The Performance of Characteristic-Sorted Portfolios: Evaluating the Past and Predicting the Future

AYDOĞAN ALTI, TRAVIS L. JOHNSON, and SHERIDAN TITMAN*

August 2026

Journal of Financial and Quantitative Analysis, Forthcoming

Abstract

We study the performance of characteristic-sorted portfolios through the lens of a statistical model that allows for persistent variation in expected returns. Allowing for the possibility of time-varying expected returns substantially weakens the evidence for long-run outperformance compared to standard approaches. Using maximum likelihood tests, we find that the value, investment, and profitability portfolios' mean returns have p-values above 9%, and only 14 of a broader set of 174 portfolios have p-values below 1%. We also use Bayesian analyses to show that with agnostic priors, posterior estimates exhibit large fluctuations in expected returns that strongly predict realized returns out-of-sample.

*All three authors are with the Department of Finance, the University of Texas at Austin. Contact information: aydogan.alti@mccombs.utexas.edu, travis.johnson@mccombs.utexas.edu, and sheridan.titman@mccombs.utexas.edu. We thank Svetlana Bryzgalova (discussant), Scott Cederburg, Eugene Fama, Kenneth French, Amit Goyal (discussant), Xindi He (discussant), Narasimhan Jegadeesh, Lubos Pastor, Allan Timmermann, and seminar participants at the American Finance Association (AFA) Meetings, Asian Bureau of Finance and Economic Research (ABFER) Conference, Financial Intermediation Research Society (FIRS) Conference, UGA Fall Finance Conference, Baylor University, Chinese University of Hong Kong, Hong Kong University, HKUST, Peking University, Thammasat University, and the University of Texas at Austin for helpful comments, and the Q Group for awarding us the 2022 Jack Treynor prize. We used artificial intelligence tools, primarily Claude, to find faster algorithms for conducting our analysis, reformat tables and figures, and suggest copy edits.

A large and growing literature links firm characteristics, such as valuation ratios, to expected stock returns. While the evidence documented in this literature convincingly rejects the CAPM, going beyond this rejection and interpreting alternatives has proven to be challenging. In particular, it remains an open question whether the historical links between characteristics and returns represent permanent economic forces that will continue to shape returns in the future, or transitory forces that will gradually dissipate.

A case in point is the value premium. The historical tendency of value stocks to outperform growth stocks, shown by Fama and French (1992) and others, has weakened in the most recent period. Yet as Fama and French (2021) point out, we cannot confidently conclude that the value premium post-1991 is different than the value premium pre-1991. Although the value premium is quite high in the early period, and is not statistically significant in the latter period, the returns are sufficiently noisy that we cannot reject the hypothesis that they are drawn from identical distributions.[1]

This paper examines the returns of characteristic-sorted portfolios through the lens of a statistical model that allows expected returns to fluctuate over time. Our interpretation is that these fluctuations are generated by persistent shocks that affect the structure of the economy, investor beliefs, and capital allocated to arbitrage activity.[2] As we show, accounting for persistent fluctuations in expected returns influences the interpretation of these portfolios historical performance as well as the forecasts of their future performance.

Our analysis provides two main contributions. First, we demonstrate how persistent conditional expected return fluctuations derail standard methods for inferences about un-

[1] Another example of long-term fluctuations in characteristic-sorted portfolio returns is the strengthening profitability effect, documented by Novy-Marx (2013).

[2] We provide further motivation for the theoretical underpinnings of our statistical model in Section 1.2.

conditional, or long-run, performance. Specifically, we derive a simple closed-form formula that adjusts the standard errors of unconditional expected return estimates to account for persistent variation in conditional expected returns. Intuitively, a return process with higher and more slowly decaying autocorrelations effectively contains fewer independent observations, resulting in less precise expected return estimates. As we show, this adjustment can substantially increase standard errors; for example, standard errors nearly double, relative to the i.i.d. case, with a one-quarter return autocorrelation of 5% that decays with a half-life of five years. Moreover, our model-based adjustment reveals, and simulations confirm, that the commonly-used Newey and West (1987) procedure adjusts standard errors too little when returns are generated by our model, leading to over-rejection of null hypotheses.

We further explore these statistical inference challenges by estimating our full model, including the persistence parameters, using maximum likelihood. We find that many unconditional expected return estimates that appear statistically significant under the traditional OLS approach may be false positives. Among a broad set of 174 characteristic-sorted portfolios from Chen and Zimmermann (2021), only 14 attain maximum-likelihood p-values below 1.0% compared to 81 with OLS. In addition, we show that traditional tests can be prone to false negatives, i.e., failing to reject false hypotheses. For instance, despite its near-zero average performance over our sample period, the size portfolio exhibits strong return autocorrelation that decays with a half-life of one to two years. As a result, our maximum likelihood estimations for the size portfolio strongly reject the CAPM null hypothesis that conditional market-adjusted returns are always zero.

The relatively imprecise maximum likelihood estimates of the model parameters motivate the second half of the paper, which presents a Bayesian analysis that incorporates information

from assumed prior beliefs and observed data. We find that, with relatively agnostic priors, Bayesian investors infer large and persistent fluctuations in conditional Sharpe ratios. For example, the value portfolio's conditional Sharpe ratio at the end of our sample period is estimated to be near zero due to recent poor performance, while the profitability portfolio's conditional Sharpe ratio is much higher than it was earlier in the sample period due to stronger recent performance. These fluctuations in conditional Sharpe ratios are substantial, often ranging between zero and 0.8 on an annualized basis, with the size portfolio exhibiting even larger swings. Furthermore, out-of-sample forecasts from the Bayesian analysis reliably predict future realized returns and subsume the information from average past returns in the broad set of 174 portfolios.

While we believe that our analysis of characteristic-sorted portfolios is novel, it builds on a well-established literature that analyzes time-variation in the expected returns of the aggregate market portfolio. For example, Ferson, Sarkissian, and Simin (2003) consider the predictability of aggregate market returns using predictor variables such as price/dividend ratios. They present simulations that show that OLS regressions can overstate the significance of such relationships in finite samples when expected returns are persistent, even when the Newey and West (1987) standard error adjustment is used. Although our application is different, our framework is similar and our standard error correction and maximum likelihood estimation could also be applied in their setting.

Pástor and Stambaugh (2009, 2012) and Avramov, Cederburg, and Lučivjanská (2018) also conduct Bayesian analyses of time-varying expected returns and, like us, find that the priors about the return generating process substantially affect the posteriors about expected returns. Our study differs from these papers in both application and focus. Specifically, our

analysis concerns characteristic-sorted portfolios rather than the aggregate market portfolio, and we put greater emphasis on prior beliefs about persistence.[3]

Among studies that apply Bayesian methods to characteristic-sorted portfolios, Smith and Timmermann (2022) is the most closely related to ours as they too focus on time-series variation in expected portfolio returns.[4] In their statistical model, occasional structural breaks result in return premia of all characteristic-sorted portfolios, as well as the market factor, change at the same time and then remain constant until the next structural break occurs. Our approach complements Smith and Timmermann (2022) by examining continuous variations in return premia that are specific to an individual characteristic-sorted portfolio. Continuous variations also permit both standard error corrections for in-sample inference and out-of-sample forecasting.

Finally, our analysis is related to the growing literature on factor momentum (see, for instance, Lewellen (2002), McLean and Pontiff (2016), Avramov et al. (2017), Gupta and Kelly (2019), Arnott et al. (2021a), and Ehsani and Linnainmaa (2022)). Studies in this literature examine dynamic trading strategies involving several characteristic-sorted portfolios that are designed to exploit relatively short-lived variations in these portfolios' conditional expected returns. In contrast, we study longer-lived fluctuations in portfolio performance by estimating a model of persistent variation in returns.

[3]Other papers in the literature that employ Bayesian methods to study aggregate market returns include Kandel and Stambaugh (1996), Barberis (2000), Wachter and Warusawitharana (2009), and Johannes, Korteweg, and Polson (2014).

[4]Bayesian methods have also been employed to address related cross-sectional asset pricing questions. Pástor (2000), Avramov (2002), and Avramov (2004) use Bayesian approaches to portfolio choice. Kozak, Nagel, and Santosh (2020) use Bayesian shrinkage to estimate a stochastic discount factor, and Bryzgalova, Huang, and Julliard (2023) develop a Bayesian framework for selecting among and aggregating linear factor models. These papers are concerned with how much confidence to place in a given pricing model or set of predictors, whereas our focus is on the dynamics of conditional expected returns.

The remainder of the paper is organized as follows. Section 1 describes our statistical model of the return-generating process. Section 2 documents the historical return performance and reduced-form autocorrelation tests for the four main characteristic-sorted portfolios we study. Section 3 presents the frequentist estimations of the model. Section 4 presents the Bayesian analysis. Section 5 concludes.

1. Statistical Model

Our empirical analyses apply a statistical model with persistent fluctuations in conditional expected returns. In this section, we present the model specification, discuss its theoretical interpretations, and describe the return patterns it generates.

1.1. Model Specification

We assume that a time series of benchmark-adjusted portfolio returns r_t satisfies:

$$r_{t+1} = \mu_t + \epsilon_{t+1}, \tag{1}$$

$$\mu_{t+1} = \mu + \lambda(\mu_t - \mu + \delta_{t+1}), \tag{2}$$

where μ_t and μ are the conditional and unconditional expected returns, respectively. The shocks to unexpected returns ϵ_t and the shocks to expected returns δ_t are i.i.d. and follow a joint normal distribution with standard deviations σ_ϵ and σ_δ, respectively, and correlation $\rho \in (-1, 1)$.[5] We expect ρ to be negative, since positive shocks to expected returns, ceteris paribus, reduce an investment's value.[6]

[5] Conrad and Kaul (1988) use a similar specification but assume $\rho = 0$.

[6] In Equation (2) we multiply the shock δ_{t+1} by λ so that returns are i.i.d. when $\lambda = 0$. With δ_{t+1} outside the parentheses and $\lambda = 0$, $\text{Cov}(r_t, r_{t-1}) = \rho\sigma_\delta^2$, meaning that one would also need to assume $\rho = 0$ for

The parameter $\lambda \geq 0$ in Equation (2) determines the persistence of shocks to μ_t. To facilitate interpretation of economic magnitudes, we express λ in our empirical analyses in terms of the annualized half-life of shocks to expected returns, H:

$$H = \frac{\log(0.5)}{\log(\lambda)} \frac{1}{N}, \tag{3}$$

where N is the number of periods per year (e.g., four for quarterly data).

1.2. Interpretation of the Return Generating Process

Since Equations (1) and (2) describe a portfolio's returns in excess of its expected return benchmark, r_t should be interpreted as an abnormal return and μ_t as its conditional expectation. In our empirical analyses, the benchmark is the CAPM and we therefore study the CAPM-adjusted returns of long-short characteristic-sorted portfolios. Under the null hypothesis of the CAPM, these risk-adjusted returns have zero means at each point in time – not just on average. In our notation, the conditional expected returns μ_t are *always* zero under the CAPM null, and thus the realized returns $r_t = \epsilon_t$ are serially uncorrelated. As we discuss in the Introduction, traditional tests, which appropriately assume independent residuals, convincingly reject the CAPM.

To understand what causes these rejections, it is important to think more explicitly about alternative hypotheses. Consider, for instance, behavioral explanations that attribute the value premium to information processing biases such as overconfidence (Daniel, Hirsleifer, and Subrahmanyam, 1998) or optimism (Shiller, 2000). As Shiller (2000) points out, waves of investor optimism about new technologies (e.g., electrification in the 1920s, the internet

returns to be independent across time.

in the late 1990s) may generate episodes in which growth stocks become overvalued and subsequently reverse. Similarly, Altı and Titman (2019) develop a dynamic behavioral model in which biased investor beliefs about disruptive innovations generate cycles of positive *and* negative conditional value premia. More generally, one would expect the abnormal performance of characteristic-sorted portfolios to fluctuate over time in response to macroeconomic, structural, and technological shocks. The statistical model above captures such fluctuations, allowing both deviations of the unconditional mean from the CAPM benchmark (i.e., $\mu \neq 0$), and time variation in conditional expected returns μ_t.[7]

The existing empirical literature addresses time variation in expected returns using the Newey and West (1987) standard error adjustment, testing the implicit null that conditional expected returns μ_t fluctuate around an unconditional mean $\mu = 0$. However, as we show in Sections 3.1 and 3.2, the Newey-West procedure rejects this null hypothesis too often. When returns generated by our model exhibit a high degree of persistence, the asymptotic Newey-West standard errors are severely downward-biased with typical sample sizes.

Our empirical analyses treat the return generating process in Equations (1) and (2) as correctly specified. The AR(1) specification for μ_t is parsimonious – it captures persistent, mean-reverting variation in expected returns with a single persistence parameter – but it excludes alternatives such as structural breaks, regime shifts, non-stationary drifts toward zero, and heavier-tailed shocks. Two observations suggest that our main conclusions are not artifacts of the AR(1) specification. First, our central inference result – that persistent variation in μ_t makes unconditional expected returns imprecisely estimated – would only be

[7]Time variation in μ_t may also result from a mis-specified benchmark rather than behavioral biases. In the rational theories of Gârleanu, Kogan, and Panageas (2012) and Kogan, Papanikolaou, and Stoffman (2020), for example, investors hold lower-returning growth stocks to hedge technology shocks; if the pertinent fundamentals in these models fluctuate, so would abnormal returns relative to the CAPM benchmark.

reinforced by richer persistence structures, which add low-frequency variation that further erodes the information in a finite sample. Second, our out-of-sample analysis in Section 4 provides a specification check that does not rely on in-sample fit: forecasts constructed from the model predict future returns and subsume the information in past return averages, which would be unlikely if the model poorly characterized expected return dynamics. Nevertheless, our estimates should be interpreted as being conditional on our specification.

1.3. Estimation and Identification

The econometrician does not observe μ_t, but can estimate it – along with other model parameters – from the observed return realizations $\mathbf{R} = [r_1, r_2, \ldots, r_T]'$. Given parameters $\Omega = [\mu, \lambda, \sigma_\epsilon, \sigma_\delta, \rho]$, $\mathbf{R}$ has the following mean and covariance matrix:

$$\mathbb{E}(\mathbf{R}|\Omega) = \mu, \tag{4}$$

$$\text{Cov}\,(\mathbf{R}|\Omega) = \mathbf{\Sigma}(\Omega), \qquad \mathbf{\Sigma}(\Omega)_{i,j} = \begin{cases} \frac{\lambda^2\sigma_\delta^2}{1-\lambda^2} + \sigma_\epsilon^2 & \text{if } i = j \\ \lambda^{|i-j|}\left(\frac{\lambda^2\sigma_\delta^2}{1-\lambda^2} + \rho\sigma_\delta\sigma_\epsilon\right) & \text{if } i \neq j \end{cases}. \tag{5}$$

Equation (5) shows that shocks to both the expected and unexpected returns contribute to the volatility of returns (the terms σ_δ^2 and σ_ϵ^2, respectively). Note also that the covariance between r_i and r_j decays at a constant rate λ as $|i - j|$ grows.

The model is over-parameterized in the sense that multiple values of Ω lead to the same predicted moments $\mathbb{E}(\mathbf{R}|\Omega) = \mu$ and $\text{Cov}(\mathbf{R}|\Omega) = \mathbf{\Sigma}(\Omega)$. To see this, consider the variance

and one-lag autocorrelation of returns:

$$\sigma_r^2(\Omega) = \text{Var}(r_t) = \frac{\lambda^2\sigma_\delta^2}{1-\lambda^2} + \sigma_\epsilon^2, \tag{6}$$

$$\gamma(\Omega) = \text{Corr}(r_{t+1}, r_t) = \lambda\frac{\lambda^2\sigma_\delta^2 + (1-\lambda^2)\rho\sigma_\delta\sigma_\epsilon}{\lambda^2\sigma_\delta^2 + (1-\lambda^2)\sigma_\epsilon^2}. \tag{7}$$

Using this alternative notation, and omitting $\cdot(\Omega)$ for brevity, the covariance matrix becomes:

$$\mathbf{\Sigma}_{i,j} = \begin{cases} \sigma_r^2 & \text{if } i = j, \\ \lambda^{|i-j|-1}\gamma\sigma_r^2 & \text{if } i \neq j. \end{cases} \tag{8}$$

Inspecting Equation (8), we see that any two parameterizations Ω and $\tilde{\Omega}$ which yield the same λ, σ_r, and γ will result in the same covariance matrix $\mathbf{\Sigma}$ for returns.

To better understand identification in our model, note that we can use the sample mean of returns to estimate the unconditional expected return μ (Equation (4)), and the average rate at which the covariance between r_i and r_j decays as $|i-j|$ grows to estimate the persistence parameter λ (Equation (5)). The identification problem arises because we have three other model parameters to be identified (σ_ϵ, σ_δ, and ρ), but only two other moments that can be estimated: the variance of returns σ_r^2 in Equation (6) and the one-lag autocorrelation γ in Equation (7).[8] Intuitively, one cannot distinguish among the channels that generate return variance and autocorrelation: an increase in the volatility of expected return shocks σ_δ increases both return variance and autocorrelation, but so do increases in the volatility of unexpected return shocks σ_ϵ and the correlation parameter ρ.

[8]Note from Equation (8) that return covariances at longer lags do not provide any additional information about the model parameters other than λ because all these covariance terms are scaled by $\gamma\sigma_r^2$.

We address this identification problem in our frequentist analysis by estimating the four moments $\theta = [\mu, \lambda, \sigma_r, \gamma]$, which we can identify, rather than the full set of underlying parameters Ω. In doing so, we apply the constraint that there must be a parameterization Ω which is consistent with θ and satisfies $\sigma_\epsilon > 0$, $\sigma_\delta \geq 0$, $\lambda \geq 0$, and $\rho \in (-1, 1)$. The identification problem does not arise in our Bayesian analysis because we compute a posterior distribution for parameter values, which is unique given a set of priors and observed data, rather than a single point estimate, which is not unique.

1.4. The Sign of Return Autocorrelations

We conclude this section with a brief discussion of the return autocorrelations that the model generates. When expected returns are highly persistent, realized returns are likely to be positively autocorrelated. Indeed, our empirical autocorrelation estimates in subsequent sections are positive. It is important to note, however, that our model can produce negative as well as positive return autocorrelations, depending on the effects of various parameters.

Specifically, a high degree of persistence in expected returns, i.e., high λ, is a source of positive autocorrelation because current and recent returns have similar conditional expectations. On the other hand, a negative correlation between expected and unexpected return shocks, i.e., $\rho < 0$, is a source of negative autocorrelation. We can see these opposing forces formally in Equation (7). The first term in the numerator, $\lambda^2\sigma_\delta^2$, represents the positive effect from persistence in expected returns. The second term, $(1 - \lambda^2)\rho\sigma_\delta\sigma_\epsilon$, represents the correlation between shocks to expected and unexpected returns and is negative when $\rho < 0$. Comparing these terms, we see that the sign of γ is determined by the relative magnitudes of $\lambda^2/(1 - \lambda^2)$ and $|\rho\sigma_\epsilon/\sigma_\delta|$.

