Submission for the Q Group Treynor Prize

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

Authors: Aydogan Alti, Travis L. Johnson, and Sheridan Titman

Research Summary:

The premise of our research agenda is that characteristic-sorted portfolio returns exhibit persistent fluctuations – such fluctuations are likely to occur as the underlying economic forces that link returns to characteristics change over time. As we show in our paper, accounting for these fluctuations has important implications for how financial economists should evaluate the significance of the historical returns of characteristic-based strategies, as well as how practitioners should incorporate information from historical returns into their strategies.

Our analysis starts with a model that provides a simple formula for adjusting the standard errors of expected return estimates relative to the traditional estimates that ignore persistence. The adjusted standard errors double under plausible assumptions. Our maximum likelihood estimates of the model also indicate that the historical data are consistent with a wide range of return processes, casting doubt on some of the traditional interpretations of the historical data. Given this wide range of plausible return processes, it is natural to apply Bayesian methods to explore how beliefs about the process generating returns, combined with the historical evidence, affects posterior beliefs and portfolio choices. We show that investors' posterior beliefs about expected returns are highly dependent on their priors about persistence, even after observing close to 60 years of data.

The insights developed in our paper suggest that many of the ever-increasing number of characteristics or factors that predict returns historically may be transient artifacts of past conditions rather than permanent features of the market. Perhaps more importantly, we provide a Bayesian methodology that can be used by practitioners to sort through the historical evidence in forming portfolios.

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

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

March 2022

Abstract

We present a statistical model that accounts for persistent fluctuations in characteristic-sorted portfolio returns. The model provides a simple formula for adjusting the standard errors of expected return estimates. With plausible parameter values, the adjusted standard errors double, casting doubt on the interpretation of the historical performance of characteristic-sorted portfolios as evidence of long-term return premia. Similarly, maximum likelihood estimates of the model indicate that the historical data are consistent with a wide range of return processes. Finally, we use Bayesian methods to show that investors' posterior beliefs about expected returns are highly dependent on their priors about persistence, even after observing close to 60 years of data.

*All three authors are with the University of Texas at Austin. We thank Scott Cederberg, Lubos Pastor, and seminar participants at Peking University and the University of Texas at Austin for helpful comments.

A large and growing literature links firm characteristics, such as valuation ratios, to expected rates of 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 is 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 eventually 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 in a recent study, we cannot confidently conclude that the value premium post-1992 is different than the value premium pre-1992. 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 vary over time. The central feature of our model, persistent variation in expected returns, is implicit in existing rational and behavioral theories of characteristic-based return predictability. The general idea is that the economic forces that generate relationships between characteristics and returns are likely to change over time.[2] Our main objective is to analyze how such persistent variation influences our inferences about the past as well as the predictions about the future returns of characteristic-sorted portfolios.

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

[2]Gârleanu, Kogan, and Panageas (2012) and Kogan, Papanikolaou, and Stoffman (2020) provide examples of rational theories where investors hold lower returning growth stocks to hedge technology shocks. Shiller (2000) provides a behavioral explanation based on investors being overly optimistic about the commercialization potential of new technologies, causing growth stocks to underperform. One might expect the economic fundamentals that drive both rational and behavioral explanations to fluctuate over time, or possibly even reverse, causing growth stocks to outperform value stocks. See Altı and Titman (2019) for a dynamic behavioral model that captures both possibilities.

Our analysis provides three main contributions. First, we derive a simple formula that adjusts the standard errors of unconditional expected return estimates to account for persistent variation in expected returns. We show that, given plausible persistence parameters, this adjustment can have a large effect on the precision of the estimates, in some cases doubling their standard errors. Second, we estimate the full model, including the persistence parameters, using maximum likelihood. The persistence parameters are estimated with very little precision; in addition, the precision of model-based estimates of unconditional expected returns are consistent with our adjusted OLS estimates. These two findings imply that the historical return data are consistent with a wide range of return processes, which include persistent expected returns that fluctuate around a zero unconditional mean, as well as i.i.d. returns with large means. Given these diverse possibilities, it is natural to ask how Bayesian investors with different prior beliefs interpret the data and forecast future return performance of characteristic-sorted portfolios. This leads to our third main finding, which is that the posterior beliefs of Bayesian investors are highly sensitive to their priors about the degree of persistence, even after observing 56 years of data.

