a,t} = \text{Cov}_t[r_{a,t+1}, r_{m,t+1}]/\text{Var}_t[r_{m,t+1}] = \text{Cov}_t[\Delta d_{a,t+1}, \Delta d_{m,t+1}]/\sigma_{m,t}^2$.

The equilibrium expected return signals (Equations [IA.38](#) and [IA.39](#)) are identical to the equilibrium expected return MLEs in the main text (Equations [18](#) and [19](#)). The only difference is in the definition of r and μ . In the main text, $r_{n,t}$ is an excess return and $\mu_{n,t}$ is the FIRE expectation of $r_{n,t+1}$. In this section, $r_{n,t}$ is an excess log return and $\mu_{n,t}$ is FIRE expectation of $r_{n,t+1}$ plus an adjustment for the Jensen's inequality term that arises in this log-Normal version of the model.

Importantly, however, this endowment version of the model clarifies how expected returns adjust in equilibrium. Specifically, we have that all terms in each $\hat{\mu}$ are exogenous to the investor's decisions, except for the dp terms. So, equilibrium expected returns adjust through the equilibrium dp . Intuitively, the price of the asset adjusts to its risk exposure, which endogenously affects the expected return the investor infers out of the exogenous endowment growth signal $\hat{g}$. So, even though the $\hat{g}$ signal is exogenous in the model, the $\hat{\mu}$ signal is endogenous (determined in equilibrium).

References for Internet Appendix

- Bryzgalova, S., M. Pelger, and J. Zhu (2024). “Forest Through the Trees: Building Cross-Sections of Stock Returns”. In: *Journal of Finance* Forthcoming.
- Campbell, J. Y., Y. L. Chan, and L. M. Viceira (2003). “A multivariate model of strategic asset allocation”. In: *Journal of Financial Economics* 67.1, pp. 41–80.
- Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997). *The Econometrics of Financial Markets*. Princeton University Press.
- Cochrane, J. H. (2005). *Asset Pricing*. Revised Edition. Princeton University Press.
- Davis, C. (2025). “The Elasticity of Quantitative Investment”. In: *Review of Financial Studies* 38.10, pp. 2845–2886.
- DeMiguel, V., L. Garlappi, and R. Uppal (2009). “Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?” In: *Review of Financial Studies* 22.5, pp. 1915–1953.
- Gibbons, M. R., S. A. Ross, and J. Shanken (1989). “A Test of the Efficiency of a Given Portfolio”. In: *Econometrica* 57.5, pp. 1121–1152.
- Gu, S., B. Kelly, and D. Xiu (2020). “Empirical Asset Pricing via Machine Learning”. In: *Review of Financial Studies* 33.5, pp. 2223–2273.
- Kozak, S., S. Nagel, and S. Santosh (2020). “Shrinking the Cross-Section”. In: *Journal of Financial Economics* 135.2, pp. 271–292.
- Merton, R. C. (1980). “On Estimating the Expected Return on the Market: An Exploratory Investigation”. In: *Journal of Financial Economics* 8, pp. 323–361.
- Newey, W. K. and K. D. West (1987). “A Simple, Positive-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix”. In: *Econometrica* 55.3, pp. 703–708.
- Newey, W. K. and K. D. West (1994). “Automatic Lag Selection in Covariance Matrix Estimation”. In: *Review of Economic Studies* 61.4, pp. 631–653.
- Tibshirani, R. (1996). “Regression Shrinkage and Selection via the Lasso”. In: *Journal of the Royal Statistical Society: Series B* 58.1, pp. 267–288.