In Appendix Table 1, we report values of γ for various combinations of H (which is monotonic in λ), ρ, and $\sigma_\epsilon/\sigma_\delta$. While we do not have a strong prior on H, it is likely that ρ is somewhat negative (e.g., in the -0.1 to -0.5 range), while $\sigma_\epsilon/\sigma_\delta$ is relatively large (i.e., unexpected return shocks are much more volatile than expected return shocks). For such parameter combinations, the model tends to generate materially positive γ for relatively high values of H. Negative values of γ obtain for very high $\sigma_\epsilon/\sigma_\delta$ and/or small H, but the magnitudes are quite small and close to zero in most cases.

2. Characteristic-Sorted Portfolios

We first describe the historical returns of portfolios sorted on value, investment, profitability, and size – the basis of the Fama and French (2015) five-factor model. Subsequent sections extend our analysis to a broader set of 174 portfolios.

2.1. Data and Characteristic Definitions

We obtain data on the historical returns of the four main characteristic-sorted portfolios we examine from Ken French's website.[9] Each of these portfolios includes a long position in a value-weighted portfolio of firms in the top quintile of the characteristic with a short position in the bottom quintile. Following Fama and French (2015), we define *Value* as the ratio of the book value of equity ($B_{i,y}$) to the market value of equity ($M_{i,y}$) at the end of the prior fiscal year y, *Investment* as the growth rate in the book value of assets ($\text{Assets}_{i,y}/\text{Assets}_{i,y-1}$), and *Profitability* as revenues minus cost of goods sold, interest expense, and selling, general, and administrative expenses in year y divided by book equity in year $y-1$. *Size* is $M_{i,y}$.

[9] `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

We also extend our analysis to a broader set of 174 characteristic-sorted portfolios. These include 170 long-short equity strategies described in the March 2022 update of the "Open Source Cross-Sectional Asset Pricing" dataset (Chen and Zimmermann, 2021) with at least 120 months of historical data, along with the four portfolios described above.

In contrast to most of the literature, which examines the monthly returns of characteristic sorted portfolios, we analyze quarterly returns. As we show in Appendix A, the monthly returns of three of the four main portfolios we study exhibit strong positive first-order autocorrelations.[10] While these short-term autocorrelations are consistent with time-varying expected returns, they could also be driven by lead-lag effects and other short-term microstructure effects. Estimating our model on quarterly returns diminishes these microstructure effects and focuses our analysis on longer-term autocorrelations driven by persistent variations in expected returns.

Since the market risk premium is not our focus, we use market-neutral versions of each portfolio's returns throughout, calculated as:

$$r_{i,t}^{\beta=0} = r_{i,t} - \hat{\beta}_i(r_{m,t} - r_{f,t}), \tag{9}$$

where $r_{i,t}$ is the quarterly return of the long-short portfolio i, $r_{m,t} - r_{f,t}$ is the quarterly excess market return, and $\hat{\beta}_i$ is the full-sample market beta.[11] The average return and Sharpe ratio of this market-neutral portfolio are equivalent to the alpha and information ratio, respectively, of the underlying long-short portfolio.

[10] This is consistent with the evidence in Gupta and Kelly (2019) that 47 of 65 characteristic-based portfolios have significantly positive first-order autocorrelation.

[11] We re-estimate $\hat{\beta}_i$ in each quarter using only past data for our out-of-sample tests.

2.2. Characteristic-Sorted Portfolio Performance and Autocorrelation

Table 1 presents summary statistics for the returns of the four characteristic-sorted portfolios. As shown in Panel A, the value, investment, and profitability portfolios exhibit large annualized mean returns and Sharpe ratios between 0.25 and 0.46, while the size portfolio's mean return is close to zero. Based on standard errors calculated under the assumption of i.i.d. returns, we strongly reject the null hypothesis of zero expected returns for the first three portfolios.[12]

We also examine the historical performance of the portfolios in the first and second halves of our sample period, 1963–1992 and 1993–2021. As discussed in McLean and Pontiff (2016), Linnainmaa and Roberts (2018), and Fama and French (2021), the value strategy's average returns are smaller in the second half of the sample and are not statistically different from zero. However, as emphasized by Fama and French (2021), the difference between the two halves of the sample is not statistically significant.[13]

The investment, profitability, and size portfolios each show different return patterns across subsamples. The investment portfolio's returns are similar and statistically significant in both halves of the sample. The profitability portfolio follows the opposite pattern to the value portfolio, performing better in the second half of the sample than the first, though again the difference is statistically insignificant. The size portfolio has small and statistically insignificant returns in both halves of the sample.

[12]We compute the i.i.d. standard errors by taking the standard deviation of sample means across simulated samples formed by re-sampling historical data with replacement.

[13]Recent studies provide two potential explanations for the decline of the value premium: book value may have worsened as a proxy for the value of assets in place (Choi, So, and Wang, 2021; Eisfeldt, Kim, and Papanikolaou, 2022; Goncalves and Leonard, 2023), or a series of shocks may have widened the difference in multiples between growth and value stocks (Israel, Laursen, and Richardson, 2020; Arnott et al., 2021b). Both imply time variation in expected returns and thus are consistent with our model.

Panel B of Table 1 reports return autocorrelation estimates for the four characteristic-sorted portfolios. We estimate these autocorrelations using regressions of quarterly returns on averages of returns over the past L quarters:

$$r_t = a + b_L \left(\frac{1}{L} \sum_{l=1}^{L} r_{t-l} \right) + \epsilon_t. \tag{10}$$

The first column in Panel B shows that the autocorrelation coefficient using the past four quarters, b_4, is positive for all four portfolios. The size portfolio exhibits the strongest autocorrelation pattern, with a coefficient of 0.34 and a p-value of 0.1%. The coefficient is significantly positive for the investment portfolio as well, while it is insignificant for the value and the profitability portfolios. The coefficient estimates using longer lags have varying signs and are generally insignificant.

We also consider a panel regression that estimates a single b_L across all four portfolios. In this setting, we reject the i.i.d. null hypothesis for one-year and five-year returns with p-values of 0.2% and 5.1%, respectively.[14] In a panel regression with only value, investment, and profitability, we reject the i.i.d. null with the one-year lag with a p-value of 5.7%.

3. OLS and Maximum Likelihood Estimations

In our model, the long-run performance of a characteristic-sorted portfolio is determined by the unconditional expected return μ. In this section, we analyze how the precision of the estimates of μ is affected by persistent variation in conditional expected returns. We first consider OLS regressions, where we use our model's structure to correct for standard

[14] We conduct this statistical test by comparing the panel data coefficient to panel coefficient estimates from samples simulated under the i.i.d. assumption.

errors given assumed values for the persistence parameters. We then estimate the full model, including the persistence parameters, using maximum likelihood.

3.1. OLS Estimation With Model-Corrected Standard Errors

We start by estimating the unconditional expected return μ and its standard error using OLS regressions of observed returns r_t on a constant. The OLS estimate $\hat{\mu}^{OLS}$ is a consistent estimator of μ, even when conditional expected returns vary, as in our model. The correct standard errors for $\hat{\mu}^{OLS}$ depend on the covariance matrix of the residuals $\psi_t = r_t - \mu$.[15]

The typical approaches used to adjust standard errors are the White (1980) correction for potential heteroskedasticity and Newey and West (1987), which corrects for both heteroskedasticity and autocorrelations up to a small number of lags (relative to the length of the time series). When expected returns are time-varying and persistent, both of these approaches produce understated standard errors. Formally, our model generates the residuals

$$\psi_t = \mu_{t-1} - \mu + \epsilon_t. \tag{11}$$

Persistent variation in conditional expected returns μ_t implies that ψ_t exhibits long-lasting autocorrelations. As we show below, the asymptotic results in Newey and West (1987) are not applicable in this case, because the number of lags that is necessary to detect these long-lasting autocorrelations is too large a fraction of the observed time series.

If the first-order return autocorrelation γ and the half-life of shocks H are known, we can correct the OLS standard errors for the resulting autocorrelation in the residuals ψ_t using

[15] We use 'correct standard errors' as an informal shorthand for the correct specification of the asymptotic distribution of $\hat{\mu}$.

the structure of our model.[16] Specifically, the standard error of $\hat{\mu}^{OLS}$ in this case is

$$\begin{aligned} \text{SE}(\hat{\mu}^{OLS}) &= \sqrt{\frac{\mathbf{1'\Sigma 1}}{T^2}} \\ &= \left(\frac{\sigma_r}{\sqrt{T}}\right)\sqrt{1+2\gamma\left[\frac{\lambda^T + T(1-\lambda) - 1}{T(1-\lambda)^2}\right]}. \end{aligned} \tag{12}$$

In Equation (12), T is the number of observations, $\mathbf{1}$ is a $T \times 1$ vector of ones, and $\mathbf{\Sigma}$ is the covariance matrix of returns. The first line is the general formula for computing standard errors with a known covariance matrix (see Section 4.5 of Cameron and Trivedi (2005)). In our model, $\mathbf{\Sigma}$ is specified by γ and λ, as shown in Equation (8). Substituting from Equation (8) results in the second line of Equation (12) after some algebraic manipulation.

Note that the first term in Equation (12), $\sigma_r/\sqrt{T}$, is the unadjusted OLS standard error. Thus, the second, square-root term is the adjustment factor, which is a function of T, γ, and λ. This adjustment factor converges to $\sqrt{1+2\gamma/(1-\lambda)}$ as $T \to \infty$. Therefore, for large T, the adjusted standard error in Equation (12) approximately equals:

$$\text{SE}(\hat{\mu}^{OLS}) \approx \left(\frac{\sigma_r}{\sqrt{T}}\right)\sqrt{1+\frac{2\gamma}{1-\lambda}}. \tag{13}$$

The standard error adjustments in Equations (12) and (13) can also be stated in terms of the annualized half-life H instead of λ by substituting $\lambda = 0.5^{\frac{1}{HN}}$, which follows from Equation (3), where N is the number of return observation periods per year.

Table 2 shows how different assumptions about γ and H affect the standard errors of

[16] Note that the model structure can also be used to obtain Generalized Least Squares (GLS) estimates. As detailed in Appendix B, we find that the GLS estimates differ only slightly from the OLS estimates for the four characteristic-sorted portfolios we study.

unconditional expected return estimates. For comparison, we also report standard errors that are adjusted using the Newey-West procedure with the number of lags matching the half-life of shocks H. For the value, investment, and profitability portfolios, we find that Newey-West standard errors differ very little from unadjusted standard errors. Newey-West standard errors are around 50% larger for the size portfolio, which is a reflection of the stronger evidence of persistence for the size portfolio presented in Table 1.

In contrast to the Newey-West adjustment, model-implied standard errors that are computed using Equation (12) tend to be larger for all four portfolios. The magnitude of this adjustment depends on both H and γ. The adjustments to standard errors are relatively small when expected return shocks are assumed to quickly mean-revert ($H = 0.5$ years). With more persistent shocks (H between 2.5 and 10 years), however, Table 2 shows that model-implied standard errors are much larger. For instance, standard errors approximately double relative to their unadjusted counterparts when $H = 5$ years and $\gamma = 5\%$. Overall, the results in Table 2 indicate that accounting for plausible magnitudes of persistent variation in expected returns results in inferences about unconditional expected returns that are materially less precise than inferences that obtain under the assumption of i.i.d. returns or with the Newey-West adjustment.

3.2. Failure of the Newey-West Correction

Table 2 shows that with historical data, Newey-West standard errors are often much smaller than those implied by our model. We next show that when we simulate data from our model, allowing us to know the true distribution of $\hat{\mu}^{OLS}$, Newey-West standard errors are severely downward biased. Specifically, for a given set of parameter values, we generate

250,000 samples with 234 observations, which is the number of quarters in our empirical sample. We consider various values for H, while γ equals 5% in all simulations.[17] In each simulated sample, we compute standard errors for $\hat{\mu}^{OLS}$ using the Newey-West adjustment with lags ranging from one to 20 years. We then measure bias by dividing these adjusted standard errors by the simulated standard error, which is the standard deviation of the estimate of $\hat{\mu}^{OLS}$ across simulations.

Panel A of Table 3 shows that the Newey-West standard errors are substantially downward biased. The bias is stronger when expected return variation is more persistent (higher H), with standard errors often 25%–50% smaller than their simulated counterparts. This leads the Newey-West procedure to reject the null too frequently, with t-statistics above the 5% critical value in 15–40% of our simulations. Interestingly, there is some bias even with $H = 0$, i.e., with i.i.d. returns. The reason is that the Newey-West adjustment embeds a small-sample estimate of autocorrelation, which is downward biased (Kendall, 1954; Marriott and Pope, 1954). This effect worsens as Newey-West uses more lags because long-run autocorrelations have even larger downward biases.

To illustrate the downward bias and imprecision of autocorrelation estimates in our model setting, we estimate the regression coefficients $\hat{b}_L$ in Equation (10) for $L = 4$ and $L = 20$ (i.e., regressions of quarterly returns on prior-year and prior five-year average returns) in each simulated sample. Panel B of Table 3 presents average values of $\hat{b}_L$ across simulated samples, and the true b_L for comparison. Our results confirm both the presence of the downward bias, and that the bias increases in both L and H. This downward bias contributes to the

[17] The remaining model parameters we use for the simulations are $\mu = 0$ and $\sigma_r = 7.57\%$. These choices have no effect on our results because autocorrelation estimates are invariant to linear transformations of r_t.

small-sample bias in Newey-West that we document in Panel A because the Newey-West adjustment uses autocorrelation estimates rather than true autocorrelations.[18] However, in untabulated tests we find that the bias in Newey-West is smaller but still present in simulated samples where the autocorrelation estimate is near the true value. This indicates that other properties of this setting, such as the long persistence relative to the sample size, contribute to the small-sample bias in Newey-West.

Panel B of Table 3 also shows that the 95% confidence intervals estimated across simulations for $\hat{b}_4$ and especially $\hat{b}_{20}$ are extremely wide. Autocorrelation coefficient estimates are inherently imprecise because realized portfolio returns are quite volatile relative to plausible variations in expected returns. The great degree of overlap in the 95% confidence intervals across simulations with different values of H, all of which encompass the point estimates for value, investment, and profitability given in Panel B of Table 1, illustrates this point.

3.3. Maximum Likelihood Tests

In this section, we present maximum likelihood estimations of our model. Unlike OLS, maximum likelihood requires distributional assumptions but allows us to estimate, rather than assume, the model parameters governing persistent variation in expected returns.

The maximum likelihood estimation of our model requires a parametric restriction due to a statistical power problem: a model of returns with small but extremely persistent autocorrelation is indistinguishable in finite samples from a model with i.i.d. returns. To see this point, consider a parameterization where the unconditional expected return μ equals zero, the first-order autocorrelation γ is small but positive, and the half-life H is very large.

[18] Hodrick (1992) documents a similar small-sample bias for Newey-West inferences in a time-series return predictability setting.

In this case, conditional expected returns can deviate substantially from zero since small shocks accumulate over time. Yet, conditional expected returns exhibit little time variation within a finite sample, thus resembling an i.i.d. process with a large (absolute) mean.

Given this statistical power problem, the maximum likelihood estimations we report below restrict the half-life parameter H to be less than or equal to 10 years. We note that this restriction is primarily an econometric device rather than an economic assumption. Some of the economic mechanisms that motivate our analysis – structural and technological change, shifts in investor composition, or the gradual deterioration of a characteristic's informativeness – could in principle generate shocks to expected returns that persist for several decades. However, since half-lives of that magnitude are nearly indistinguishable in our sample from permanent shifts in unconditional expected returns, allowing them would leave the boundary between μ and μ_t empirically undefined. We report the sensitivity of our results to the restriction on the half-life parameter H below.

3.3.1. Hypothesis Tests for Unconditional Expected Returns

We use maximum likelihood estimates of the model to test the hypothesis that the unconditional expected return μ is zero. Specifically, for each characteristic-sorted portfolio, we estimate the model under two alternative specifications. The first assumes that returns are i.i.d., and the second allows expected returns to be time-varying. In each case, we estimate the model parameters twice; first with no restrictions on μ and then under the restriction that $\mu = 0$. We then use these estimates to conduct a likelihood ratio test for the $\mu = 0$ null.[19] These maximum likelihood estimates are reported in Panel A of Table 4.

[19]The small-sample downward bias in autocorrelation estimates discussed in Section 3.1 also affects maximum likelihood estimates. To account for this bias, we compute likelihood ratio p-values from the finite-

The first three columns in Panel A of Table 4 test the hypothesis that $\mu = 0$ using model specifications that do not admit any time variation in expected returns (i.e., where returns are assumed to be i.i.d.). Testing $\mu = 0$ under this assumption is analogous to an OLS test with no standard error correction. Not surprisingly, the likelihood ratio tests strongly reject the hypothesis that $\mu = 0$ for the investment and profitability portfolios, with p-values less than 0.1%. For the value portfolio, the hypothesis is rejected with a p-value of 5.5%.

The next five columns in Panel A of Table 4 relax the i.i.d. assumption and estimate the values of γ and H that maximize the likelihood of observing the historical data.[20] We find that the estimate for the autocorrelation parameter γ is positive for all four portfolios, ranging from 7.65% for the profitability portfolio to 14.03% for the size portfolio. The estimate for the half-life parameter H is about one year or less for all four portfolios. However, as we discuss in Section 3.3.2, these estimates are quite imprecise.

Importantly, likelihood ratio tests fail to reject the null hypothesis that $\mu = 0$ for any of the four portfolios when expected returns are allowed to be time varying. For instance, likelihood ratio p-values for the investment and profitability portfolios increase to 9.5% and 10.4%, respectively, which are in strong contrast with less than 0.1% in the i.i.d. case. Intuitively, admitting time-varying expected returns increases the likelihood that the historical return performance of these portfolios during the sample period reflects higher than average realizations of conditional expected returns.

It should be noted that under the $\mu = 0$ null hypothesis, the estimated H for the in-

sample distribution of the statistic, simulated from the model with parameters estimated under the null. p-values based on the asymptotic distribution, or small-sample i.i.d. simulations, are available upon request.

[20] As discussed in Section 1.3, we estimate $\theta = [\mu, H, \sigma_r, \gamma]$ rather than the underlying parameters $\Omega = [\mu, \lambda, \sigma_\epsilon, \sigma_\delta, \rho]$ because the latter are not fully identified. We restrict the admissible values of θ to the range for which there exists at least one possible Ω yielding the same covariance matrix of returns as θ.

vestment and profitability portfolios equals 10 years; i.e., the limit we impose on H binds. Given this, we examine how our conclusions depend on this limit. Figure 1 plots, for upper bounds H_{ub} ranging from five to 30 years, the p-value of the $\mu = 0$ likelihood ratio test, along with the likelihood of the observed data as a function of H with the remaining parameters re-estimated at each H.

The likelihood profiles in Figure 1 illustrate why the data cannot resolve the choice of H_{ub}: for the value, investment, and profitability portfolios, the likelihood is nearly flat across the upper range of H, so extending the admissible range adds parameterizations that fit the data essentially as well as the unrestricted maximum. Notably, the size portfolio is the only one for which the likelihood declines steeply at long half-lives, reflecting the stronger evidence of relatively short-lived persistence for this portfolio discussed below.