The focus of our empirical analysis is on the returns of the four characteristic-sorted portfolios described in Fama and French (2015) – value, investment, profitability, and size. We first show that all four exhibit positive return autocorrelation over yearly horizons, and while only the size portfolio's autocorrelation estimate is significant on its own, a joint test of the four portfolios strongly rejects the zero-autocorrelation null hypothesis. These autocorrelation estimates, however, are sufficiently imprecise that they are consistent with a wide variety of parameterizations of our model, from persistence lasting months to decades.

We then apply our model to analyze the precision of expected return estimates. In contrast to the prior literature, which employs standard errors that are either unadjusted or corrected using Newey and West (1987), we adjust our standard errors using the covariance structure implied by our model. Specifically, given a return autocorrelation at one lag and its decay rate over time, the model provides a simple, closed-form formula for adjusting the standard errors of unconditional expected return estimates. This adjustment results in higher standard errors when returns are more persistent, that is, when the autocorrelation is more positive and decays more slowly. Intuitively, a persistent return time series exhibits less independent variation, generating less precise estimates of expected returns.

As our estimates illustrate, different assumptions about the magnitude of persistent variation in characteristic-sorted portfolio returns, all consistent with the historical data, generate substantially different inferences about unconditional expected returns. If returns are assumed to be i.i.d., meaning they have no persistence, standard errors are low and the hypothesis of zero unconditional expected returns is strongly rejected. However, if returns are assumed to exhibit plausible degrees of persistence, the model-adjusted standard errors are substantially higher and the hypothesis is often not rejected. For example, standard errors nearly double, relative to the case where returns are i.i.d., when return autocorrelation is 5% at one-quarter lag and decays with a half-life of five years. Our analysis also shows that the commonly-used Newey and West (1987) procedure is ineffective when applied to persistent return time series and results in much smaller standard error corrections compared to our model-based adjustments.

In addition to estimating standard errors using assumed parameters, we use maximum likelihood to estimate the full set of model parameters, including those that determine the magnitude of persistent variation in returns. Consistent with the highly noisy autocorrelation estimates in our reduced-form regressions, we find that the persistence parameters are imprecisely estimated. Given these imprecise estimates, we cannot rule out a wide variety of a plausible alternative explanations for the historical performance of the characteristic-sorted portfolios. Indeed, for several of the portfolios, likelihood ratio tests fail to reject the hypothesis that unconditional expected returns are zero, but conditional expected returns are both variable and highly persistent.

While the results discussed so far are expressed from a frequentist statistical perspective, our focus on how assumptions about time-variation affect inferences about unconditional expected returns has a natural Bayesian interpretation. Specifically, our OLS estimates are similar to Bayesian analyses with dogmatic priors about the degree of persistent variation in expected returns, while the maximum likelihood estimations are similar to Bayesian analyses with agnostic priors. This analogy is limited, however, to priors about persistence, as the frequentist analysis assumes fully agnostic priors about unconditional expected returns.

To assess how priors about unconditional expected returns and expected return persistence interact, we embed the model into a Bayesian framework. Specifically, we assume that different investors start with different, but economically plausible priors, and that they update their beliefs based on the same return data that we observe in our sample period. The value of this exercise is to quantify the extent to which individuals with different perspectives on the determinants of expected returns may differ in their posterior beliefs after observing the historical data. For example, investors guided by predictions of the CAPM may have the prior belief that the unconditional expected returns of characteristic-sorted portfolios cannot deviate substantially from zero, while others may have less confidence in the CAPM.