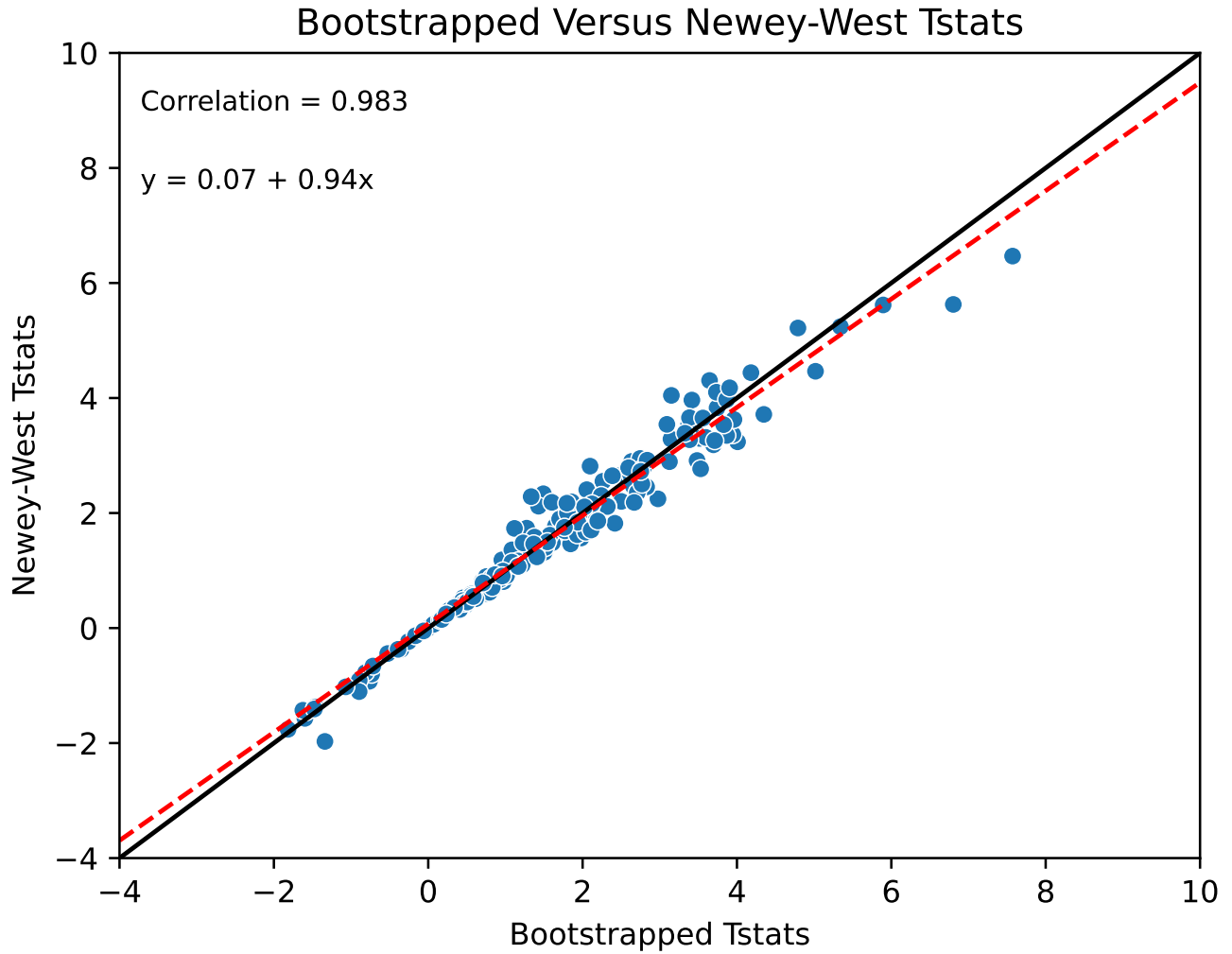


Figure IA.1

Comparing Bootstrap and Newey-West t-stats of Factor Regression Alphas Post-Publication

The figure plots t-stats of factor regression alphas post-publication based on Newey and West (1987, 1994) (in the y-axis) and the bootstrap method described in Section B of this Internet Appendix (in the x-axis). In the main text, we focus on bootstrap t-statistics since we cannot compute t-statistics for real time investor alphas using Newey and West (1987, 1994). However, the graph shows that the two procedures produce similar t-statistics for factor regression alphas over the post-publication period. These results are discussed in Subsection 2.2.