The p-values in Figure 1, in contrast, do depend on H_{ub} for the investment and profitability portfolios, rising from 1.5% and 4.2% when $H_{ub} = 5$ years to 18.1% and 26.3% when $H_{ub} = 30$ years, as looser limits admit more persistent processes that can explain these portfolios' performance with a zero unconditional mean. The p-values for the value and size portfolios are insensitive to H_{ub}, remaining near 13.3% and 97.4%, respectively. We formalize the implications of beliefs about H in our Bayesian analysis of Section 4.

3.3.2. Hypothesis Tests for H and γ

Next, we use maximum likelihood estimates to assess the plausibility of various hypotheses for the magnitude of time variation in expected returns. Specifically, we consider the assumptions on H and γ that we examined in Section 3. For each assumption, we re-estimate the model while restricting H and γ to their assumed values, and calculate likelihood ratios

relative to the unrestricted model to test whether we can reject the restriction.

Panel B of Table 4 presents the results. For the value, investment, and profitability portfolios, we find that *none* of the hypotheses on time-variation is rejected: the likelihood ratio p-values for the hypothesized restrictions all exceed 10%, and are often greater than 30%. Thus, historical returns of these portfolios are consistent with no time-variation (i.e., i.i.d. returns), as well as large and persistent variation in expected returns.

While the results for the value, investment, and profitability portfolios suggest that the historical data offer little guidance about the magnitude of time-variation in expected returns, the results for the size portfolio show that this is not always the case. Consistent with the reduced-form autocorrelation estimates in Table 1, the estimates for the size portfolio in Table 4 point to strong positive autocorrelation in returns that decays with a half-life between one and two years. As a result, the likelihood-ratio tests strongly reject the i.i.d. hypothesis, while also rejecting $H = 10$ with p-values less than 10%.

Importantly, the rejection of the i.i.d. hypothesis implies that the size portfolio returns deviate significantly from the CAPM benchmark. In contrast, traditional asset pricing tests, which solely focus on average returns and ignore time variation, fail to reject the CAPM for the size portfolio, since the portfolio's historical average performance is close to zero (see Table 1). Thus, by ignoring time variation in expected returns, traditional tests may generate not only false positives too frequently (rejecting $\mu = 0$ when in fact returns exhibit persistent variation around zero unconditional means), but also false negatives too frequently (failing to reject the CAPM when in fact conditional expected returns deviate from it).

3.3.3. Hypothesis Tests for the Broader Set of Characteristic-Sorted Portfolios

We now extend our analysis to the broader set of 174 characteristic-sorted portfolios described in Section 2.1. The analysis of Chen and Zimmermann (2021) illustrates the apparent ease with which research uncovers significant relationships between firm characteristics and average stock returns. As Harvey, Liu, and Zhu (2016) and others discuss, this evidence may be partly due to p-hacking; i.e., researchers testing many different characteristics and reporting only those that turn out to have a statistically significant relation to future returns. Harvey, Liu, and Zhu (2016) suggests addressing p-hacking by requiring higher t-statistic (or lower p-value) thresholds for rejecting the no-predictability null hypothesis. As Chen and Zimmermann (2021) show, however, higher significance thresholds do not completely eliminate these anomalies. For instance, 58 of the 174 portfolios in our sample have expected return estimates with unadjusted OLS p-values below 0.1%, which is extremely unlikely under the null, even with p-hacking.

To assess how persistent return variation affects the precision of unconditional expected return estimates more generally, we estimate our model using maximum likelihood for each of the 174 portfolios, and test the $\mu = 0$ null hypothesis using maximum likelihood as in Section 3.3.1. Figure 2 summarizes the results of this exercise, showing the number of portfolios that attain various statistical significance levels. For comparison, the figure also summarizes the results with unadjusted and Newey-West adjusted OLS regressions.[21] As the figure shows, rejections of the $\mu = 0$ null hypothesis are substantially weaker with the maximum likelihood estimates that allow for persistent variation in returns. The gap between the MLE and OLS

[21] The complete results for our analysis of the 174 portfolios are reported in Appendix Table 4.

tests is particularly large for p-values of 0.1% and 1%, which MLE achieves for only four and 14 portfolios, respectively, compared to 58 and 81 using OLS.

3.3.4. Sample Size and Test Power

The results in Table 4 and Figure 2 indicate that inferences from historical data about unconditional expected returns and the magnitude of persistent variation in returns are generally quite imprecise. A natural question is how large the sample of returns needs to be to obtain more precise parameter estimates.

We address this question by simulating random samples of varying lengths for two different parameterizations of our model: one where returns are i.i.d., and the other with a persistent process where $H = 5$ years and $\gamma = 5\%$. In both parameterizations we assume that the unconditional expected return $\mu = 5\%$ and the annualized return volatility $\sigma_r = 15\%$. For each simulated sample, we estimate the model with maximum likelihood and test the hypothesis that $\mu = 0$. We also test the 'opposite' hypothesis about persistence (i.e., the hypothesis that $H = 5$ years and $\gamma = 5\%$ when the true data generating process is i.i.d., and the i.i.d. hypothesis when the true data generating process has $H = 5$ and $\gamma = 5\%$). In these hypothesis tests using likelihood ratios, the econometrician does not know the true data generating process and has to estimate the model exactly as we do in Table 4. The results are summarized in Figure 3, which plots the fraction of simulated samples in which the tested hypothesis is rejected with a p-value of 5% or less.[22]

Panel A of Figure 3 shows the rejection rates for the $\mu = 0$ hypothesis. With 50-year

[22]Obtaining small-sample p-values using the restricted parameter estimates (i.e., similar to those reported in Table 4) for the analysis in Figure 3 is not feasible due to computational cost. For this reason, the analysis in Figure 3 is based on p-values that obtain with asymptotic likelihood-ratio distributions. We note that the sample lengths considered in Figure 3 (i.e., several hundred years) mitigate the concern about the reliability of asymptotic p-values.

samples, which is close to the length of the historical sample, the rejection rate for $\mu = 0$ is around 70% when the true data generating process is i.i.d. Thus, there is a 30% chance that one would fail to reject $\mu = 0$ with 50 years of data despite the fact that $\mu = 5\%$. When the return data are generated by the persistent process, the failure to reject $\mu = 0$ becomes even more acute, with only about 37% of samples rejecting. Longer samples allow i.i.d. samples to converge to the true values relatively more quickly: with 200 years, more than 95% of the samples reject $\mu = 0$. In contrast, reaching a similar rejection rate requires 800 years when returns are generated by the persistent process.

Panel B of Figure 3 shows the rejection rates for hypotheses about the magnitude of persistent variation in returns. The i.i.d. and the persistent return processes that generate the simulated data are quite different from each other in terms of the economic interpretation of the model. Yet there is little power to distinguish between them with 50 years of data: only about 40% (20%) of the samples reject the persistent (i.i.d.) process when returns are in fact i.i.d. (persistent). As the figure shows, reliably distinguishing between these alternative return generating processes requires at least 400 years of data.

To summarize, our simulation findings indicate that obtaining precise estimates of unconditional expected returns requires very long samples, especially when returns exhibit persistent variation. Furthermore, distinguishing between i.i.d. and persistent returns is challenging in the absence of such long samples. The imprecision of inferences from historical data has implications for investors' portfolio decisions, which we analyze next.

4. Bayesian Analysis

Our results in the previous section suggest that, given the possibility of persistent variation in expected returns, reaching precise conclusions about model parameters may require a longer time series than is available. This does not imply, however, that data from shorter samples are uninformative. As Kandel and Stambaugh (1996), Wachter and Warusawitharana (2009), and others have observed, Bayesian investors will use data to substantially update their beliefs about the return generating process even if statistical tests of data do not reject null hypotheses at conventional significance levels. Therefore, a natural next step for our analysis is to ask how investors with prior beliefs about the return generating process interpret historical data, and how this in turn influences their investment decisions. We examine these normative questions in this section by conducting a Bayesian analysis which delivers posterior distributions for both model parameters and moments that affect investors' portfolio timing decisions, such as the conditional Sharpe ratios at each point in time.

4.1. Prior Beliefs About Model Parameters

To facilitate economic interpretation of magnitudes, our Bayesian analysis specifies prior beliefs on a transformation of the model parameters expressed in terms of annualized Sharpe ratios rather than expected returns.[23] Specifically, we specify priors over μ_{sr}, the unconditional Sharpe ratio of portfolio returns; σ_r, the volatility of portfolio returns; H, the half-life of shocks to expected returns; σ_{sr}, the volatility of the conditional Sharpe ratio; and ρ, the correlation between unexpected and expected return shocks. The mapping between these

[23] Another advantage of this transformation is that the same Sharpe ratio priors can be applied across all characteristic-sorted portfolios regardless of their volatility levels.

annualized parameters and the underlying model parameters $\Omega = [\mu, \lambda, \sigma_\epsilon, \sigma_\delta, \rho]$ is given by the following equations, where $N = 4$ is the number of periods per year:

$$\mu_{sr} = \frac{\mu}{\sigma_\epsilon}\sqrt{N}, \qquad \sigma_r = \sqrt{\left(\frac{\lambda^2\sigma_\delta^2}{1-\lambda^2} + \sigma_\epsilon^2\right)N}, \tag{14}$$

$$H = \frac{\log(0.5)}{\log(\lambda)}\frac{1}{N}, \qquad \sigma_{sr} = \sqrt{\left(\frac{\lambda^2\sigma_\delta^2/(1-\lambda^2)}{\sigma_\epsilon^2}\right)N}, \qquad \rho = \rho.$$

Next we describe the prior beliefs for the transformed parameters in Equation (14). For the unconditional Sharpe ratio μ_{sr}, we specify the prior as a normal distribution with a mean of zero and standard deviation of 0.4.[24] To illustrate how prior beliefs about the half-life of shocks H influence Bayesian inferences about expected returns, we consider three dogmatic and three agnostic priors. The dogmatic priors assert that $H = 0$ (making returns i.i.d.), $H = 2.5$ years, or $H = 5$ years with certainty. The three agnostic priors view H as unknown and uniformly distributed between zero and 10, 20, or 30 years.

We consider uniform priors over wide but plausible ranges for the remaining parameters. The prior for the volatility of annual returns, σ_r, is uniformly distributed between 10% and 20%. The prior for the volatility of the conditional Sharpe ratio, σ_{sr}, is uniformly distributed between 0 and 1. The correlation between unexpected and expected return shocks, ρ, is likely to be negative given the inverse relation between prices and expected returns. In contrast to aggregate market returns, however, characteristic-sorted portfolio returns should be driven primarily by cash-flow news rather than discount rate news. Based on these observations, we specify the prior on ρ to be uniformly distributed in the interval $[-0.5, 0]$.

To provide intuition for economic magnitudes, Panel A of Table 5 reports summary

[24] In unreported results we also considered priors for μ_{sr} that are uniformly distributed between -2 and 2.

statistics for a number of moments that are implied by the priors described above.[25] The prior for μ_{sr} implies a 95% confidence interval $[-0.78, 0.78]$. Thus, the Bayesian investor has a neutral prior for unconditional return performance but does allow for substantial deviations from the zero benchmark. The prior distribution of the one-lag return autocorrelation γ has a mean around 5% and a 95% confidence interval that ranges from small negative values to about 17% in all of the prior specifications for H.

The parameter σ_{sr} governs the volatility of the true conditional Sharpe ratio, defined as the conditional expected return μ_t in Equation (2) divided by the standard deviation of unexpected return shocks σ_ϵ. The conditional expected return μ_t is unlikely to be perfectly observable to investors in practice. For an investor who knows the model parameters and observes an infinite return history, but not μ_t itself, the relevant measure of time variation is $\hat{\sigma}_{sr}$, the volatility of the Sharpe ratio conditional on past realized returns. The last two columns in Panel A of Table 5 report summary statistics for $\hat{\sigma}_{sr}$ for each prior specification we consider. They show that the means and the 95% confidence intervals for $\hat{\sigma}_{sr}$ are about half as large as those for σ_{sr}.[26]

4.2. Posteriors for Unconditional Moments

We compute posterior distributions for 24 prior-data pairs (six prior specifications for H, four characteristic-sorted portfolios). For each pair, we generate 250,000 posterior draws using the Markov Chain Monte Carlo (MCMC) procedure detailed in Appendix C. Although we compute posteriors for the full parameter vector Ω, Panels B through E of Table 5 report

[25] When necessary, we calculate these summary statistics by simulating 250,000 draws from each prior.

[26] Recall that we specify the prior for σ_{sr} to be uniformly distributed between 0 and 1, which implies a mean of 0.5 and a 95% confidence interval of $[0.025, 0.975]$.

only the parameters and moments of primary economic interest.

The first two columns in Table 5 report summary statistics for the posterior distribution of the unconditional Sharpe ratio, μ_{sr}. The results show that different priors about H result in substantially different posterior beliefs about μ_{sr}. For instance, the 95% confidence intervals for μ_{sr} become wider as priors for H increase, making the possibility of zero or negative unconditional Sharpe ratios much more likely. The intuition for this result is the same as in the frequentist analysis above: the data do not strongly reject the possibility that the historical return performance of characteristic-sorted portfolios is explained by persistent but dissipating positive shocks to conditional expected returns.

The remaining columns in Table 5 report summary statistics for the posterior distributions of H, γ, and $\hat{\sigma}_{sr}$. With the agnostic priors for H, the posterior confidence intervals for the value, investment, and profitability portfolios are almost as wide as the prior confidence interval. Thus, the data offer very little information about H for these three portfolios. This conclusion is not an artifact of the $U(0, 10)$ prior's upper bound: under the wider $U(0, 20)$ and $U(0, 30)$ priors, the posteriors for H continue to resemble the priors, with 97.5th percentiles that remain essentially equal to 97.5% of the corresponding prior upper bounds (for example, 19.51 and 29.29 years for the value portfolio, all but identical to the prior percentiles of 19.50 and 29.25 years). The data therefore provide essentially no evidence against even multi-decade half-lives for these portfolios. A similar conclusion emerges for γ and $\hat{\sigma}_{sr}$ as well: the posterior means for these statistics are somewhat lower than prior means, but the 95% posterior confidence intervals are still quite wide.[27]

[27]Because our conclusions regarding the value, investment, and profitability portfolios highlight the similarity between priors and posteriors, it is important to verify that this similarity reflects the information content of the data rather than poor mixing of the MCMC chain. Appendix Table 3 reports mixing diagnostics for the sampler. Acceptance rates range from 18% to 38% across the 24 prior-portfolio pairs, and the

In contrast, the posteriors for the size portfolio differ more substantially from the priors. The posterior mean of H is well below the prior mean of five years with the agnostic prior $U(0, 10)$. This pattern strengthens under the wider priors: the posterior means of H under the $U(0, 20)$ and $U(0, 30)$ priors are 4.65 and 5.96 years, well below the prior means of 10 and 15 years, and the posterior 97.5th percentiles of 17.17 and 25.64 years pull away from the prior upper bounds. Thus, unlike for the other three portfolios, the data actively push back against large values of H for the size portfolio. Furthermore, unlike the other three portfolios, the posteriors for the size portfolio's γ and $\hat{\sigma}_{sr}$ have much higher means compared to priors and confidence intervals that exclude zero, indicating that the data tilt the posterior beliefs in favor of positive autocorrelation.

4.3. Posteriors for Conditional Sharpe Ratios

In addition to unconditional Sharpe ratios, the Bayesian analysis allows us to compute posterior distributions of conditional Sharpe ratios at each point in time during our sample period. The fluctuations in the means of conditional Sharpe ratio distributions provide a gauge for how Bayesian investors can combine priors with data to engage in characteristic timing strategies. In addition, the conditional Sharpe ratios at the end of our sample period are informative about the likely performance of the portfolios in the future.

For the analysis of this section, we utilize the agnostic prior for H that is uniformly distributed between zero and 10 years. The priors for other parameters are as described in Section 4.1. We compute the posterior conditional Sharpe ratio distributions two ways:

effective sample size of the posterior draws exceeds 10,000 for every parameter in every pair, and is typically between 25,000 and 70,000. These diagnostics indicate that the sampler mixes well and that the posterior distributions we report are estimated precisely.

in-sample, using the full sample of returns, and out-of-sample, using only past data available at each point in time. We require at least 10 years of available data to compute the out-of-sample posteriors.

Figure 4 plots the time series of posterior means of conditional Sharpe ratios for the four characteristic-sorted portfolios. As the figure shows, conditional Sharpe ratios exhibit substantial and persistent fluctuations over time. These fluctuations are especially pronounced for the size portfolio, where the conditional Sharpe ratios range from -0.75 to above one. These relatively larger fluctuations are a reflection of the more positive return autocorrelations that the size portfolio exhibits during our sample period.

Another takeaway from Figure 4 is that in-sample and out-of-sample conditional Sharpe ratios are highly correlated with each other, particularly for the Investment and Size portfolios. The out-of-sample forecasts are more optimistic for value, and more pessimistic for profitability, because only the in-sample posteriors have knowledge of the downward trend in value's performance or the upward trend in profitability's performance. The large fluctuations that obtain with the out-of-sample analysis imply that Bayesian investors with standard utility functions would engage in characteristic-timing strategies that exhibit large changes in optimal portfolio weights over time.

A comparison of the conditional Sharpe ratios in Figure 4 and the unconditional Sharpe ratios in Table 5 indicates that the predicted future performance of the characteristic-sorted portfolios can be quite different from both their historical and predicted unconditional performance. For instance, because the value portfolio performed poorly toward the end of our sample period, its conditional Sharpe ratio at the end of 2021 is substantially lower than its unconditional Sharpe ratio of 0.20. Thus, the Bayesian predictions suggest continued weak

performance for value, followed by some recovery as the conditional Sharpe ratio reverts toward the unconditional. The opposite pattern obtains for the profitability portfolio. Due to its strong recent performance, the profitability portfolio is predicted to exhibit a very high conditional Sharpe ratio in the near future, followed by a reversion toward its unconditional Sharpe ratio of 0.39. The conditional Sharpe ratios for the investment and size portfolios at the end of 2021 are similar to their unconditional counterparts.

4.4. Out-of-Sample Return Forecast Accuracy

In our final analysis, we consider the accuracy of the out-of-sample Bayesian forecasts for the broader set of 174 portfolios. To facilitate this analysis, for each calendar year y in our sample starting 10 years after data become available for a given portfolio, we calculate posterior belief distributions about model parameters given all quarterly data available through y, the agnostic prior belief for H that is uniform between zero and 10 years, and the priors for other model parameters described in Section 4.1. Next, for each of the 250,000 posterior draws and each quarter in year $y+1$, we generate 10 random observations from the quarterly return distribution conditional on all prior quarters.[28] Given these 2,500,000 total draws for each quarter, we compute out-of-sample forecasts for average quarterly returns $\hat{r}_{i,t}$.