We find that prior beliefs about persistence substantially affect how investors, after observing the return history, update their beliefs about unconditional expected returns. If investors have strong priors that expected returns fluctuate very little, then their posterior beliefs about unconditional expected returns tend to be relatively precise and not highly sensitive to priors about unconditional expected returns. However, if investors' priors put more weight on the possibility of persistent fluctuations, then their posterior beliefs about unconditional expected returns become more diffuse and more sensitive to their priors. In this way, persistence makes investors learn less from data and rely more on their priors.

We also use our Bayesian approach to generate estimates of conditional expected returns and Sharpe ratios in 2020, the year after our sample period ends. These estimates speak directly to the ongoing debate about the recent performance of value strategies and the effectiveness of factor timing strategies.[3] We find that Bayesian investors who account for time-varying expected returns will have very different views of near-term and long-term expected returns. The reason is that while data from the early and later parts of the sample are equally important for estimating unconditional expected returns, near-term conditional expected return estimates put more weight on the more recent observations when returns are persistent. For example, because the value portfolio performed particularly poorly towards the end of the sample, the value premium in 2020, measured as the mean of the posterior, is generally around a quarter of the unconditional value premium, and is close to zero in many specifications. The profitability portfolio exhibits the opposite pattern: because the returns were stronger in recent decades, the posteriors for conditional expected returns in 2020 are higher than posteriors for the unconditional expected return.

In our final analysis, we repeat the Bayesian estimation procedure at the end of each calendar year using only past data to form posterior beliefs about the distribution of next-year returns. Based on these beliefs, we compute optimal portfolio weights assuming constant relative risk aversion (CRRA) utility. We find that investors with different priors about unconditional expected returns choose substantially different portfolios early in the sample but quickly converge to similar choices. Investors with different priors about persistence, by contrast, have ongoing differences in portfolio choices driven by the extent of market timing they pursue. Priors tilted towards longer persistence result in more-aggressive dependence of portfolio choices on recent performance. This timing behaviour is particularly strong for the size portfolio, which despite not having near-zero average returns in the full sample has large variations in conditional expected returns.[4]

[3]See Arnott et al. (2021b), Asness et al. (2021), Choi, So, and Wang (2021), Eisfeldt, Kim, and Papanikolaou (2022), Fama and French (2021), and Goncalves and Leonard (2021).

[4]Brown, Kleidon, and Marsh (1983), Conrad and Kaul (1988), and Guidolin and Timmermann (2008) also find evidence of time-variation in conditional size portfolio returns.

While we are among the first to analyze time variation in the expected returns of characteristic-sorted portfolios, there is a well-established literature that explores a number of these issues within the context of the aggregate equity market portfolio. For example, Ferson, Sarkissian, and Simin (2003) consider the predictability of aggregate market returns using persistent 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 errors adjustment is used. Although our application is different, the problem we address is similar and we also show that the Newey and West (1987) correction is not effective in dealing with the problem. In addition, we propose some frequentist remedies, such as OLS standard error corrections when the level of persistence in expected returns is known, or maximum likelihood when the level of persistence is estimated.

Pástor and Stambaugh (2009, 2012) and Avramov, Cederburg, and Lučivjanská (2018) also conduct Bayesian analyses of time-varying expected returns and find that the priors about the return generating process substantially affect the posteriors about expected returns, a similar conclusion to the one we reach. Our study differs from these papers in both application and focus, examining returns of characteristic-sorted portfolios instead of timing the aggregate equity market portfolio based on imperfect predictors.[5]

We are aware of only two studies that use Bayesian methods to study time variation in characteristic-sorted portfolio returns: Pástor (2000) and Smith and Timmermann (2022). In contrast to our analysis, persistence plays no role in Pástor (2000) as expected returns are assumed to be constant. Smith and Timmermann (2022) instead assume that there are occasional structural breaks where 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

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

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 have the advantage of permitting out-of-sample forecasting – structural breaks can only be identified in hindsight – as well as standard error corrections for in-sample inference.