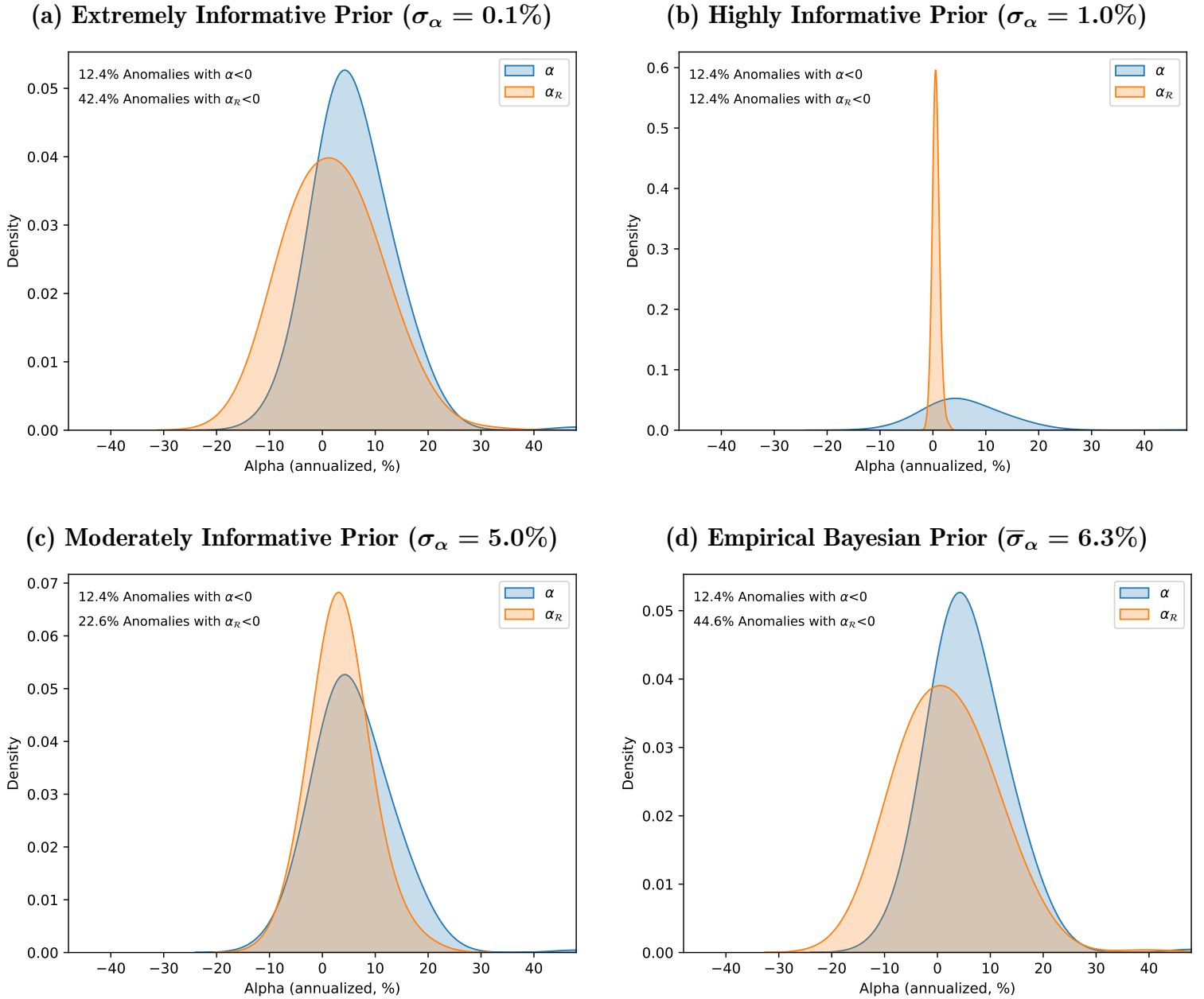


Figure IA.2
Distribution of Alphas: Bayesian Weight Estimation

This figure contains density plots for the distributions of CAPM alphas across 177 anomaly strategies. We form each anomaly strategy as the long-short decile portfolio with weights as per the original publication. We consider factor regression alphas (α) and real time investor alphas (α_R). These alphas differ based on the weights used to combine the anomaly strategy with the market portfolio. Factor regression alphas (real time investor alphas) use weights calculated from the post-publication (pre-publication) period to build a portfolio over the post-publication period, with the pre-publication period ending in December of the publication year. Consequently, factor regression alphas are equivalent