We compare the Bayesian forecasts to forecasts based on maximum likelihood estimation of our model, as well as to simple forecasts computed as average realized returns over all past quarters, past five years, and past one year. For the MLE forecasts, at each forecast date we

[28]For example, when computing conditional expected returns for Q3 of 2002, we use posterior distributions for model parameters estimated using MCMC on data through Q4 of 2001 combined with realized returns through Q2 of 2002. This combination allows us to use all available return data while only conducting the computationally-expensive MCMC procedure for calculating posterior beliefs about parameters once per year. Given the slow-moving changes in expected returns we are interested in, we would not expect large quarter-to-quarter variations in conditional distributions of model parameters.

re-estimate the model parameters by maximum likelihood using the portfolio's past market-adjusted returns, and compute the one-quarter-ahead conditional expected return implied by the point estimates. The MLE forecasts differ from the Bayesian forecasts in that they condition on a single point estimate of the model parameters, and thus ignore parameter uncertainty, rather than integrating over the posterior distribution of the parameters.

To assess the accuracy of the forecasts for realized next-quarter returns, we utilize a Fama and MacBeth (1973) regression approach. Specifically, at the end of each quarter t, we compute each forecast for all characteristic-sorted portfolios with at least 10 years of past returns, which is 132 to 173 portfolios depending on the date. We then use a cross-sectional regression of future realized quarterly returns $r_{i,t+1}$ on the Bayesian forecast $\hat{r}_{i,t}$ and/or alternative forecasts $\hat{r}^{\text{alt}}_{i,t}$:

$$r_{i,t+1} = a_t + b_t \cdot \hat{r}_{i,t} + c_t \cdot \hat{r}^{\text{alt}}_{i,t} + \epsilon_{i,t+1}. \tag{15}$$

We report the time-series averages and standard errors of b_t and c_t in Panel A of Table 6.

Column (1) of Panel A shows that our model's Bayesian forecasts strongly predict future realized returns. The reported standard errors imply a t-statistic of 5.1 for the no-predictability null and a t-statistic of 1.2 for the one-for-one predictability null. Column (2) shows that the MLE forecasts also predict future returns, but with a coefficient of 0.44 that is well below one. The MLE forecasts are thus too dispersed, overstating cross-sectional differences in expected returns. This pattern is consistent with the forecasts conditioning on noisy point estimates of the model parameters, and illustrates the value of integrating over parameter uncertainty as the Bayesian forecasts do. Columns (3) through (5) indicate that

simple equal-weighted averages of past returns are also strong predictors, particularly when using the full history of returns in Column (3).[29]

The multivariate specifications in Columns (6) through (8) show that the Bayesian forecasts subsume the information in the alternatives. When the Bayesian forecast is included alongside the three past-return averages in Column (6), its coefficient remains near one while the coefficient on the full-sample average return falls to zero, indicating that the model's forecasts combine most of the relevant information in past returns. In contrast, Column (7) shows that the MLE forecast contains essentially no incremental information beyond the past-return averages. Column (8), which includes all five forecasts, confirms that the Bayesian forecast is the dominant predictor. The negative coefficient on past five-year average return $\overline{r}_i$ in Columns (6) and (8) suggests that the Bayesian forecast $\hat{r}_{i,t}$ overweights returns in this window relative to the optimal forecast.

While the Fama-MacBeth regressions establish statistical predictability, trading strategies make the economic magnitudes more transparent. Panel B of Table 6 reports the performance of long-short strategies based on each forecast. At the end of each quarter, we sort the available portfolios by their forecast conditional Sharpe ratios and form an equally-weighted strategy that is long the top and short the bottom quintile, decile, or ventile of portfolios.[30] The strategies based on the Bayesian forecasts earn between 7.16% per year using quintiles and 13.87% per year using ventiles, in each case exceeding the corresponding strategies based on the MLE and past-average forecasts. The final column sorts portfolios on the component of the Bayesian forecast that is orthogonal to all three past-return

[29]The full past sample result, and the past-year result, are similar to the findings in McLean and Pontiff (2016) and Gupta and Kelly (2019), respectively.

[30]For each forecast, we divide the forecast expected return by the corresponding forecast return volatility; for the three past-average forecasts, we use the volatility of all past returns.

averages, isolating the model's information beyond simple averages of past returns. These strategies earn between 2.42% and 4.10% per year, all statistically significant, indicating that the incremental information in the Bayesian forecasts is economically meaningful.

Finally, Panel C of Table 6 evaluates the utility gains that a mean-variance investor could realize from the model's forecasts, in the spirit of Campbell and Thompson (2008). At the end of each quarter t, we form a mean-variance optimal portfolio that combines all available characteristic-sorted portfolios with the market portfolio. Letting $\hat{r}_t$ stack one of the five forecasts $\hat{r}_{i,t}$ and the market's expected return, the portfolio weights solve

$$w_t = \arg\max_{w} \; w'\hat{r}_t - \frac{A}{2} w'\hat{\Sigma}_t w \quad \text{subject to} \quad \mathbf{1}'w = 0 \;\text{ and }\; |w_i| \leq 0.05, \tag{16}$$

where $\hat{\Sigma}_t$ is the estimated covariance matrix and A is the coefficient of risk aversion.[31] The first constraint makes the strategy a zero-net-investment long-short portfolio (we therefore add the realized risk-free rate when reporting total returns); the weight caps limit leverage.

The final column reports the performance of the market-only investor as a benchmark. Using the Bayesian forecasts, the mean-variance portfolio earns an average total return of 13.62% per year with an annualized Sharpe ratio of 1.01, compared to 9.40% and 0.49 for the market-only benchmark. We also compute each strategy's certainty equivalent return, $\overline{r}_p - \frac{A}{2}\widehat{\text{var}}(r_p)$, using the same A. The Bayesian strategy's certainty equivalent is 12.21% per year – 4.5 percentage points above the market-only benchmark, and 0.5 to 2.6 above the

[31] We set $A = 2.997$ so that the optimal weight in the market portfolio is one for an investor combining the market portfolio with a risk-free asset. We estimate $\hat{\Sigma}_t$ using all past quarterly returns, shrinking correlations halfway toward the identity matrix, and require at least 20 past quarters and 20 available portfolios. In every column, the market's expected return is the average of past quarterly market excess returns, so columns differ only in their characteristic-sorted portfolio forecasts.

strategies based on past return averages. Consistent with the Fama-MacBeth results, the MLE forecasts produce a much smaller benefit: their certainty equivalent of 7.77% per year is comparable to that of the market-only benchmark. These calculations indicate that the Bayesian forecasts translate into economically meaningful utility gains.

In summary, our analysis shows that out-of-sample Bayesian return forecasts with priors that allow for time variation in expected returns do in fact predict future returns, subsume the information in both maximum likelihood forecasts and simple averages of past returns, and generate economically meaningful utility gains for investors who use them. This success validates that, for the typical characteristic-sorted portfolio, expected returns are indeed likely to vary over time in a manner captured by our statistical model.

5. Conclusion

This paper proposes a statistical model that allows for time-varying expected returns and applies it to the returns of characteristic-sorted portfolios. Two main takeaways emerge from our analysis. First, accounting for time variation in expected returns materially affects statistical inferences about these portfolios' returns. For the value, investment, and profitability portfolios, the historical data are consistent with starkly different return generating processes. These include both i.i.d. returns with significant positive means, and persistent conditional expected returns that fluctuate around zero unconditional means. The same conclusion holds for many other characteristic-sorted portfolios in the literature. The size portfolio, in contrast, exhibits significant evidence of persistent variation in conditional expected returns, and as such rejects the CAPM despite its near-zero average performance.

Our second takeaway is that, although statistical inference from the historical data is imprecise, the data remain informative for investors. As our Bayesian analyses indicate, investors with agnostic priors, open to the possibility of persistent variation in returns, will form posterior beliefs about the conditional expected returns that fluctuate substantially over time and rise with recent performance. Our findings on out-of-sample forecast accuracy show that investors can indeed benefit from such an approach.

Our results have implications for practitioners as well as for academics who are interested in characteristic-sorted portfolios. Financial institutions now offer a multitude of relatively passive investment products, such as ETFs and mutual funds, that focus on the long-term links between returns and characteristics. At the same time, there exist active hedge funds that attempt to time short-term variations in characteristic return premia. Our framework offers guidance for both groups regarding the extent to which the past performance of these portfolios should be incorporated in forecasts of future performance.

The framework we develop can be extended in a number of ways. For example, while our model allows abnormal returns to fluctuate over relatively long cycles, the process governing returns is stationary. The possibility of a non-stationary process in which abnormal returns converge towards zero as markets become more efficient – or jump towards zero upon the publication of relevant academic research – can be incorporated into a model like ours. Another natural extension would be to incorporate an observable process, for example spreads in short interest or valuation ratios between the long and short sides of each portfolio, that is correlated with the unobserved expected return process. Future research incorporating these or other extensions could shed further light on what causes the time variations in conditional expected returns we study here.

References

Altı, Aydoğan and Sheridan Titman, 2019, A dynamic model of characteristic-based return predictability, *Journal of Finance* 74, 3187–3216.

Arnott, Robert D, Mark Clements, Vitali Kalesnik, and Juhani T Linnainmaa, 2021a, Factor momentum, Available at SSRN 3116974 .

Arnott, Robert D, Campbell R Harvey, Vitali Kalesnik, and Juhani T Linnainmaa, 2021b, Reports of values death may be greatly exaggerated, *Financial Analysts Journal* 77, 44–67.

Avramov, Doron, 2002, Stock return predictability and model uncertainty, *Journal of Financial Economics* 64, 423–458.

Avramov, Doron, 2004, Stock return predictability and asset pricing models, *The Review of Financial Studies* 17, 699–738.

Avramov, Doron, Scott Cederburg, and Katarína Lučivjanská, 2018, Are stocks riskier over the long run? taking cues from economic theory, *Review of Financial Studies* 31, 556–594.

Avramov, Doron, Si Cheng, Amnon Schreiber, and Koby Shemer, 2017, Scaling up market anomalies, *Journal of Investing* 26, 89–105.

Barberis, Nicholas, 2000, Investing for the long run when returns are predictable, *Journal of Finance* 55, 225–264.

Bryzgalova, Svetlana, Jiantao Huang, and Christian Julliard, 2023, Bayesian solutions for the factor zoo: We just ran two quadrillion models, *The Journal of Finance* 78, 487–557.

Cameron, A Colin and Pravin K Trivedi, *Microeconometrics: methods and applications* (Cambridge University Press 2005).

Campbell, John Y and Samuel B Thompson, 2008, Predicting excess stock returns out of sample: Can anything beat the historical average?, *The Review of Financial Studies* 21, 1509–1531.

Chen, Andrew Y and Tom Zimmermann, 2021, Open source cross-sectional asset pricing, *Critical Finance Review, Forthcoming* .

Choi, Ki-Soon, Eric C So, and Charles CY Wang, 2021, Going by the book: Valuation ratios and stock returns, SSRN working paper 3854022 .

Conrad, Jennifer and Gautam Kaul, 1988, Time-variation in expected returns, *Journal of Business* 409–425.

Daniel, Kent, David Hirsleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and overreactions, *Journal of Finance* 53, 1839–1885.

Ehsani, Sina and Juhani T Linnainmaa, 2022, Factor momentum and the momentum factor, *Journal of Finance* 77, 1877–1919.

Eisfeldt, Andrea L, Edward Kim, and Dimitris Papanikolaou, 2022, Intangible value, *Critical Finance Review* 11, 299–332.

Fama, Eugene F and Kenneth R French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.

Fama, Eugene F and Kenneth R French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Fama, Eugene F and Kenneth R French, 2021, The value premium, *Review of Asset Pricing Studies* 11, 105–121.

Fama, Eugene F and James D MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of political economy* 81, 607–636.

Ferson, Wayne E, Sergei Sarkissian, and Timothy T Simin, 2003, Spurious regressions in financial economics?, *Journal of Finance* 58, 1393–1413.

Gârleanu, Nicolae, Leonid Kogan, and Stavros Panageas, 2012, Displacement risk and asset returns, *Journal of Financial Economics* 105, 491–510.

Goncalves, Andrei and Gregory Leonard, 2023, The fundamental-to-market ratio and the value premium decline, *Journal of Financial Economics* 147, 382–405.

Gupta, Tarun and Bryan Kelly, 2019, Factor momentum everywhere, *Journal of Portfolio Management* 45, 13–36.

Harvey, Campbell R, Yan Liu, and Heqing Zhu, 2016, and the cross-section of expected returns, *Review of Financial Studies* 29, 5–68.

Hodrick, Robert J, 1992, Dividend yields and expected stock returns: Alternative procedures for inference and measurement, *The Review of Financial Studies* 5, 357–386.

Israel, Ronen, Kristoffer Laursen, and Scott Richardson, 2020, Is (systematic) value investing dead?, *Journal of Portfolio Management* 47, 38–62.

Johannes, Michael, Arthur Korteweg, and Nicholas Polson, 2014, Sequential learning, predictability, and optimal portfolio returns, *Journal of Finance* 69, 611–644.

Kandel, Shmuel and Robert F Stambaugh, 1996, On the predictability of stock returns: an asset-allocation perspective, *Journal of Finance* 51, 385–424.

Kendall, Maurice G, 1954, Note on bias in the estimation of autocorrelation, *Biometrika* 41, 403–404.

Kogan, Leonid, Dimitris Papanikolaou, and Noah Stoffman, 2020, Left behind: Creative destruction, inequality, and the stock market, *Journal of Political Economy* 128, 855–906.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2020, Shrinking the cross-section, *Journal of Financial Economics* 135, 271–292.

Lewellen, Jonathan, 2002, Momentum and autocorrelation in stock returns, *Review of Financial Studies* 15, 533–564.

Linnainmaa, Juhani T and Michael R Roberts, 2018, The history of the cross-section of stock returns, *Review of Financial Studies* 31, 2606–2649.

Marriott, FHC and JA Pope, 1954, Bias in the estimation of autocorrelations, *Biometrika* 41, 390–402.

McLean, R David and Jeffrey Pontiff, 2016, Does academic research destroy stock return predictability?, *Journal of Finance* 71, 5–32.

Newey, Whitney K and Kenneth D West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1–28.

Pástor, L'uboš, 2000, Portfolio selection and asset pricing models, *Journal of Finance* 55, 179–223.

Pástor, L'uboš and Robert F Stambaugh, 2009, Predictive systems: Living with imperfect predictors, *Journal of Finance* 64, 1583–1628.

Pástor, L'uboš and Robert F Stambaugh, 2012, Are stocks really less volatile in the long run?, *Journal of Finance* 67, 431–478.

Shiller, Robert J, *Irrational exuberance* (Princeton university press 2000).

Smith, Simon and Allan Timmermann, 2022, Have risk premia vanished?, *Journal of Financial Economics* 145, 553–576.

Wachter, Jessica A and Missaka Warusawitharana, 2009, Predictable returns and asset allocation: Should a skeptical investor time the market?, *Journal of Econometrics* 148, 162–178.

White, Halbert, 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* 817–838.

Figure 1: Maximum Likelihood with Alternative Upper Bounds on H

This figure presents p-values from likelihood ratio tests (left axis, black numbered line) for the null hypothesis that the unconditional expected return μ is zero, plotted as a function of H_{ub}, the maximum half-life for shocks to expected returns (H) allowed when estimating the model. For each H_{ub}, we compute the ratio of the maximum likelihood constrained so that $\mu = 0$ and $H \leq H_{ub}$ to the maximum likelihood constrained only by $H \leq H_{ub}$. We compute the p-values for this ratio using small-sample distributions obtained by simulating the model under the restricted parameter estimates. The figure also shows the likelihood of the observed data conditional on H, allowing all other parameters to be re-estimated for each H, scaled by the maximum likelihood value (right axis, grey line). The four panels show estimates for the value, investment, profitability, and size portfolios, which are defined in the header of Table 1. The sample consists of 234 quarterly observations from Q3 1963 through Q4 2021.

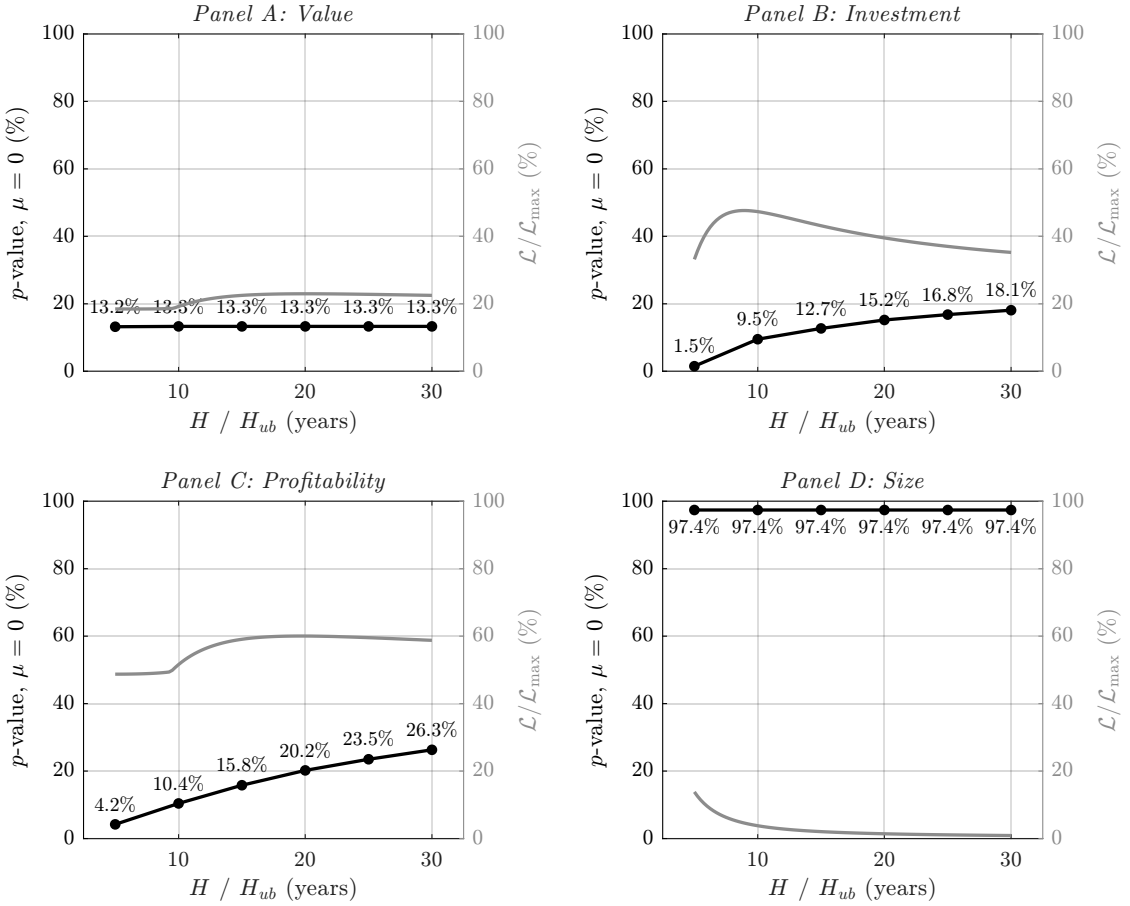


Figure 2: Unconditional Expected Return Inferences for 174 Portfolios

This figure reports the number of characteristic-sorted portfolios for which the null hypothesis that the unconditional expected return μ is zero is rejected at different p-value thresholds. The dark gray bars use OLS with White (1980) standard errors. The medium gray bars use OLS with Newey-West standard errors and 12 lags. The light gray bars use likelihood ratio tests based on the maximum likelihood estimates of our statistical model, as described in Section 3.3.1. We conduct these tests for quarterly returns of 174 long-short equity portfolios, which include the four portfolios that we focus on in our main analysis and 170 portfolios analyzed by Chen and Zimmermann (2021). The sample period varies across portfolios and consists of at most 234 quarterly observations from Q3 1963 through Q4 2021.