Finally, our analysis is related to a number of recent papers that combine evidence from several characteristic-sorted portfolios to show that there is persistent time-variation in conditional expected returns, e.g., Lewellen (2002), McLean and Pontiff (2016), Avramov et al. (2017), Gupta and Kelly (2019), Arnott et al. (2021a), and Ehsani and Linnainmaa (2021). We contribute to this literature by providing model-based estimates of the magnitude of persistent variation in returns, and by studying the implications of persistence for inferences on 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 performance and the evidence of return autocorrelation for the characteristic-sorted portfolios. Section 3 presents the OLS and maximum likelihood estimations of the model. Section 4 presents the Bayesian analysis. Section 5 concludes.

1. Statistical Model

In this section, we describe and analyze a simple statistical model of the return-generating process that we later apply to characteristic-sorted portfolio returns. The important feature of the model is the assumption that the conditional expected returns of the portfolio exhibits persistent variations around an unconditional mean.

1.1. Specification

We assume that the time-series of the zero-cost 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 ϵ_t and δ_t are i.i.d. and follow a joint normal distribution with variances σ_ϵ and σ_δ, respectively, and correlation $\rho \in (-1, 1)$.[6] We expect ρ to be negative, since shocks to expected rates of return, ceteris paribus, reduce an investment's value.[7]

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).

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

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

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

[6] Conrad and Kaul (1988) uses the same specification but assume $\rho = 0$.

[7] In Equation (2) we multiply the shock δ_{t+1} by λ so that returns are i.i.d. when $\lambda = 0$.

Equation (4) 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 constant rate λ as $|i-j|$ grows.

1.2. Identification

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

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

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

Using this alternative notation, the covariance matrix becomes:

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

Note that any two parameterizations Ω and $\tilde{\Omega}$ satisfying

$$[\mu, \lambda, \sigma_r, \gamma] = \left[\tilde{\mu}, \tilde{\lambda}, \tilde{\sigma_r}, \tilde{\gamma}\right] \tag{9}$$

result in identical mean and covariance matrix for $\mathbf{R}$ (μ and $\mathbf{\Sigma}$). Because we have five parameters to satisfy four equations, many distinct Ω and $\tilde{\Omega}$ generate any given μ and $\mathbf{\Sigma}$.

To better understand identification in our model, note that the sample mean of returns identifies the unconditional expected return μ (Equation (4)), while the rate at which the covariance between r_i and r_j decays as $|i-j|$ grows identifies the persistence parameter λ (Equation (5)). This leaves three other 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, the identification problem arises because one cannot distinguish between different 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 same increases can also be generated from increases in the volatility of unexpected return shocks σ_ϵ and the correlation parameter ρ.

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 $\sigma_\epsilon > 0$, $\sigma_\delta > 0$, $\lambda \geq 0$, and $\rho \in (-1, 1)$. The identification problem is not an issue for 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.

1.3. The sign of return autocorrelations

Although our model can generate both positive and negative return autocorrelations, negative autocorrelations tend to occur only when expected returns exhibit very little persistence. Time variation in expected returns contributes to positive autocorrelation because current and recent past returns have similar levels of expected returns. A negative return autocorrelation requires a negative correlation between realized returns and *changes* in expected returns. However, unless shocks to expected returns are quickly mean-reverting, the information in past returns about a single period's change in expected return is less relevant than the information about the ongoing level of expected returns.

We can see this formally using Equation (7), which specifies the first-order return au-

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

tocorrelation γ as a function of λ, ρ, σ_ϵ, and σ_δ. The first term in the numerator, $\lambda^2\sigma_\delta^2$, represents the effect of persistent expected return variations and is always positive. The second term in the numerator, $(1-\lambda^2)\rho\sigma_\delta\sigma_\epsilon$, represents the effect of correlated shocks to expected and realized returns and is negative when $\rho < 0$. Comparing these two terms reveals that when λ is close to one, as is the case for the highly-persistent variations that motivate our analysis, $\gamma > 0$. To obtain $\gamma < 0$, λ needs to be small relative to $|\rho\sigma_\epsilon/\sigma_\delta|$.