to intercepts from factor regressions over the post-publication period, but real time investor alphas are not. The optimal weights for the real time investor alphas are constructed using Bayesian shrinkage with the figure panels varying by the level of prior informativeness. Panels A to C consider different fixed priors while Panel D considers an empirical Bayesian prior, with the details described in Subsection C.1. Results are based on monthly returns and annualized (approximately) by multiplying by 12. Section 1 presents the alpha definitions, Section 2 provides more details about the estimation of alphas, and Subsection C.1 adjusts the estimation for this Bayesian application and discusses the results from this figure.

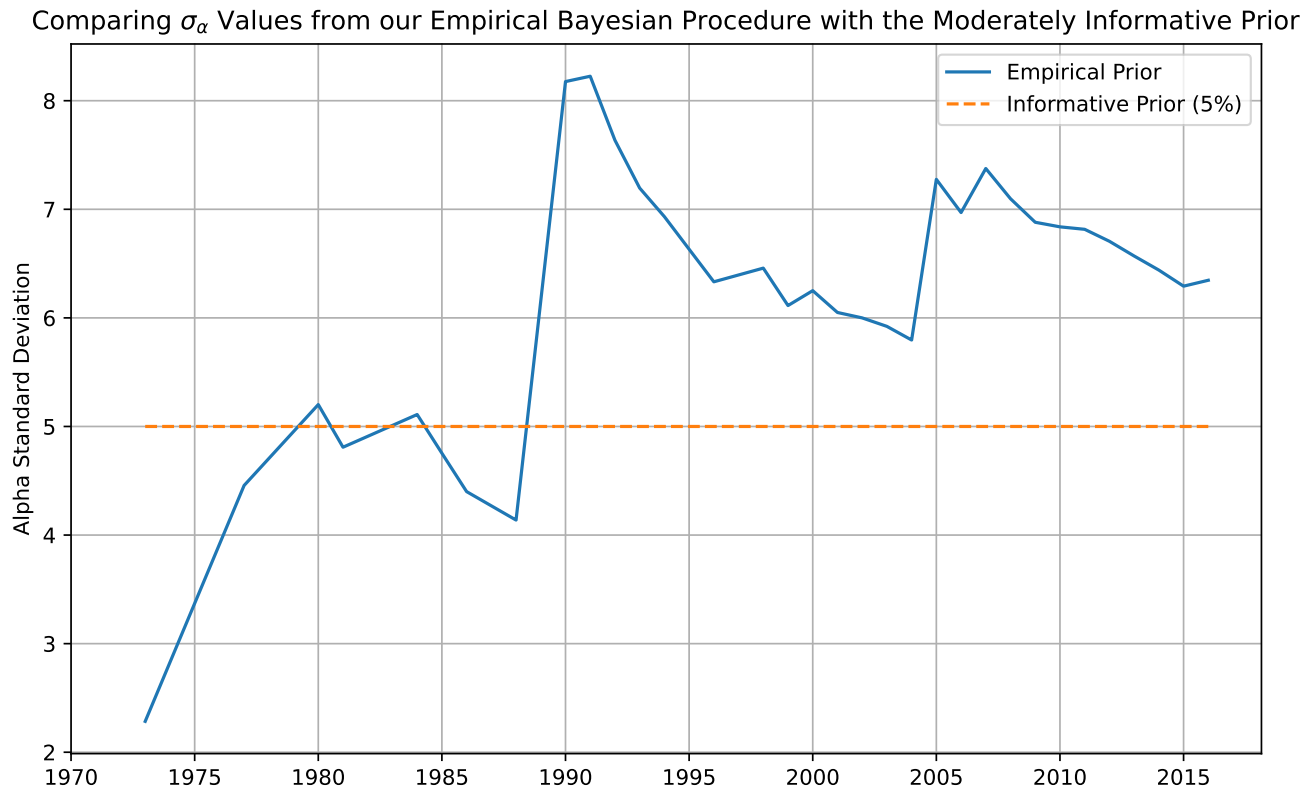
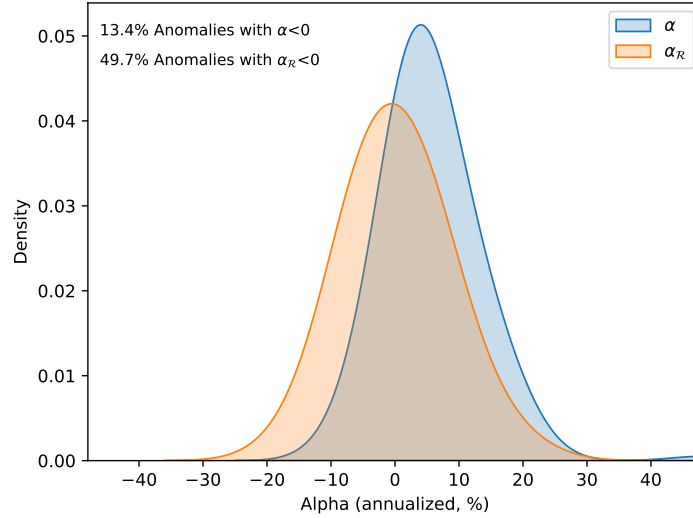