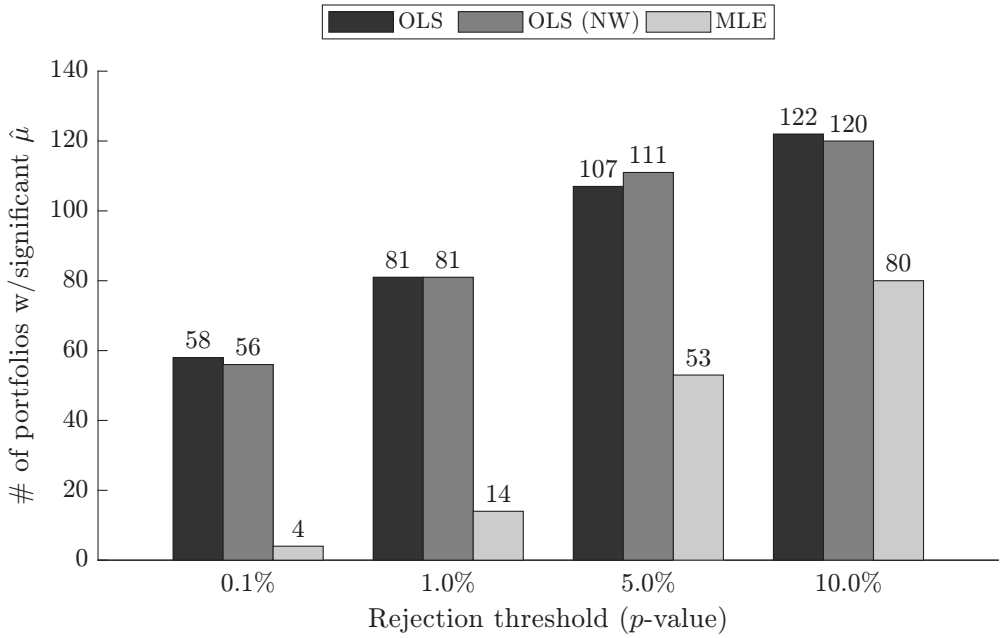
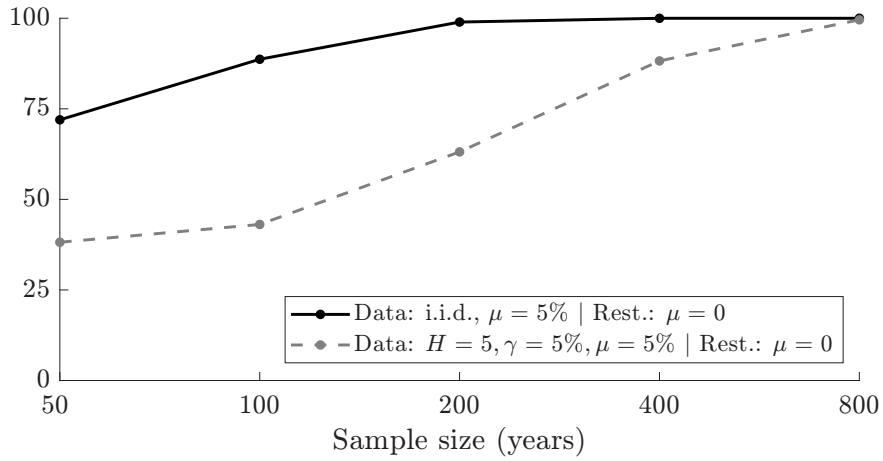


Figure 3: Maximum Likelihood Tests in Simulated Samples

This figure reports the rates at which likelihood ratio tests reject parameter restrictions in simulated samples of varying lengths. Each sample consists of quarterly data simulated under a specific parameterization of our model. For each parameterization of each length we simulate 10,000 samples. In all simulated samples, the annualized return volatility $\sigma_r = 15\%$. The parameter μ is the annualized unconditional expected return, H is the annualized half-life of shocks to expected return, and γ is the one-quarter autocorrelation in returns. The parameters of the process that generates the simulated data are given in the "Data" label for each line. The parameter restrictions for the tested hypothesis are given in the "Rest." label for each line, with any unspecified parameter left unrestricted. The lines in the figure report the fraction of simulated samples, in percent, for which an asymptotic likelihood ratio test rejects the restriction at the 5% level.

Panel A: % of simulated samples rejecting restrictions on μ at the 5% level



Panel B: % of simulated samples rejecting restrictions on H and γ at the 5% level

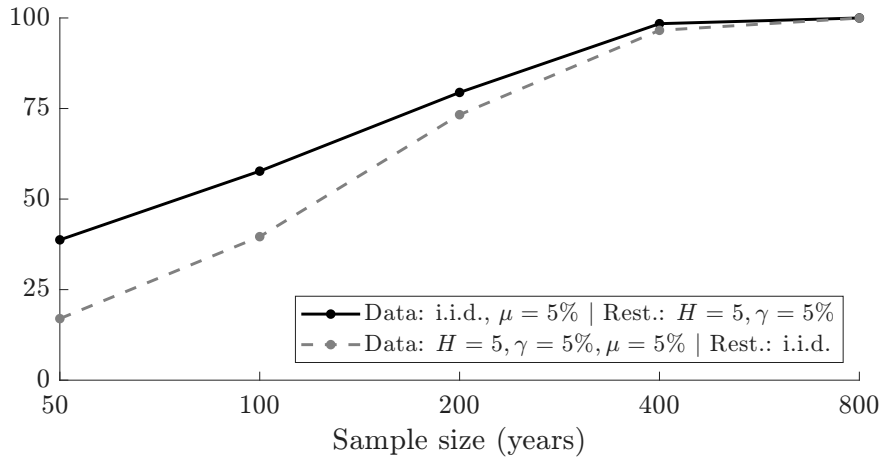
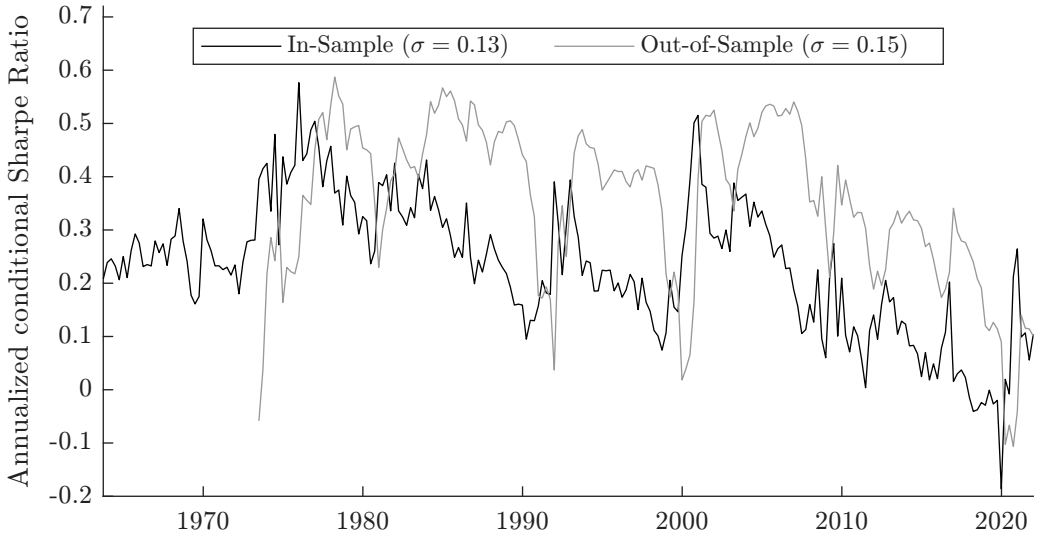


Figure 4: Posteriors for Conditional Sharpe Ratios Across Time

This figure shows means of posterior belief distributions for annualized conditional Sharpe ratios of value, investment, profitability, and size portfolios, as defined in the header of Table 1. Posterior distributions are computed both in-sample (black line) and out-of-sample (grey line). The in-sample approach uses the full sample of historical data and presents posterior beliefs about the conditional Sharpe ratio for each quarter. The out-of-sample approach repeats the Bayesian estimation at the end of each calendar year using only past data, and uses the resulting posteriors along with all past realized returns to compute conditional next-quarter Sharpe ratios for each quarter in the following year. For both approaches, prior beliefs are that the unconditional Sharpe ratio μ_{sr} is distributed $N(0, 0.4)$ and the half-life of shocks to expected returns H is distributed $U(0, 10)$. The legends provide the time-series standard deviation of conditional Sharpe ratios. The sample consists of 234 quarterly observations from Q3 1963 through Q4 2021.

Panel A: Value



Panel B: Investment

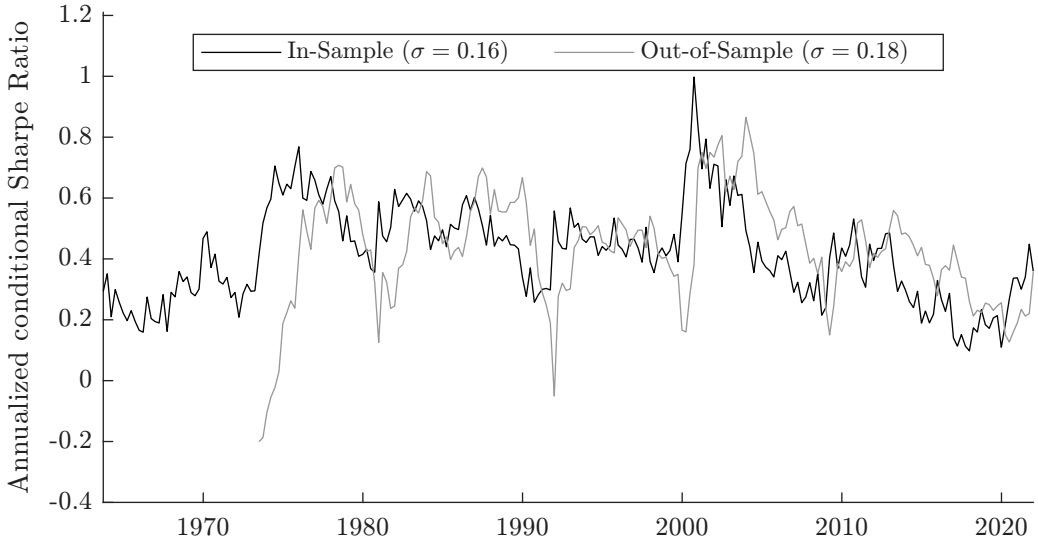
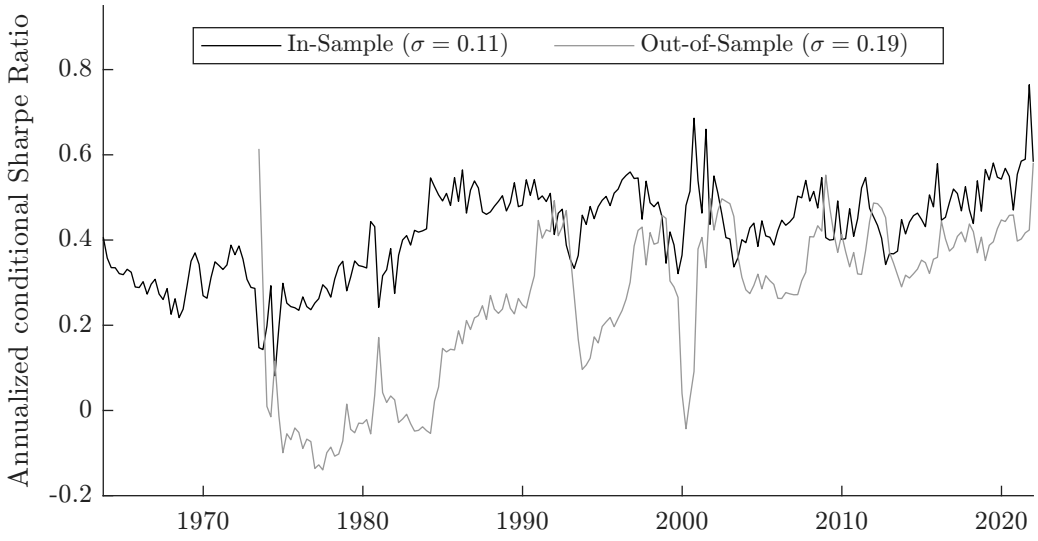


Figure 4: Posteriors for Conditional Sharpe Ratios Across Time (continued)

Panel C: Profitability



Panel D: Size

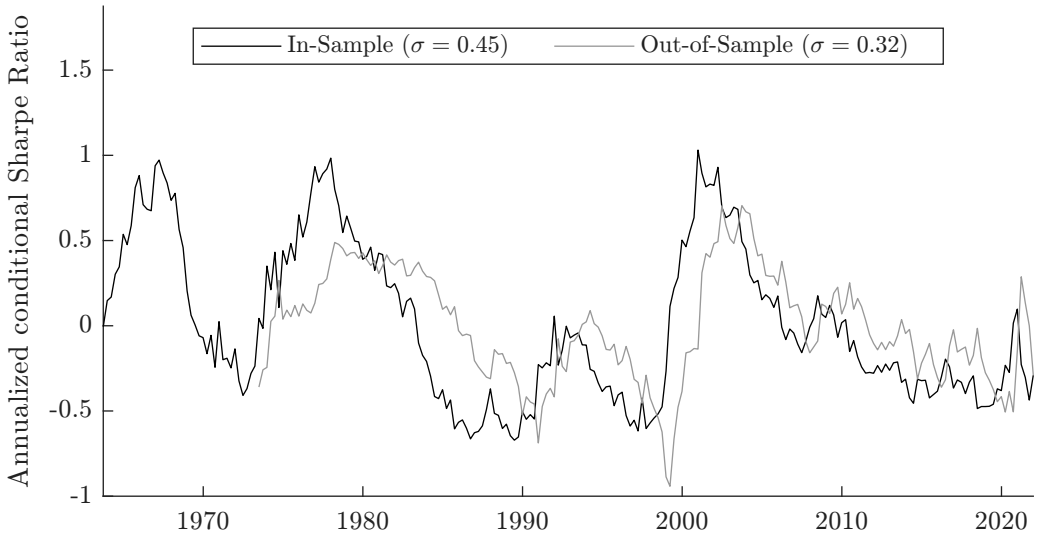


Table 1: Performance and Autocorrelation of Characteristic-Sorted Portfolios

This table presents statistics summarizing the historical performance of value-weighted quintile portfolios formed on value, investment, profitability, and size characteristics. The value portfolio is based on sorting firms by their book-to-market ratios, the investment portfolio on sorting by the annual growth rate of total assets, the profitability portfolio on sorting by operating profits divided by book equity, and the size portfolio on sorting by market capitalization, as in Fama and French (2015). The investment portfolio is long firms in the lowest quintile and short firms in the highest quintile, while the other three portfolios are long the highest quintile and short the lowest. For each portfolio, we compute market-neutral returns by hedging out market risk using the full-sample market β. Panel A reports the mean annualized quarterly return in percentage terms and the annualized Sharpe ratio for the full sample, two subsamples, and the difference between the subsamples. t-statistics based on i.i.d. re-sampling of the calendar quarters in each sample are in parentheses. Panel B reports estimates of overlapping time-series regressions of quarterly returns r_t on a constant and rolling averages of past returns from quarters $t-L$ through $t-1$:

$$r_t = a + b_L \left(\frac{1}{L} \sum_{l=1}^{L} r_{t-l} \right) + \epsilon_t, \tag{17}$$

along with p-values for the $b_L = 0$ null hypothesis based on i.i.d. re-sampling of the calendar months in our sample, reported in the rows labeled i.i.d. p-value. The "Pooled" and "Pooled (without size)" rows show estimates of b_L from panel regressions including all four portfolios, and the three portfolios other than size, respectively. The sample consists of 234 quarterly observations from Q3 1963 through Q4 2021.

Panel A: Portfolio Performance

	Mean (annualized %)				Sharpe Ratio (annualized)			
	All	1963–1992	1993–2021	Diff	All	1963–1992	1993–2021	Diff
Value	3.80	6.09	1.47	-4.62	0.25	0.45	0.09	-0.36
	(1.92)	(2.20)	(0.52)	-(1.17)	(1.91)	(2.40)	(0.47)	-(1.34)
Investment	4.75	4.81	4.70	-0.12	0.45	0.48	0.41	-0.07
	(3.41)	(2.47)	(2.36)	-(0.04)	(3.39)	(2.61)	(2.19)	-(0.26)
Profitability	4.92	3.40	6.45	3.05	0.46	0.36	0.54	0.18
	(3.48)	(1.72)	(3.20)	(1.08)	(3.46)	(1.92)	(2.86)	(0.69)
Size	0.15	1.96	-1.69	-3.65	0.01	0.12	-0.12	-0.24
	(0.07)	(0.69)	-(0.59)	-(0.90)	(0.07)	(0.63)	-(0.63)	-(0.89)

Panel B: Portfolio Autocorrelation Tests

	b_4	b_{10}	b_{20}
Value	0.09	-0.07	0.10
i.i.d. p-value	17.1%	50.2%	21.1%
Investment	0.22	0.07	-0.19
i.i.d. p-value	2.0%	25.5%	55.0%
Profitability	0.05	-0.25	-0.49
i.i.d. p-value	29.2%	78.4%	81.4%
Size	0.34	0.40	0.09
i.i.d. p-value	0.1%	0.3%	21.8%
Pooled	0.20	0.15	0.00
i.i.d. p-value	0.2%	5.1%	26.5%
Pooled (without size)	0.11	-0.07	-0.10
i.i.d. p-value	5.7%	53.5%	45.7%

Table 2: OLS with Time-Varying Expected Returns

This table presents estimates of unconditional expected returns of the value, investment, profitability, and size portfolios, which are defined in the header of Table 1, along with standard errors calculated using a variety of approaches. The first row reports the point estimate of the unconditional expected return, $\hat{\mu}$, in annualized percentage terms. The 'Unadjusted SE' row reports the typical OLS standard error, which is calculated under the assumption of independently distributed returns. The next three rows report standard errors calculated using the Newey-West adjustment with 10, 20, or 40 quarterly lags. The remaining rows report OLS standard errors corrected for autocorrelation using Equation (12) with different values of H, the half-life of shocks to expected returns in years, and γ, the first-order autocorrelation of quarterly returns. The sample consists of 234 quarterly observations from Q3 1963 through Q4 2021.