We quantify how small λ (or, equivalently, H) needs to be for γ to be negative in Appendix Table 1. We assume $\rho = -1$ to provide an upper bound on the prevalence of negative autocorrelations, and consider various combinations of either H and $\sigma_\epsilon/\sigma_\delta$, or H and the time-series volatility of conditional Sharpe ratios, σ_{sr}.[9] We find that materially negative γ values obtain only when H is one quarter of a year or less.

2. Characteristic-Sorted Portfolios

We apply our statistical model to study four portfolios that are formed by sorting stocks based on firm characteristics: value, investment, profitability, and size. We focus on these portfolios because they are the basis of the Fama and French (2015) five-factor model and illustrate a variety of channels through which time-varying expected returns affect our understanding of portfolio returns.

2.1. Data and characteristic definitions

We use data on the historical returns of characteristic-sorted portfolios from Ken French's website.[10] Each portfolio combines a long position in a value-weighted portfolio of firms in one extreme quintile of the characteristic with a short position in the other extreme.

The characteristics are defined following Fama and French (2015). *Value* is the ratio of

[9]See Section 4.1 for details on how σ_{sr} maps to σ_ϵ, σ_δ, and H.
[10]http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

the book value of equity ($B_{i,y}$) to the market value of equity ($M_{i,y}$) as of the end of the prior fiscal year y. *Investment* is the growth rate in the book value of assets ($\text{Assets}_{i,y}/\text{Assets}_{i,y-1}$). *Profitability* is 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 $M_{i,y}$.

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 portfolios we study exhibit strong positive first-order autocorrelations.[11] 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. To focus our analysis on longer-term autocorrelations driven by persistent variations in expected returns, we estimate our model using quarterly returns.

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

$$r_{i,t}^{\beta=0} = r_{i,t} - \hat{\beta}_i(r_{m,t} - r_{f,t}), \tag{10}$$

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. The average return and the Sharpe ratio of this market-neutral portfolio are equivalent to the alpha and information ratio, respectively, of the underlying long-short portfolio.

2.2. Historical performance of characteristic-sorted portfolios

Table 1 summarizes the historical performance of the characteristic-sorted portfolios. The value, investment, and profitability portfolios have annualized mean returns above 4% and Sharpe ratios between 0.3 and 0.5, while the size portfolio's mean return is close to zero.

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

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, 1963–1990 and 1991–2019. 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 patterns across subsamples. The investment portfolio's returns are largely consistent over time and statistically significant in both halves of the sample. The profitability portfolio follows the opposite pattern as the value portfolio, performing worse in the first half of the sample than the second, though again the difference is statistically insignificant. The size portfolio has small and statistically insignificant returns in both halves of the sample.[14]

2.3. Autocorrelation estimates

As discussed in the Introduction, a number of recent studies document short-term autocorrelation in characteristic-sorted portfolio returns, or more generally analyze portfolio timing strategies that are inherently based on time-variation in expected returns. Before we estimate our model we present similar reduced form analyses for the four characteristic sorted portfolios that we analyze.

[12]We compute the i.i.d. standard errors by taking the standard deviation 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, 2021), 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 return and thus are consistent with our model.

[14]The absence of a significant size effect in the early part of the sample is in part due to our use of value weights in portfolio construction.

Estimates of the autocorrelation structure for individual portfolios are inherently imprecise because realized returns are quite volatile relative to the plausible variations in expected returns. Furthermore, autocorrelation estimates have a well-document downward bias in small samples (Kendall, 1954; Marriott and Pope, 1954). To illustrate these difficulties, we first present autocorrelation estimates that we generate with model-simulated data samples under a variety of assumptions about H and γ. In these simulations, we fix the model parameters and generate 50,000 samples with 226 observations, which is the number of quarters in our empirical sample.[15] For each simulated return series, we estimate return autocorrelations using regressions of quarterly returns on averages of returns over the previous L quarters:

$$r_t = a + b_L \left(\frac{1}{L} \sum_{l=1}^{L} r_{t-l} \right) + \epsilon_t. \tag{11}$$