Figure IA.3

Comparing σ_α Values from our Empirical Bayesian Procedure with the Moderately Informative Prior

The figure plots the $\sigma_\alpha = 5.0\%$ of our moderately informative prior (in orange) against the time-series of the σ_α estimated from our empirical Bayesian procedure (in blue). The details of our Bayesian procedure and a discussion of the results in this graph are provided in Subsection [C.1](#).

(a) Estimating Weights using KNS (Kozak, Nagel, and Santosh (2020))



(b) Estimating Weights using BPZ (Bryzgalova, Pelger, and Zhu (2024))

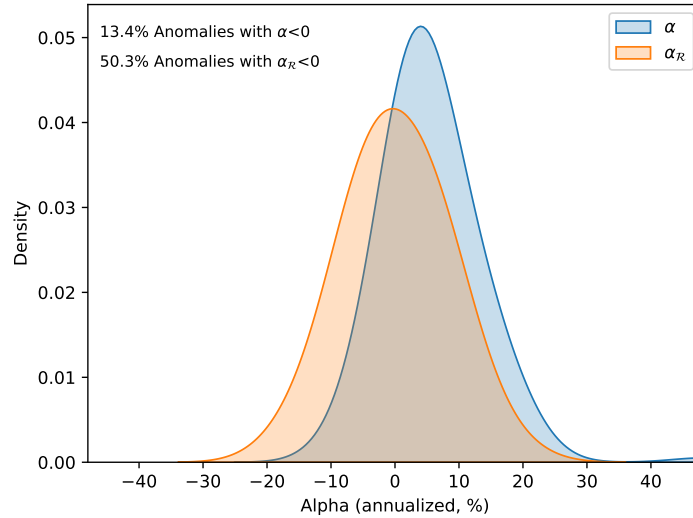


Figure IA.4
Distribution of Alphas: Machine Learning Weight Estimation

This figure contains density plots for the distributions of CAPM alphas across 149 anomaly strategies. These 149 anomaly strategies reflect the subset of our original 177 anomaly strategies that have return data for at least 20 years in the pre-publication period. We form each anomaly strategy as the long-short decile portfolio with weights as per the original publication. We consider factor regression alphas (α) and real time investor alphas (α_R). These alphas differ based on the weights used to combine the anomaly strategy with the market portfolio. Factor regression alphas (real time investor alphas) use weights calculated from the post-publication (pre-publication) period to build a portfolio over the post-publication period, with the pre-publication period ending in December of the publication year. Consequently, factor regression alphas are equivalent to intercepts from factor regressions over the post-publication period, but real time investor alphas are not. The optimal weights for the real time investor alphas are constructed using Machine Learning methods, with the details described in Subsection C.1. Following Davis (2025), we only explore two machine learning methods to compute portfolio weights: the KNS method from Kozak, Nagel, and Santosh (2020) (in Panel A) and the BPZ method from Bryzgalova, Pelger, and Zhu (2024) (in Panel B). Results are based on monthly returns and annualized (approximately) by multiplying by 12. Section 1 presents the alpha definitions, Section 2 provides more details about the estimation of alphas, and Subsection C.2 adjusts the estimation for this Machine Learning application and discusses the results from this figure.