	Value	Investment	Profitability	Size
$\hat{\mu}$ (annualized %)	3.80	4.75	4.92	0.15
Unadjusted SE	(1.97)	(1.39)	(1.41)	(2.02)
Newey-West SE (10 lags)	(2.04)	(1.58)	(1.42)	(2.89)
Newey-West SE (20 lags)	(2.03)	(1.48)	(1.29)	(3.08)
Newey-West SE (40 lags)	(2.01)	(1.36)	(1.36)	(2.78)
Model-based SE ($H = 0.5$, $\gamma = 2.5\%$)	(2.14)	(1.50)	(1.53)	(2.19)
Model-based SE ($H = 0.5$, $\gamma = 5\%$)	(2.29)	(1.61)	(1.63)	(2.34)
Model-based SE ($H = 0.5$, $\gamma = 10\%$)	(2.56)	(1.80)	(1.83)	(2.62)
Model-based SE ($H = 2.5$, $\gamma = 2.5\%$)	(2.58)	(1.81)	(1.84)	(2.64)
Model-based SE ($H = 2.5$, $\gamma = 5\%$)	(3.06)	(2.15)	(2.19)	(3.13)
Model-based SE ($H = 2.5$, $\gamma = 10\%$)	(3.85)	(2.71)	(2.75)	(3.94)
Model-based SE ($H = 5$, $\gamma = 2.5\%$)	(2.99)	(2.10)	(2.13)	(3.06)
Model-based SE ($H = 5$, $\gamma = 5\%$)	(3.74)	(2.62)	(2.67)	(3.82)
Model-based SE ($H = 10$, $\gamma = 2.5\%$)	(3.54)	(2.48)	(2.53)	(3.62)
Model-based SE ($H = 10$, $\gamma = 5\%$)	(4.60)	(3.23)	(3.28)	(4.70)

Table 3: Downward Bias in Newey-West Standard Errors

This table uses small-sample simulations to demonstrate how the downward bias in Newey-West standard errors and autocorrelation estimates depends on the true data-generating process. Panel A shows how the bias varies as a function of H, the half-life of shocks to conditional expected returns (in years). Except in the i.i.d. case, we assume that γ, the first-order autocorrelation of returns, is 5%. Each of the 250,000 simulations has 234 quarters, zero unconditional expected return μ, and return volatility equal to the volatility of the value portfolio. For each H, we calculate Newey-West standard errors for μ using 1, 3, 5, 10, or 20 years of lags. We show the ratio of the average Newey-West standard error to the simulated standard error, defined as the standard deviation of the point estimate across simulated samples. We also show the fraction of simulated samples in which the t-statistic for $\mu = 0$ is above 1.96. Panel B shows the average and 95% confidence interval for autocorrelation estimates $\hat{b}_4$ and $\hat{b}_{20}$, defined in Table 1, across simulated samples. We also show the true values of b_4 and b_{20} for comparison.

Panel A: Bias in Newey-West Standard Errors

		Number of lags (years)				
H (years)		1	3	5	10	20
i.i.d.	NW SE/Sim. SE	0.99	0.96	0.94	0.89	0.79
	|NW t-stat| > 1.96	5.6%	6.7%	7.8%	10.6%	16.1%
0.5	NW SE/Sim. SE	0.91	0.93	0.92	0.87	0.78
	|NW t-stat| > 1.96	7.8%	7.8%	8.6%	11.2%	16.8%
2.5	NW SE/Sim. SE	0.68	0.74	0.76	0.77	0.72
	|NW t-stat| > 1.96	18.4%	15.9%	15.1%	15.9%	20.6%
5	NW SE/Sim. SE	0.56	0.60	0.63	0.66	0.64
	|NW t-stat| > 1.96	27.8%	24.7%	23.2%	22.6%	25.9%
10	NW SE/Sim. SE	0.44	0.48	0.50	0.52	0.52
	|NW t-stat| > 1.96	38.4%	35.6%	34.0%	32.6%	34.4%

Panel B: Bias in Autocorrelation Estimates b_4 and b_{20}

H (years)	True b_4	Avg $\hat{b}_4$	95% CI $\hat{b}_4$	True b_{20}	Avg $\hat{b}_{20}$	95% CI $\hat{b}_{20}$
i.i.d.	0.00	-0.03	[-0.28,0.20]	0.00	-0.16	[-0.88,0.34]
0.5	0.11	0.08	[-0.16,0.30]	0.17	0.04	[-0.58,0.46]
2.5	0.14	0.10	[-0.17,0.33]	0.34	0.18	[-0.45,0.59]
5	0.14	0.09	[-0.18,0.33]	0.39	0.19	[-0.47,0.62]
10	0.15	0.07	[-0.20,0.31]	0.42	0.15	[-0.54,0.61]

Table 4: Maximum Likelihood Hypothesis Tests

This table presents parameter estimates and hypothesis tests based on maximum-likelihood estimations of our model for the value, investment, profitability, and size portfolios, which are defined in the header of Table 1. Panel A reports estimates of μ, the unconditional expected return in annualized percentage terms; σ_r, the volatility of returns in annualized percentage terms; H, the half-life of shocks to expected returns in years; and γ, the first-order autocorrelation of quarterly returns. The rows labeled $\mu = 0$ re-estimate the model with μ restricted to be zero, and report the p-value for this restriction based on a likelihood ratio test. Panel B presents a variety of hypothesis tests for different restrictions on H and γ. In each case, we report the likelihood ratio p-value of the (H, γ) restriction. We compute p-values for likelihood ratios using small-sample distributions computed by simulating the model under the restricted parameter estimates. The sample consists of 234 quarterly observations from Q3 1963 through Q4 2021.

Panel A: Tests for $\mu = 0$

	i.i.d.			Time-varying expected returns				
	μ (%)	σ_r (%)	$\mu = 0$ p-value	μ (%)	σ_r (%)	H (years)	γ (%)	$\mu = 0$ p-value
Value	3.80	15.10		3.77	15.10	0.14	11.82	
$\mu = 0$	0.00	15.22	5.5%	0.00	15.22	0.17	13.32	13.3%
Investment	4.75	10.60		4.78	10.60	0.34	9.21	
$\mu = 0$	0.00	10.87	0.0%	0.00	10.80	10.00	4.81	9.5%
Profitability	4.92	10.78		4.98	10.78	0.21	7.65	
$\mu = 0$	0.00	11.05	0.0%	0.00	10.98	10.00	4.31	10.4%
Size	0.15	15.44		-0.13	15.43	1.23	14.03	
$\mu = 0$	0.00	15.44	94.3%	0.00	15.43	1.23	14.02	97.4%

Panel B: Hypothesis Tests for H and γ (p-values)

		$H = 2.5$		$H = 5$		$H = 10$	
	i.i.d.	$\gamma = 2.5\%$	$\gamma = 5\%$	$\gamma = 2.5\%$	$\gamma = 5\%$	$\gamma = 2.5\%$	$\gamma = 5\%$
Value	34.0%	27.5%	17.5%	26.5%	16.9%	29.2%	20.1%
Investment	45.7%	41.9%	32.9%	35.0%	24.3%	33.3%	23.7%
Profitability	74.8%	48.9%	29.6%	49.1%	29.8%	55.0%	36.2%
Size	1.2%	8.9%	21.4%	5.6%	10.6%	3.4%	5.3%

Table 5: Bayesian Prior and Posterior Beliefs

This table presents summary statistics for the prior and posterior distributions of key moments from our model for the value, investment, profitability, and size portfolios, which are defined in the header of Table 1. We consider six possible priors, each differing only in its belief about H, the half-life of shocks to expected returns in years: that it equals 0, 2.5, or 5 years; or that it is uniformly distributed as $U(0, 10)$, $U(0, 20)$, or $U(0, 30)$ years. Panel A shows the mean and 95% confidence interval for prior beliefs about μ_{sr}, the annualized unconditional Sharpe ratio; H; γ, the first-order autocorrelation of quarterly returns; and $\hat{\sigma}_{sr}$, the volatility of the Sharpe ratio conditional on past realized returns. Panels B through E show means and 95% confidence intervals for posterior beliefs about the same moments. We compute the posterior beliefs using an MCMC procedure described in Appendix C. Our sample consists of 234 quarterly observations from Q3 1963 through Q4 2021.

Panel A: Prior Beliefs

H Prior	μ_{sr}		H		γ (%)		$\hat{\sigma}_{sr}$	
0	0.00	[-0.78,0.78]	0.00	-	-	-	-	-
2.5	0.00	[-0.78,0.78]	2.50	-	4.79	[-0.59,15.44]	0.19	[0.00,0.52]
5	0.00	[-0.78,0.78]	5.00	-	5.59	[-0.26,16.52]	0.25	[0.00,0.61]
$U(0, 10)$	0.00	[-0.78,0.78]	5.00	[0.25,9.75]	4.92	[-1.82,16.42]	0.23	[0.00,0.63]
$U(0, 20)$	0.00	[-0.78,0.78]	10.00	[0.50,19.50]	5.62	[-0.65,17.17]	0.28	[0.00,0.69]
$U(0, 30)$	0.00	[-0.78,0.78]	15.00	[0.75,29.25]	5.93	[-0.34,17.50]	0.30	[0.00,0.72]

Panel B: Posterior Beliefs for Value

H Prior	μ_{sr}		H		γ (%)		$\hat{\sigma}_{sr}$	
0	0.23	[-0.01,0.48]	0.00	-	-	-	-	-
2.5	0.21	[-0.12,0.52]	2.50	-	2.74	[-0.60,11.66]	0.12	[0.00,0.42]
5	0.20	[-0.19,0.54]	5.00	-	2.75	[-0.28,11.85]	0.14	[0.00,0.49]
$U(0, 10)$	0.20	[-0.18,0.53]	4.66	[0.29,9.74]	2.83	[-0.84,11.80]	0.14	[0.00,0.48]
$U(0, 20)$	0.18	[-0.27,0.55]	9.71	[0.44,19.51]	3.02	[-0.45,12.40]	0.17	[0.00,0.55]
$U(0, 30)$	0.17	[-0.31,0.56]	14.84	[0.60,29.29]	3.19	[-0.27,12.85]	0.19	[0.00,0.59]

Panel C: Posterior Beliefs for Investment

H Prior	μ_{sr}		H		γ (%)		$\hat{\sigma}_{sr}$	
0	0.41	[0.17,0.66]	0.00	-	-	-	-	-
2.5	0.37	[-0.01,0.70]	2.50	-	4.05	[-0.52,13.49]	0.17	[0.00,0.47]
5	0.34	[-0.10,0.70]	5.00	-	3.58	[-0.27,13.72]	0.18	[0.00,0.54]
$U(0, 10)$	0.36	[-0.08,0.70]	4.33	[0.33,9.68]	3.83	[-0.61,13.36]	0.17	[0.00,0.52]
$U(0, 20)$	0.32	[-0.18,0.70]	8.92	[0.47,19.42]	3.77	[-0.34,13.76]	0.19	[0.00,0.58]
$U(0, 30)$	0.31	[-0.24,0.70]	13.71	[0.59,29.18]	3.75	[-0.22,13.90]	0.21	[0.00,0.61]

Table 5: Bayesian Prior and Posterior Beliefs (continued)

Panel D: Posterior Beliefs for Profitability

H Prior	μ_{sr}		H		γ (%)		$\hat{\sigma}_{sr}$	
0	0.42	[0.18,0.67]	0.00	-	-	-	-	-
2.5	0.40	[0.08,0.71]	2.50	-	2.21	[-0.63,10.89]	0.10	[0.00,0.40]
5	0.39	[0.02,0.72]	5.00	-	2.37	[-0.29,11.16]	0.13	[0.00,0.47]
$U(0,10)$	0.39	[0.01,0.71]	4.75	[0.28,9.75]	2.38	[-1.15,11.23]	0.12	[0.00,0.46]
$U(0,20)$	0.36	[-0.09,0.73]	10.03	[0.46,19.53]	2.75	[-0.58,11.99]	0.16	[0.00,0.54]
$U(0,30)$	0.34	[-0.16,0.74]	15.39	[0.65,29.33]	3.04	[-0.34,12.62]	0.19	[0.00,0.59]

Panel E: Posterior Beliefs for Size

H Prior	μ_{sr}		H		γ (%)		$\hat{\sigma}_{sr}$	
0	0.01	[-0.23,0.25]	0.00	-	-	-	-	-
2.5	-0.00	[-0.45,0.45]	2.50	-	10.64	[2.94,16.86]	0.39	[0.14,0.56]
5	-0.00	[-0.52,0.51]	5.00	-	10.69	[1.64,17.52]	0.44	[0.11,0.63]
$U(0,10)$	-0.00	[-0.47,0.46]	3.32	[0.76,9.01]	10.33	[2.07,16.96]	0.38	[0.10,0.62]
$U(0,20)$	0.00	[-0.48,0.47]	4.65	[0.79,17.17]	10.19	[1.32,17.05]	0.39	[0.09,0.65]
$U(0,30)$	0.00	[-0.48,0.48]	5.96	[0.77,25.64]	9.98	[0.86,17.13]	0.39	[0.07,0.66]

Table 6: Out-of-Sample Bayesian Forecasts

This table presents evidence on the out-of-sample forecasting performance of Bayesian forecasts of quarterly characteristic-sorted portfolio returns, along with several benchmarks. Panel A shows results from Fama-MacBeth regressions of next-quarter portfolio returns ($r_{i,t+1}$) on forecasts derived from our Bayesian framework (Bayesian $\hat{r}_{i,t}$); maximum likelihood estimations of our model (MLE $\hat{r}_{i,t}$); the average return in the full prior sample (All past $\overline{r}_i$); the average return in the past five years (Past five years $\overline{r}_i$); and the average return in the past year (Past year $\overline{r}_i$). Each of these forecasts uses only past return data for the portfolio in question, made market neutral using past market portfolio returns. Realized next-quarter returns are net of the past beta times the realized market return. Panel B presents average returns for portfolios combining long positions in the characteristic-sorted portfolios predicted to have the best future returns with short positions in those predicted to have the worst. It additionally includes predictions from our Bayesian framework after orthogonalizing within each period with respect to the three past return measures (Bayesian $\hat{r}^{\perp}_{i,t}$). Panel C presents the mean return, annualized Sharpe ratio, and certainty equivalent of mean-variance (MV) optimal portfolios that combine all available characteristic-sorted portfolios together with the market as an additional asset. Each column forms the MV-optimal portfolio from one of the five return forecasts, and the final column reports a market-only benchmark. The mean return and certainty equivalent are computed on total returns (adding back the realized risk-free rate), while the Sharpe ratio is computed on returns in excess of the risk-free rate. Standard errors are in parentheses in Panels A and B. The sample period is 194 quarters from Q3 of 1973 through Q4 of 2021, with the number of available characteristic-sorted portfolios varying from 132 to 173.

Panel A: Fama-MacBeth Regressions of $r_{i,t+1}$ on Forecasts

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bayesian $\hat{r}_{i,t}$	0.81***	–	–	–	–	1.04***	–	1.16***
	(0.16)					(0.39)		(0.44)
MLE $\hat{r}_{i,t}$	–	0.44***	–	–	–	–	0.06	-0.10
		(0.08)					(0.06)	(0.07)
All past $\overline{r}_i$	–	–	0.71***	–	–	0.01	0.54***	0.04
			(0.12)			(0.21)	(0.12)	(0.21)
Past five years $\overline{r}_i$	–	–	–	0.30***	–	-0.25*	-0.04	-0.27*
				(0.09)		(0.14)	(0.11)	(0.14)
Past year $\overline{r}_i$	–	–	–	–	0.20***	0.06	0.11**	0.05
					(0.05)	(0.05)	(0.05)	(0.05)

Panel B: Cross Characteristic Long-Short Strategy Returns (%/yr)

	Bayesian $\hat{r}_{i,t}$	MLE $\hat{r}_{i,t}$	All past $\overline{r}_i$	Past 5y $\overline{r}_i$	Past 1y $\overline{r}_i$	Bayesian $\hat{r}^{\perp}_{i,t}$
Quintiles	7.16***	5.77***	7.01***	4.23**	4.97**	2.42**
	(1.81)	(1.21)	(1.45)	(1.96)	(2.03)	(1.16)
Deciles	10.20***	7.58***	9.84***	6.28**	5.83**	3.68***
	(2.27)	(1.62)	(1.56)	(2.50)	(2.57)	(1.39)
Ventiles	13.87***	8.77***	12.19***	9.89***	6.66**	4.10**
	(2.55)	(2.23)	(1.97)	(2.84)	(3.22)	(1.68)

Table 6: Out-of-Sample Bayesian Forecasts (continued)

Panel C: Cross Characteristic MV-Optimal Strategy Performance

	Market + characteristic-sorted portfolios					Market only
Expected return:	Bayesian $\hat{r}_{i,t}$	MLE $\hat{r}_{i,t}$	All past $\overline{r}_i$	Past 5y $\overline{r}_i$	Past 1y $\overline{r}_i$	
Mean (%/yr)	13.62	8.95	12.91	11.83	12.80	9.40
Sharpe ratio (ann.)	1.01	0.56	1.02	0.71	0.61	0.49
Certainty equiv. (%/yr)	12.21	7.77	11.72	9.97	9.66	7.71

Appendix A. Autocorrelations in Monthly Returns

As described in Section 2.1, we analyze quarterly returns of characteristic-sorted portfolios rather than the monthly returns that are typically studied in the literature. We do so because monthly returns exhibit strong first-order autocorrelations, which may be caused by lead-lag effects, under-reaction, or some other transitory source of persistence. While interesting on its own, this form of autocorrelation differs from the main focus of this paper, namely, slow-moving but persistent variations in expected returns. To avoid biasing our estimates towards large but quickly-reverting variations, we use quarterly rather than monthly data.

Appendix Figure 1 illustrates the magnitude of the autocorrelations at monthly lags $l = 1$ through $l = 60$ for the four portfolios we study, as well as the autocorrelations estimated in a pooled regression including all four portfolios. The first-order autocorrelation is the single largest coefficient for any of the 60 months for value, investment, and profitability, as well as in pooled regressions. Quarterly data do not exhibit a strong first-order autocorrelation (see Table 1) because the two- and three-month autocorrelations in Appendix Figure 1 are much smaller and statistically insignificant.

Appendix B. Generalized Least Squares (GLS) Estimations

In addition to the OLS standard error correction described in Section 3.1, we estimate unconditional expected returns μ in GLS regressions that use the covariance matrix Σ implied by γ and H. These GLS estimations adjust both the point estimates $\hat{\mu}$ and the standard errors. While the OLS estimates are based on an equally-weighted average of the returns in the sample, the GLS estimates utilize an average weighted by the amount of orthogonal information each observation contains about the unconditional expected return. When conditional expected returns exhibit persistent variations, the observations in the middle of the sample are relatively more redundant because they 'over-sample' the same epoch of conditional expected returns. As illustrated by Appendix Figure 2, GLS therefore over-weights observations at the beginning and end of the sample. This effect is larger for larger values of H, and reverses when γ is negative.