Panel A of Table 2 presents average values, as well as 95% confidence intervals, for the autocorrelation coefficient $\hat{b}_L$ across samples that are simulated under a variety of plausible parametric assumptions. Specifically, we include shocks with half-lives, H, that are 2.5, 5, and 10 years, and first-order autocorrelations, γ, that equal 2.5%, 5%, and 10%. For each parameterization, we report the standard deviation of annualized Sharpe ratios conditional on past returns, σ(Cond. Sharpe), that is implied by that parameterization. These values indicate that even seemingly small values for γ (2.5% or 5%) can generate economically meaningful – though plausible – variations in conditional Sharpe ratios.

The first row of Panel A in Table 2 reports the coefficient estimates when $H = 0$ and thus returns are distributed i.i.d. As expected, the estimated average autocorrelation coefficients are negative due to the aforementioned downward bias that arises because we are using the in-sample mean to calculate the autocorrelation, generating what looks like mean-reversion

[15] The remaining model parameters we use for the simulations are $\mu = 0$ and $\sigma_r = 6.89\%$, the full-sample return standard deviation for the value portfolio. These choices have no effect on our results as autocorrelation estimates from regressions with intercepts are invariant to linear transformations of r_t.

even when there is no true mean reversion. The bias is stronger the longer the past return window because there is a smaller effective sample size. To take an extreme example, if we divide the sample in two, one half will appear above-average and the other below-average, suggesting mean reversion, in every random sample.

The other rows in Panel A of Table 2 show the autocorrelation estimates that obtain when expected returns exhibit persistent variation, i.e., $H > 0$ and $\gamma > 0$. Two observations emerge from the reported estimates. First, the downward bias continues to affect autocorrelation estimates even when expected returns are persistent, especially with longer past return windows. Second, and more importantly, the confidence intervals for the autocorrelation estimates are quite wide and include negative values in every parameterization – even those with true autocorrelations that are quite high.

Panel B of Table 2 presents estimates of Equation (11) for the historical samples of the quarterly returns of the value, investment, profitability, and size portfolios. The first column shows that all four portfolios have positive b_4, indicating that past-year returns positively predict next-quarter returns. The second and third columns show that longer past-return windows have point estimates with varying signs.

We do not report standard errors or bias corrections in Panel B of Table 2 because Panel A shows that both depend heavily on the magnitude and persistence of the variations in expected returns. In most cases, the confidence intervals for different parameterizations in Panel A include all of the point estimates in Panel B, which means that one cannot reject any of the posited autocorrelation structures. Even economically large coefficients, such as $\hat{b}_{40} = -0.73$ for the investment portfolio, lie in the 95% confidence interval for every parameterization from i.i.d. to $H = 10$ and $\gamma = 5\%$.

For three out of the four characteristic-sorted portfolios, the autocorrelation estimates do not reject any reasonable parameterization of our model. The size portfolio is the exception; we convincingly reject the i.i.d. null in the regression with the one-year lag. In addition, in a pooled regression that combines the returns of the four portfolios, the i.i.d. null hypothesis is rejected for one-year and ten-year returns with p-values of 0.0% and 4.7%, respectively.[16] In a pooled regression with only value, investment, and profitability, we reject the i.i.d null with the one year lag with a p-value of 0.1%. The reason we can jointly but not individually reject the null is that the coefficients for the value, investment, and profitability portfolios are each above the i.i.d. benchmark, but not by enough to reject on an individual basis.

3. OLS and Maximum Likelihood Estimations

As the previous section shows, the autocorrelation patterns observed in the historical returns of characteristic-sorted portfolios are consistent with a range of assumptions about the magnitude of persistent variation in conditional expected returns. In this section, we formally analyze the impact of such variation on the estimates of unconditional expected returns using OLS regressions and maximum likelihood estimations.