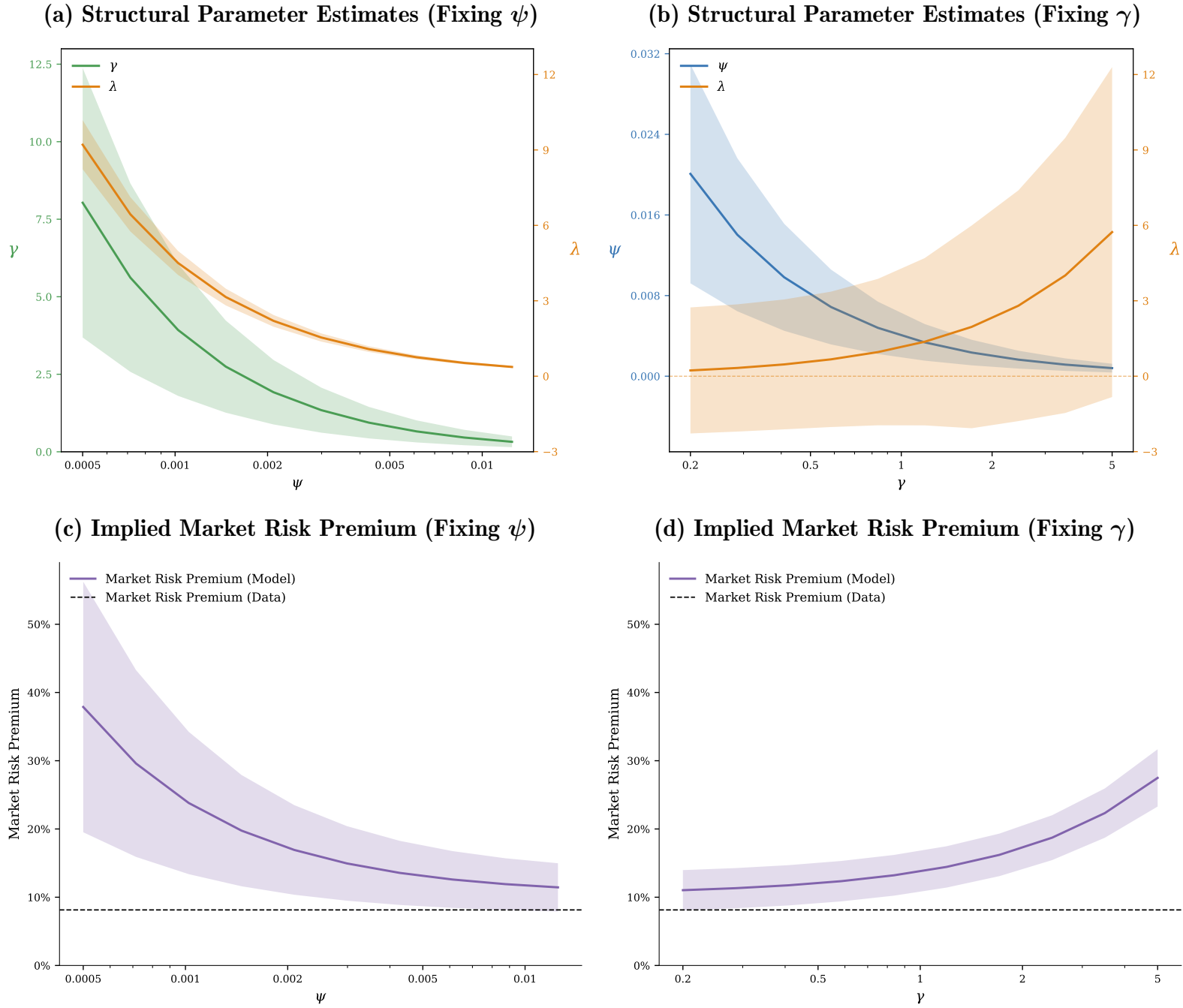


Figure IA.5
Model Estimation: Structural Parameters and Implied Market Risk Premium
(using pre- and post-publication data from 177 Anomalies)