Appendix Table 2 shows the GLS point estimates and t-statistics, along with the corresponding OLS statistics, for a variety of assumptions about γ and H. We find that the GLS corrections to point estimates are generally small, but somewhat more negative in some specifications for the value and the investment portfolios. The reason for this negative effect is that these portfolios had unusually low returns at the beginning and/or the end of the sample, which GLS infers as providing more independent information relative to the observations in the middle. Overall, the main conclusion from the OLS analysis, that t-statistics can be half as large for reasonable values of H and γ relative to uncorrected t-statistics, remains unchanged when using GLS.

Appendix C. Sampling Bayesian Posteriors

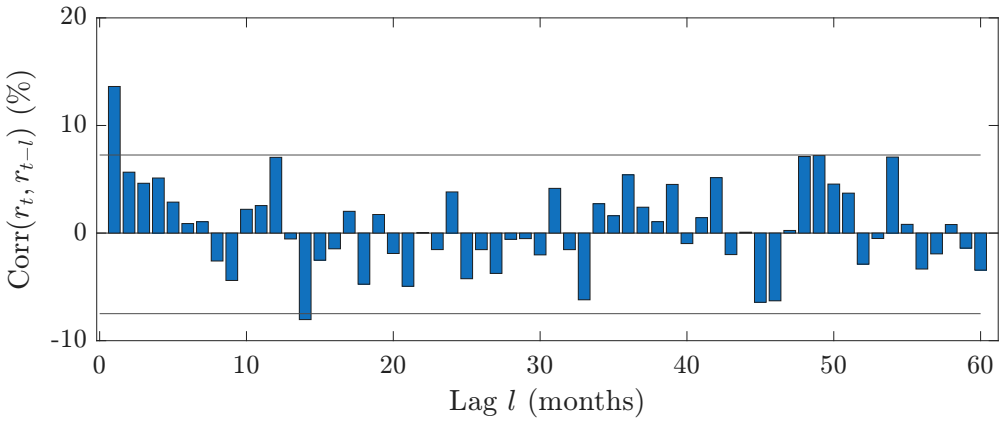
We draw samples of $N = 250,000$ observations from the posterior distribution of model parameters $\Omega^{\text{posterior}}$ using the following procedure:

1. Draw N observations Ω_i^{prior}, $i \in [1, N]$ from the prior distribution.
2. Accept Ω_1^{prior} as the first observation of the posterior distribution $\Omega_1^{\text{posterior}}$.
3. For observations $i = 2 \dots N$:
   (a) Evaluate the conditional likelihood of the data D given the ith draw from the prior parameters as well as the $(i-1)$th draw from the posterior parameters:
   $$\mathcal{L}^{\text{propose}} = \mathcal{L}(D|\Omega = \Omega_i^{\text{prior}}), \tag{18}$$
   $$\mathcal{L}^{\text{previous}} = \mathcal{L}(D|\Omega = \Omega_{i-1}^{\text{posterior}}). \tag{19}$$
   (b) If $\mathcal{L}^{\text{propose}} \geq \mathcal{L}^{\text{previous}}$, accept $\Omega_i^{\text{posterior}} = \Omega_i^{\text{prior}}$.
   (c) If $\mathcal{L}^{\text{propose}} < \mathcal{L}^{\text{previous}}$, accept $\Omega_i^{\text{posterior}} = \Omega_i^{\text{prior}}$ with probability $\frac{\mathcal{L}^{\text{propose}}}{\mathcal{L}^{\text{previous}}}$, and otherwise retain $\Omega_i^{\text{posterior}} = \Omega_{i-1}^{\text{posterior}}$.

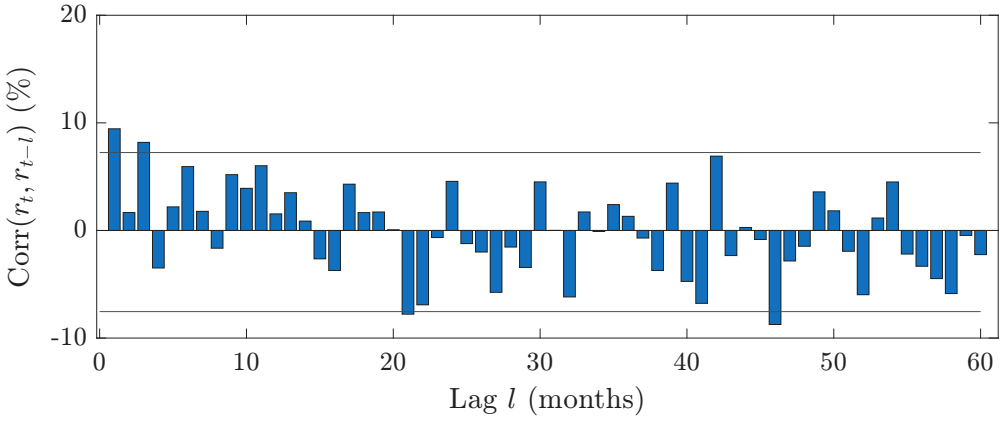
Appendix Figure 1: Monthly Autocorrelograms

This figure presents autocorrelations of monthly returns for value, investment, profitability, and size portfolios, as defined in the header of Table 1, as well as the autocorrelation estimated in a pooled regression containing all four portfolios. We estimate the autocorrelation for each lag l independently. The horizontal lines represent the 95% confidence interval for autocorrelation coefficients under the zero-autocorrelation null hypothesis. The sample consists of 702 monthly observations from Q3 1963 through Q4 2021.

Panel A: Value

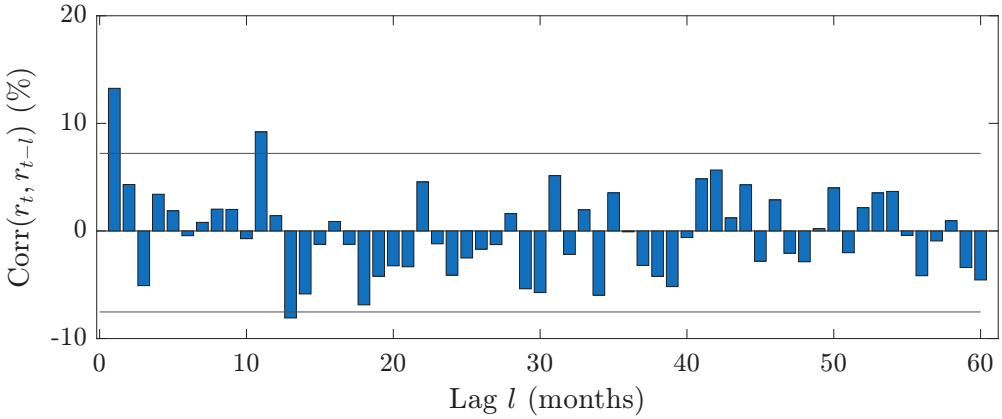


Panel B: Investment

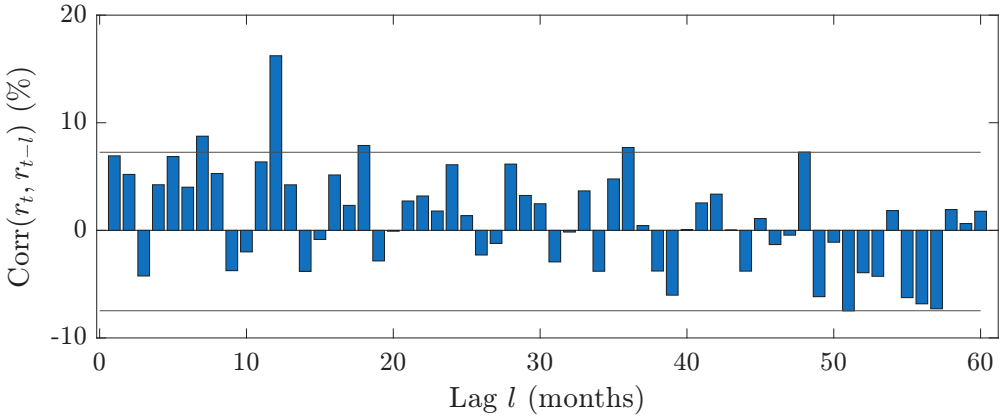


Appendix Figure 1: Monthly Autocorrelograms (continued)

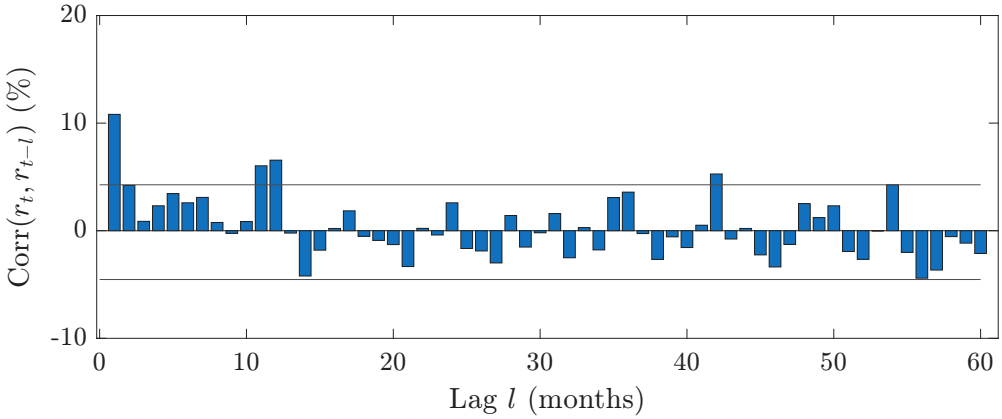
Panel C: Profitability



Panel D: Size

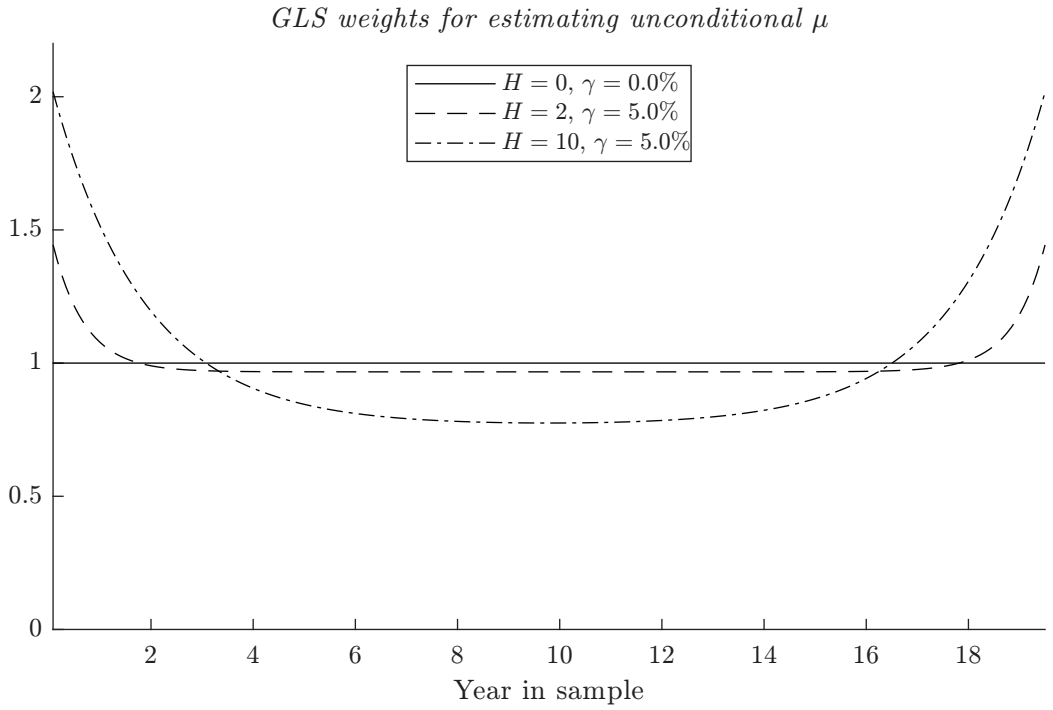


Panel E: Pooled



Appendix Figure 2: GLS Influence Functions

This figure presents the GLS weights that apply in the estimation of the unconditional expected return μ under different assumptions about the autocorrelation structure of returns. The parameter H is the half-life of shocks to expected returns in years, and γ is the first-order autocorrelation of quarterly returns.



Appendix Table 1: Return Autocorrelations for Different Parameterizations

This table shows the first-order autocorrelation of quarterly returns in percentage terms, $\gamma = \text{corr}(r_t, r_{t-1})$, implied by various combinations of H, the half-life of shocks to expected returns; $\sigma_\epsilon/\sigma_\delta$, the ratio of the volatilities of unexpected and expected return shocks; and ρ, the correlation of unexpected and expected return shocks.

Panel A: $\rho = -0.1$

H (years)	$\sigma_\varepsilon/\sigma_\delta$						
	1	2	4	10	20	100	1000
0.5	31.8	11.3	2.5	0.0	-0.2	-0.1	-0.0
2.5	80.0	56.8	26.0	5.0	1.1	-0.0	-0.0
5	89.5	74.0	43.7	11.0	2.8	0.0	-0.0
10	94.6	85.5	61.9	20.9	6.0	0.2	-0.0
25	97.8	93.8	80.7	40.9	14.7	0.6	-0.0

Panel B: $\rho = -0.5$

H (years)	$\sigma_\varepsilon/\sigma_\delta$						
	1	2	4	10	20	100	1000
0.5	17.7	0.0	-4.2	-2.8	-1.6	-0.3	-0.0
2.5	75.2	49.8	19.4	1.5	-0.8	-0.4	-0.0
5	86.9	69.7	38.5	7.6	0.9	-0.3	-0.0
10	93.3	83.1	58.4	17.9	4.2	-0.2	-0.0
25	97.3	92.7	78.9	38.6	13.0	0.2	-0.0

Panel C: $\rho = -1$

H (years)	$\sigma_\varepsilon/\sigma_\delta$						
	1	2	4	10	20	100	1000
0.5	0.0	-14.1	-12.5	-6.3	-3.4	-0.7	-0.1
2.5	69.1	41.1	11.2	-2.9	-3.0	-0.9	-0.1
5	83.7	64.3	32.1	3.3	-1.4	-0.8	-0.1
10	91.6	80.1	54.0	14.1	1.9	-0.7	-0.1
25	96.6	91.4	76.6	35.7	10.9	-0.3	-0.1

Appendix Table 2: GLS with Time-Varying Expected Returns

This table presents estimates of unconditional expected returns of value, investment, profitability, and size portfolios, as defined in the header of Table 1, under a variety of assumptions about the magnitude and persistence of variations in conditional expected returns. The model-implied autocorrelation structure of returns is summarized by H, the half-life of shocks to expected returns in years, and γ, the first-order autocorrelation of quarterly returns. We estimate unconditional expected returns μ using both OLS and GLS and calculate t-statistics using the model-implied correlation matrix of the regression error terms. The sample consists of 234 quarterly observations from Q3 1963 through Q4 2021.

Panel A: Value

	H (years)	2.5			5		10	
	γ (%)	2.5	5	10	2.5	5	2.5	5
OLS	$\hat{\mu}$ (%)	3.80	3.80	3.80	3.80	3.80	3.80	3.80
	Model t-stat	(1.47)	(1.24)	(0.99)	(1.27)	(1.02)	(1.07)	(0.83)
GLS	$\hat{\mu}$ (%)	3.68	3.62	3.57	3.50	3.36	3.28	3.03
	Model t-stat	(1.43)	(1.19)	(0.93)	(1.18)	(0.91)	(0.94)	(0.67)

Panel B: Investment

	H (years)	2.5			5		10	
	γ (%)	2.5	5	10	2.5	5	2.5	5
OLS	$\hat{\mu}$ (%)	4.75	4.75	4.75	4.75	4.75	4.75	4.75
	Model t-stat	(2.63)	(2.21)	(1.76)	(2.26)	(1.81)	(1.91)	(1.47)
GLS	$\hat{\mu}$ (%)	4.65	4.63	4.64	4.47	4.36	4.23	4.00
	Model t-stat	(2.57)	(2.16)	(1.72)	(2.14)	(1.68)	(1.72)	(1.26)

Panel C: Profitability

	H (years)	2.5			5		10	
	γ (%)	2.5	5	10	2.5	5	2.5	5
OLS	$\hat{\mu}$ (%)	4.92	4.92	4.92	4.92	4.92	4.92	4.92
	Model t-stat	(2.67)	(2.25)	(1.79)	(2.30)	(1.84)	(1.95)	(1.50)
GLS	$\hat{\mu}$ (%)	5.05	5.16	5.37	5.07	5.24	5.05	5.23
	Model t-stat	(2.75)	(2.37)	(1.96)	(2.39)	(1.98)	(2.02)	(1.62)

Panel D: Size

	H (years)	2.5			5		10	
	γ (%)	2.5	5	10	2.5	5	2.5	5
OLS	$\hat{\mu}$ (%)	0.15	0.15	0.15	0.15	0.15	0.15	0.15
	Model t-stat	(0.06)	(0.05)	(0.04)	(0.05)	(0.04)	(0.04)	(0.03)
GLS	$\hat{\mu}$ (%)	0.11	0.06	-0.06	0.14	0.10	0.22	0.18
	Model t-stat	(0.04)	(0.02)	-(0.01)	(0.05)	(0.03)	(0.06)	(0.04)

Appendix Table 3: Markov Chain Monte Carlo Mixing Diagnostics

This table reports mixing diagnostics for the Markov chain Monte Carlo (MCMC) procedure described in Appendix C, which we use to compute the posterior beliefs reported in Table 5. Each panel reports diagnostics for one of the value, investment, profitability, and size portfolios, as defined in the header of Table 1, and each row corresponds to one of the six priors for H considered in Table 5. The first column reports the acceptance rate of the sampler. The remaining columns report the effective sample size (ESS) of the posterior draws for μ_{sr}, the unconditional Sharpe ratio; H, the half-life of shocks to expected returns; γ, the first-order autocorrelation of quarterly returns; and $\hat{\sigma}_{sr}$, the volatility of the Sharpe ratio conditional on past realized returns. A dash marks a parameter that is constant under the corresponding prior. Our sample consists of 234 quarterly observations from Q3 1963 through Q4 2021.