This figure reports structural parameter estimates and the market risk premium implied by the model in Subsection 4.1. The results are based on the structural estimation in the last four columns of the first two rows of Table 3, which uses the 177 anomalies and differentiates between pre-publication and post-publication returns when calculating the moments needed for estimation. Because Equation 23 identifies the products $A = \psi \cdot \lambda$ and $C = \psi \cdot \gamma$, but does not separately identify ψ , γ , and λ , the figure fixes one structural parameter and recovers the other two. Panel (a) fixes ψ over the interval $[\bar{\psi}/5, \bar{\psi} \cdot 5]$ where $\bar{\psi} = 0.0025$, and reports the corresponding estimates of γ and λ . Panel (b) fixes γ over the interval $[\bar{\gamma}/5, \bar{\gamma} \cdot 5]$ where $\bar{\gamma} = 1$, and reports the corresponding estimates of ψ and λ . Panels (c) and (d) report the market risk premium implied by Equation 18 using the structural parameters from Panels (a) and (b), respectively. The solid lines report point estimates, while the shaded areas report 95% confidence intervals. Implied market risk premia are annualized approximately by multiplying monthly values by 12. Subsection 4.2 provides more details about the estimation.

Table IA.1
Distribution of Alphas: Bayesian Weight Estimation

This table reports the results for the distributions of CAPM alphas and their t-statistics across the 177 anomaly strategies we study, with each anomaly strategy reflecting a long-short decile portfolio with weights as per the original anomaly publication. We consider real time investor alphas, which use weights calculated from the pre-publication period to build a portfolio over the post-publication period, with the pre-publication period ending in December of the publication year. Consequently, real time investor alphas are not equivalent to intercepts from factor regressions over the post-publication period. The optimal weights for the real time investor alphas are constructed using Bayesian shrinkage with the table panels varying by the level of prior informativeness. Panels A to C consider different fixed priors while Panel D considers an empirical Bayesian prior, with the details described in Subsection C.1. Results are based on monthly returns and annualized (approximately) by multiplying by 12. Section 1 presents the alpha definitions, Section 2 provides more details about the estimation of alphas, and Subsection C.1 adjusts the estimation for this Bayesian application and discusses the results from this table.

PANEL A - Extremely Informative Prior ($\sigma_\alpha = 0.1\%$)

	Mean	Median	Min	Q10%	Q25%	Q75%	Q90%	Max
(Bayesian) Optimal Weight	0.01	0.01	0.00	0.00	0.00	0.02	0.04	0.11
Real Time Investor Alpha ($\alpha_{\mathcal{R}}$)	0.6%	0.5%	-1.0%	-0.1%	0.3%	0.9%	1.3%	2.9%
$\alpha_{\mathcal{R}}$ t-stat	(0.67)	(0.80)	(-1.05)	(-0.39)	(0.47)	(1.04)	(1.26)	(2.12)

PANEL B - Highly Informative Prior ($\sigma_\alpha = 1.0\%$)

	Mean	Median	Min	Q10%	Q25%	Q75%	Q90%	Max
(Bayesian) Optimal Weight	0.33	0.29	0.00	0.05	0.12	0.51	0.68	0.89
Real Time Investor Alpha ($\alpha_{\mathcal{R}}$)	3.5%	3.1%	-9.7%	-2.5%	0.7%	6.3%	9.7%	21.1%
$\alpha_{\mathcal{R}}$ t-stat	(0.82)	(0.97)	(-2.25)	(-0.86)	(0.23)	(1.48)	(2.09)	(8.05)

PANEL C - Moderately Informative Prior ($\sigma_\alpha = 5.0\%$)

	Mean	Median	Min	Q10%	Q25%	Q75%	Q90%	Max
(Bayesian) Optimal Weight	0.58	0.59	0.02	0.28	0.46	0.75	0.84	1.05
Real Time Investor Alpha ($\alpha_{\mathcal{R}}$)	2.0%	2.3%	-14.6%	-7.0%	-4.0%	7.4%	12.7%	29.8%
$\alpha_{\mathcal{R}}$ t-stat	(0.29)	(0.54)	(-2.57)	(-1.47)	(-0.89)	(1.14)	(1.97)	(7.01)