Panel A: Value

H Prior	Accept (%)	ESS(μ_{sr})	ESS(H)	ESS(γ)	ESS($\hat{\sigma}_{sr}$)
0	26.9	46,638	-	-	-
2.5	33.2	64,456	-	50,474	48,082
5	33.5	64,112	-	47,133	43,794
$U(0, 10)$	32.9	60,552	36,399	46,133	46,681
$U(0, 20)$	35.8	67,211	40,158	49,902	46,598
$U(0, 30)$	37.8	70,501	41,834	49,843	45,937

Panel B: Investment

H Prior	Accept (%)	ESS(μ_{sr})	ESS(H)	ESS(γ)	ESS($\hat{\sigma}_{sr}$)
0	18.5	29,212	-	-	-
2.5	29.2	47,680	-	42,015	40,807
5	30.1	46,853	-	36,776	34,130
$U(0, 10)$	26.7	37,686	24,295	32,408	35,133
$U(0, 20)$	29.8	41,810	24,699	33,435	34,033
$U(0, 30)$	31.9	43,505	26,979	35,367	33,983

Panel C: Profitability

H Prior	Accept (%)	ESS(μ_{sr})	ESS(H)	ESS(γ)	ESS($\hat{\sigma}_{sr}$)
0	18.1	29,114	-	-	-
2.5	21.2	34,607	-	28,392	27,413
5	22.0	34,791	-	28,403	26,091
$U(0, 10)$	22.3	34,041	22,689	29,928	29,006
$U(0, 20)$	25.3	36,493	26,507	31,130	28,404
$U(0, 30)$	27.8	40,123	29,588	34,711	30,888

Panel D: Size

H Prior	Accept (%)	ESS(μ_{sr})	ESS(H)	ESS(γ)	ESS($\hat{\sigma}_{sr}$)
0	32.2	60,758	-	-	-
2.5	27.8	46,688	-	34,875	36,328
5	38.4	66,214	-	50,380	51,247
$U(0, 10)$	21.5	27,280	18,102	20,986	27,660
$U(0, 20)$	20.2	18,502	14,000	14,134	19,019
$U(0, 30)$	20.2	15,227	11,685	11,952	17,015

Appendix Table 4: Inference and Posteriors for 174 Portfolios

This table presents summary statistics and estimation results for the four portfolios that we focus on in our main analysis (the first four rows) and 170 long-short equity portfolios analyzed by Chen and Zimmermann (2021) (the remaining rows). The first two columns report $E(r_t)$, the in-sample average of annualized quarterly returns, and SR, the in-sample annualized Sharpe ratio. The next two columns report the p-values for the hypothesis that the unconditional expected return μ equals zero using OLS regressions with Newey-West standard errors, and maximum likelihood estimations (MLE) that allow for time-varying expected returns. The next three columns report the means of Bayesian posterior distributions for μ_{sr}, the unconditional Sharpe ratio; $\mu_{sr,2022}$, the conditional Sharpe ratio for Q1 of 2022; and $\hat{\sigma}_{sr}$, the volatility of the Sharpe ratio conditional on past returns.

			$\mu = 0$ p-values		Bayesian Posterior Means		
Portfolio	$E(r_t)$	SR	OLS (NW)	MLE	μ_{sr}	$\mu_{sr,2022}$	$\hat{\sigma}_{sr}$
Value (FF)	3.80	0.25	5.8%	13.3%	0.20	0.11	0.14
Investment (FF)	4.75	0.45	0.2%	9.5%	0.36	0.37	0.17
Profitability (FF)	4.92	0.46	0.0%	10.4%	0.39	0.60	0.12
Size (FF)	0.15	0.01	96.0%	97.4%	-0.00	-0.31	0.38
AM	1.34	0.08	54.8%	63.8%	0.06	-0.01	0.12
AOP	3.90	0.35	2.2%	7.0%	0.28	0.39	0.12
AbnormalAccruals	4.92	0.53	0.0%	0.0%	0.46	0.53	0.08
Accruals	4.60	0.47	0.0%	0.0%	0.42	0.43	0.05
Activism1	-0.60	-0.10	71.3%	73.4%	-0.07	-0.13	0.19
AdExp	5.79	0.38	2.1%	4.9%	0.27	0.13	0.23
AgeIPO	13.80	0.59	0.0%	0.3%	0.46	0.51	0.12
AnalystRevision	5.40	0.57	0.0%	10.9%	0.47	0.57	0.13
AnalystValue	5.39	0.32	3.9%	9.0%	0.26	0.28	0.15
AnnouncementReturn	8.37	1.05	0.0%	1.1%	0.80	1.00	0.18
AssetGrowth	4.39	0.40	0.9%	18.2%	0.29	0.24	0.21
BM	2.29	0.11	42.2%	57.1%	0.08	0.00	0.12
BMdec	3.78	0.27	8.1%	13.2%	0.18	-0.07	0.25
BPEBM	1.75	0.25	3.5%	17.4%	0.23	0.32	0.10
Beta	-7.18	-0.36	1.6%	4.3%	-0.29	-0.50	0.19
BetaFP	-7.74	-0.53	0.0%	9.1%	-0.42	-0.42	0.16
BetaLiquidityPS	1.56	0.16	24.4%	36.6%	0.11	-0.03	0.15
BetaTailRisk	-0.17	-0.01	93.5%	92.9%	-0.01	-0.23	0.24
BetaVIX	8.39	0.72	0.1%	16.6%	0.41	0.53	0.34
BidAskSpread	-10.96	-0.46	0.1%	24.2%	-0.36	-0.52	0.19
BookLeverage	-0.48	-0.03	76.7%	84.2%	-0.02	-0.03	0.08
BrandInvest	-0.87	-0.06	64.7%	51.0%	-0.04	-0.06	0.09
CBOperProf	6.88	0.74	0.0%	4.1%	0.62	0.94	0.16
CF	6.36	0.44	0.1%	5.4%	0.37	0.38	0.10
Cash	2.83	0.20	11.2%	7.3%	0.17	0.15	0.08
CashProd	1.57	0.12	34.2%	46.6%	0.11	0.05	0.10
Cfp	6.86	0.50	0.0%	2.6%	0.43	0.46	0.07
ChAssetTurnover	2.06	0.26	4.4%	18.7%	0.22	0.15	0.10
ChEQ	5.33	0.50	0.0%	1.4%	0.44	0.46	0.08
ChInv	5.79	0.70	0.0%	2.4%	0.56	0.60	0.17
ChInvIA	3.31	0.41	0.0%	0.0%	0.38	0.62	0.14
ChNNCOA	3.01	0.42	0.1%	4.2%	0.36	0.34	0.08
ChNWC	2.89	0.39	0.1%	10.6%	0.32	0.29	0.14
ChTax	3.68	0.38	0.2%	28.7%	0.32	0.24	0.13

Appendix Table 4: Inference and Posteriors for 174 Portfolios (continued)

			$\mu = 0$ p-values		Bayesian Posterior Means		
Portfolio	$E(r_t)$	SR	OLS (NW)	MLE	μ_{sr}	$\mu_{sr,2022}$	$\hat{\sigma}_{sr}$
ChangeInRecommendation	2.14	0.28	21.2%	30.1%	0.20	0.39	0.22
CompEquIss	3.85	0.30	1.0%	2.5%	0.26	0.30	0.07
CompositeDebtIssuance	1.87	0.25	10.7%	18.2%	0.18	0.03	0.18
CoskewACX	0.67	0.06	60.6%	67.3%	0.05	0.04	0.10
Coskewness	1.71	0.17	18.0%	21.8%	0.15	0.13	0.11
CustomerMomentum	9.93	0.39	0.6%	25.7%	0.31	0.31	0.12
DNoa	6.27	0.72	0.0%	0.2%	0.51	0.45	0.28
DelBreadth	1.90	0.19	12.0%	9.6%	0.16	0.20	0.08
DelCOA	2.71	0.27	4.0%	27.4%	0.22	0.20	0.11
DelCOL	-0.91	-0.08	46.1%	47.7%	-0.07	-0.10	0.06
DelDRC	4.51	0.40	4.5%	20.1%	0.25	0.30	0.16
DelEqu	3.48	0.30	2.1%	5.1%	0.26	0.29	0.09
DelFINL	3.91	0.63	0.0%	9.3%	0.50	0.55	0.18
DelLTI	1.75	0.33	1.5%	14.7%	0.29	0.44	0.12
DelNetFin	1.57	0.22	9.1%	20.5%	0.18	0.06	0.12
DolVol	1.25	0.09	55.3%	61.8%	0.06	-0.10	0.18
EBM	2.60	0.22	9.8%	19.0%	0.19	0.16	0.13
EP	3.47	0.25	5.0%	3.6%	0.22	0.25	0.09
EarnSupBig	3.43	0.25	0.8%	3.7%	0.23	0.22	0.06
EarningsConsistency	3.42	0.27	2.1%	6.0%	0.24	0.21	0.09
EarningsForecastDisparity	3.68	0.25	2.2%	1.2%	0.20	0.20	0.08
EarningsStreak	8.37	0.93	0.0%	0.9%	0.71	0.82	0.14
EarningsSurprise	2.43	0.33	0.6%	1.7%	0.29	0.29	0.09
EntMult	5.82	0.39	0.4%	1.4%	0.32	0.26	0.12
EquityDuration	7.82	0.48	0.1%	9.4%	0.38	0.36	0.18
ExclExp	2.68	0.54	0.0%	3.7%	0.43	0.46	0.10
FEPS	11.91	0.77	0.0%	0.9%	0.61	0.87	0.15
FR	1.20	0.15	49.9%	58.4%	0.14	0.55	0.38
Fgr5yrLag	7.72	0.43	0.4%	2.1%	0.35	0.42	0.12
FirmAge	-1.20	-0.14	30.7%	42.8%	-0.12	-0.30	0.15
FirmAgeMom	18.00	0.78	0.0%	6.4%	0.61	0.96	0.24
ForecastDispersion	6.14	0.38	0.3%	6.2%	0.32	0.42	0.10
Frontier	1.79	0.09	50.5%	24.8%	0.07	-0.04	0.12
GP	4.47	0.39	0.4%	10.8%	0.34	0.58	0.15
GrAdExp	3.48	0.29	4.7%	4.8%	0.24	0.28	0.11
GrLTNOA	1.72	0.21	6.5%	1.1%	0.18	0.19	0.07
GrSaleToGrInv	3.02	0.37	1.7%	26.6%	0.28	0.28	0.29
GrSaleToGrOverhead	0.63	0.07	46.8%	59.4%	0.06	0.04	0.06
Grcapx	4.08	0.41	0.5%	2.5%	0.33	0.28	0.16
Grcapx3y	2.93	0.29	4.4%	7.1%	0.24	0.14	0.13
Herf	0.77	0.07	56.6%	69.9%	0.06	0.03	0.10
HerfAsset	0.84	0.08	52.7%	59.8%	0.07	0.07	0.11
HerfBE	0.78	0.07	56.3%	68.8%	0.06	0.03	0.13
High52	15.24	0.76	0.0%	2.2%	0.64	0.70	0.12
Hire	2.23	0.22	11.4%	21.4%	0.17	0.10	0.13
IdioRisk	15.96	0.78	0.0%	9.0%	0.51	0.91	0.39

Appendix Table 4: Inference and Posteriors for 174 Portfolios (continued)

			$\mu = 0$ p-values		Bayesian Posterior Means		
Portfolio	$E(r_t)$	SR	OLS (NW)	MLE	μ_{sr}	$\mu_{sr,2022}$	$\hat{\sigma}_{sr}$
IdioVol3F	15.75	0.74	0.0%	4.6%	0.52	0.87	0.35
IdioVolAHT	13.98	0.56	0.0%	10.1%	0.42	0.61	0.25
Illiquidity	3.01	0.25	20.5%	51.5%	0.12	-0.85	0.45
IndMom	2.88	0.19	19.5%	26.1%	0.16	0.19	0.17
IndRetBig	20.46	1.26	0.0%	10.2%	0.67	0.48	0.46
IntMom	11.93	0.63	0.0%	15.9%	0.44	0.15	0.24
IntanBM	0.30	0.02	90.3%	93.6%	0.01	-0.17	0.15
IntanCFP	2.43	0.14	23.6%	28.8%	0.11	0.05	0.08
IntanEP	1.92	0.12	33.1%	28.4%	0.10	0.07	0.07
IntanSP	-1.16	-0.06	69.7%	78.1%	-0.05	-0.13	0.15
InvGrowth	5.14	0.49	0.1%	21.5%	0.39	0.49	0.19
InvestPPEInv	4.92	0.49	0.1%	5.9%	0.36	0.21	0.22
Investment	1.16	0.12	26.1%	46.8%	0.11	0.08	0.07
LRreversal	1.67	0.08	57.1%	65.0%	0.07	0.05	0.15
Leverage	1.83	0.10	42.4%	53.0%	0.08	0.05	0.11
MRreversal	3.59	0.24	13.0%	24.5%	0.19	0.10	0.20
MaxRet	11.42	0.65	0.0%	0.4%	0.50	0.76	0.30
MeanRankRevGrowth	3.62	0.35	0.7%	13.3%	0.29	0.25	0.10
Mom12m	15.28	0.68	0.0%	3.0%	0.57	0.56	0.10
Mom12mOffSeason	8.88	0.44	0.0%	0.1%	0.39	0.40	0.05
Mom6m	10.46	0.55	0.0%	4.8%	0.48	0.58	0.11
Mom6mJunk	5.34	0.30	5.9%	28.2%	0.22	0.23	0.15
MomOffSeason	5.29	0.30	2.2%	26.5%	0.25	0.17	0.13
MomOffSeason06YrPlus	8.80	0.73	0.0%	8.4%	0.58	0.69	0.17
MomOffSeason11YrPlus	3.70	0.35	0.7%	1.8%	0.30	0.27	0.09
MomOffSeason16YrPlus	2.82	0.31	0.6%	2.7%	0.27	0.27	0.08
MomSeason	4.91	0.42	4.7%	21.3%	0.26	0.32	0.48
MomSeason06YrPlus	6.16	0.58	0.1%	11.1%	0.41	0.38	0.33
MomSeason11YrPlus	5.76	0.61	0.0%	2.3%	0.52	0.52	0.08
MomSeason16YrPlus	3.09	0.32	0.2%	1.2%	0.29	0.27	0.10
MomSeasonShort	5.76	0.46	0.2%	38.2%	0.32	0.01	0.26
NOA	4.23	0.41	0.8%	10.2%	0.32	0.34	0.24
NetDebtFinance	3.55	0.54	0.0%	7.8%	0.43	0.50	0.14
NetDebtPrice	2.40	0.15	18.2%	9.1%	0.13	0.11	0.07
NetEquityFinance	5.36	0.55	0.0%	2.9%	0.46	0.52	0.10
NetPayoutYield	7.02	0.53	0.0%	3.1%	0.43	0.74	0.16
NumEarnIncrease	2.69	0.47	0.0%	4.6%	0.40	0.49	0.09
OPLeverage	3.41	0.28	1.0%	6.3%	0.25	0.27	0.07
OperProf	6.20	0.47	0.1%	0.9%	0.40	0.58	0.15
OperProfRD	7.05	0.58	0.1%	8.0%	0.42	1.33	0.37
OptionVolume1	3.35	0.28	28.5%	32.3%	0.12	-0.09	0.31
OptionVolume2	2.95	0.38	5.0%	35.8%	0.24	0.19	0.18
OrderBacklog	-1.07	-0.07	60.5%	68.7%	-0.05	0.06	0.12
OrderBacklogChg	-0.52	-0.04	72.7%	65.1%	-0.03	-0.06	0.09
OrgCap	5.12	0.55	0.0%	0.3%	0.48	0.49	0.07
PS	8.41	0.42	0.1%	3.0%	0.34	0.30	0.10
PayoutYield	4.32	0.33	1.7%	4.0%	0.28	0.44	0.13

Appendix Table 4: Inference and Posteriors for 174 Portfolios (continued)

			$\mu = 0$ p-values		Bayesian Posterior Means		
Portfolio	$E(r_t)$	SR	OLS (NW)	MLE	μ_{sr}	$\mu_{sr,2022}$	$\hat{\sigma}_{sr}$
PctAcc	2.34	0.26	1.1%	7.5%	0.23	0.23	0.08
PctTotAcc	0.83	0.10	45.8%	57.4%	0.08	0.07	0.10
PredictedFE	2.90	0.19	16.9%	34.2%	0.16	0.21	0.10
Price	-5.93	-0.21	22.2%	38.9%	-0.15	-0.35	0.33
PriceDelayRsq	-1.75	-0.16	26.6%	10.8%	-0.14	-0.20	0.11
PriceDelaySlope	-0.90	-0.09	51.4%	54.9%	-0.07	-0.08	0.11
PriceDelayTstat	0.24	0.04	77.1%	31.3%	0.03	0.02	0.09
ProbInformedTrading	15.77	1.21	0.0%	10.9%	0.63	0.81	0.29
RD	1.49	0.09	58.0%	67.0%	0.07	0.08	0.20
RDAbility	5.35	0.26	12.8%	12.1%	0.21	0.28	0.13
RDS	2.01	0.35	2.3%	9.2%	0.26	0.21	0.15
RDcap	1.60	0.07	69.0%	82.0%	0.03	-0.37	0.23
REV6	4.36	0.20	21.4%	49.8%	0.13	-0.06	0.19
Realestate	3.22	0.31	3.7%	13.2%	0.24	0.19	0.15
ResidualMomentum	5.96	0.57	0.0%	4.2%	0.48	0.43	0.10
RetConglomerate	6.45	0.46	0.0%	2.9%	0.37	0.36	0.12
ReturnSkew	-0.09	-0.01	92.8%	96.1%	0.01	0.33	0.25
ReturnSkew3F	-1.77	-0.33	1.5%	42.1%	-0.26	-0.19	0.16
RevenueSurprise	1.51	0.20	7.5%	22.0%	0.18	0.19	0.09
RoE	2.89	0.26	4.6%	10.7%	0.23	0.34	0.13
Roaq	6.66	0.40	0.4%	3.1%	0.31	0.51	0.15
SP	4.76	0.27	5.4%	13.1%	0.22	0.17	0.14
STreversal	-0.19	-0.01	92.7%	94.2%	-0.02	-0.15	0.14
Sfe	10.96	0.56	0.0%	0.1%	0.47	0.48	0.09
ShareIss1Y	6.53	0.83	0.0%	1.8%	0.66	0.71	0.20
ShareIss5Y	4.96	0.59	0.0%	0.0%	0.48	0.47	0.16
ShortInterest	4.41	0.41	0.2%	0.8%	0.35	0.41	0.09
Size	-0.06	-0.00	98.7%	97.8%	0.00	-0.15	0.31
Skew1	5.90	0.47	1.3%	23.3%	0.33	0.41	0.16
SmileSlope	13.06	1.34	0.0%	18.8%	0.65	0.88	0.39
Std_turn	11.39	0.70	0.0%	15.0%	0.45	0.50	0.36
Tang	0.60	0.05	66.7%	41.4%	0.04	0.07	0.07
Tax	4.34	0.53	0.0%	2.8%	0.43	0.78	0.22
TotalAccruals	2.48	0.29	4.2%	25.4%	0.25	0.34	0.13
TrendFactor	11.88	0.74	0.0%	20.4%	0.51	0.21	0.25
VarCF	1.50	0.10	51.2%	65.4%	0.09	0.31	0.23
VolMkt	4.81	0.34	1.8%	1.3%	0.28	0.34	0.13
VolSD	4.10	0.52	0.0%	3.6%	0.44	0.46	0.09
VolumeTrend	4.69	0.42	1.4%	6.8%	0.32	0.48	0.26
XFIN	9.93	0.70	0.0%	9.9%	0.53	0.91	0.23
Zerotrade	5.06	0.51	0.1%	6.7%	0.41	0.50	0.17
ZerotradeAlt1	4.31	0.39	2.6%	8.1%	0.28	0.56	0.30
ZerotradeAlt12	5.18	0.52	0.1%	10.0%	0.41	0.46	0.18