PANEL D - Empirical Bayesian Prior ($\bar{\sigma}_\alpha = 6.3\%$)

	Mean	Median	Min	Q10%	Q25%	Q75%	Q90%	Max
(Bayesian) Optimal Weight	0.59	0.61	0.03	0.33	0.48	0.75	0.84	1.06
Real Time Investor Alpha ($\alpha_{\mathcal{R}}$)	1.8%	1.9%	-15.2%	-7.7%	-4.1%	7.0%	12.6%	39.8%
$\alpha_{\mathcal{R}}$ t-stat	(0.25)	(0.48)	(-2.57)	(-1.48)	(-0.96)	(1.10)	(1.86)	(7.00)

Table IA.2
Distribution of Alphas: Machine Learning Weight Estimation

This table reports the results for the distributions of CAPM alphas and their t-statistics across 149 of the anomaly strategies we study, with each anomaly strategy reflecting a long-short decile portfolio with weights as per the original anomaly publication. These 149 anomaly strategies reflect the subset of our original 177 anomaly strategies that have return data for at least 20 years in the pre-publication period. We consider real time investor alphas, which use weights calculated from the pre-publication period to build a portfolio over the post-publication period, with the pre-publication period ending in December of the publication year. Consequently, real time investor alphas are not equivalent to intercepts from factor regressions over the post-publication period. The optimal weights for the real time investor alphas are constructed using Machine Learning methods, with the details described in Subsection C.1. Following Davis (2025), we only explore two machine learning methods to compute portfolio weights: the KNS method from Kozak, Nagel, and Santosh (2020) (in Panel B) and the BPZ method from Bryzgalova, Pelger, and Zhu (2024) (in Panel C). For comparison, Panel A provides results using our baseline weight estimation applied to the same 149 anomaly strategies. Results are based on monthly returns and annualized (approximately) by multiplying by 12. Section 1 presents the alpha definitions, Section 2 provides more details about the estimation of alphas, and Subsection C.2 adjusts the estimation for this Machine Learning application and discusses the results from this table.

PANEL A - Estimating Weights using our Baseline Method

	Mean	Median	Min	Q10%	Q25%	Q75%	Q90%	Max
Optimal Weight	0.61	0.63	0.15	0.39	0.52	0.74	0.82	0.98
Real Time Investor Alpha ($\alpha_{\mathcal{R}}$)	0.8%	0.8%	-16.0%	-8.7%	-5.9%	6.2%	11.6%	25.3%
$\alpha_{\mathcal{R}}$ t-stat	(0.01)	(0.15)	(-2.57)	(-1.60)	(-1.05)	(0.91)	(1.58)	(4.22)

PANEL B - Estimating Weights using the KNS method (Kozak, Nagel, and Santosh (2020))

	Mean	Median	Min	Q10%	Q25%	Q75%	Q90%	Max
(ML) Optimal Weight	0.60	0.62	0.00	0.29	0.50	0.74	0.83	1.00
Real Time Investor Alpha ($\alpha_{\mathcal{R}}$)	0.2%	0.0%	-19.2%	-8.7%	-5.4%	5.6%	9.3%	23.2%
$\alpha_{\mathcal{R}}$ t-stat	(-0.02)	(0.00)	(-2.41)	(-1.51)	(-0.95)	(0.86)	(1.27)	(4.47)

PANEL C - Estimating Weights using the BPZ method (Bryzgalova, Pelger, and Zhu (2024))

	Mean	Median	Min	Q10%	Q25%	Q75%	Q90%	Max
(ML) Optimal Weight	0.61	0.62	0.00	0.38	0.52	0.76	0.84	1.00
Real Time Investor Alpha ($\alpha_{\mathcal{R}}$)	0.2%	0.0%	-17.3%	-9.4%	-5.2%	6.2%	9.3%	19.5%
$\alpha_{\mathcal{R}}$ t-stat	(-0.02)	(0.00)	(-2.39)	(-1.49)	(-1.01)	(0.86)	(1.26)	(4.65)