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 $\psi_t = r_t - \mu$.[17]

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. When expected returns are time-varying and persistent, both of the standard approaches produce understated stan-

[16] We conduct this test comparing the sum of the individual estimates from observed data to the distribution of this sum in samples simulated under the i.i.d. assumption.

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

dard errors. The reason is that persistent variations in expected returns generate small but long-lasting correlations in ψ_t that extend beyond the windows considered by Newey and West (1987). Formally, our model generates the residuals

$$\psi_t = \mu_{t-1} - \mu + \epsilon_t, \tag{12}$$

which implies that ψ_t have long-lasting autocorrelations due to persistent variations in μ_t. As we show below, even if we extend the number of lags in Newey and West (1987) to match or exceed the half-lives of shocks to expected returns, standard errors remain under-estimated because the lags are too large a fraction of the observed time series for the asymptotic results in Newey and West (1987) to hold.

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 the structure of our model.[18] Specifically, the asymptotic standard error of $\hat{\mu}^{OLS}$ satisfies

$$\begin{aligned} \text{SE}(\hat{\mu}^{OLS}) &= \sqrt{\frac{\mathbf{1}'\mathbf{\Sigma}\mathbf{1}}{T^2}} \\ &= \left(\frac{\sigma_r}{\sqrt{T}}\right)\sqrt{1 + 2\gamma\left[\frac{\lambda^T + T(1-\lambda) - 1}{T(1-\lambda)^2}\right]}. \end{aligned} \tag{13}$$

In Equation (13), 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 fully specified by γ and λ, as shown in Equation (8). Substituting from Equation (8) results in the formula in the second line of Equation (13) after some algebraic manipulation. Note that the first term in this formula, $\sigma_r/\sqrt{T}$, is the unadjusted OLS

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

standard error. Thus, the second term in square roots is the adjustment factor as a function of T, γ, and λ. For large T, Equation (13) approximately equals

$$\mathrm{SE}(\hat{\mu}^{OLS}) \approx \left(\frac{\sigma_r}{\sqrt{T}}\right)\sqrt{1+\frac{2\gamma}{1-\lambda}}. \tag{14}$$

The standard error adjustments in Equations (13) and (14) 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 periods per year.

Table 3 shows how different assumptions about γ and H affect the t-statistics for unconditional expected return estimates that obtain with model-adjusted standard errors given by Equation (13). For comparison, we also report the Newey-West corrected t-statistics that are calculated using numbers of lags that match the half-life of shocks H that we consider. Across all four portfolios, we find that Newey-West t-statistics differ very little from the unadjusted t-statistics. Model-implied t-statistics, by contrast, are substantially lower. For instance, the model adjustment cuts t-statistics approximately by half when $H = 5$ years and $\gamma = 5\%$. Overall, the results in Table 3 indicate that accounting for plausible magnitudes of persistent variation in expected returns can result in inferences about unconditional expected returns that differ materially from the inferences that obtain with the assumption of i.i.d. returns.

3.2. Maximum likelihood tests

This subsection describes maximum likelihood estimates of our statistical model. In contrast to the least squares regressions described in the previous subsections, maximum likelihood estimation requires distributional assumptions, but allows us to estimate all of the model parameters. In particular, we will be estimating rather than assuming the magnitudes of the various persistence parameters.

3.2.1. Tests for $\mu = 0$

We test the null hypothesis that the unconditional expected return μ equals zero under a variety of assumptions about the structure of time-variations in conditional expected returns. For each assumption, we estimate the model using maximum likelihood twice, first with no constraints on μ and then under the restriction of $\mu = 0$. In both cases we restrict H to be less than or equal to 20 years because μ is not identified when H approaches infinity. Using these estimates, we compute the p-value for the $\mu = 0$ null using a likelihood ratio test.

The first set of columns in Panel A of Table 4 test whether $\mu = 0$ assuming no time variation in expected returns (i.e. $H = 0$ so returns are i.i.d.). Testing $\mu = 0$ under this assumption is analogous to using OLS with no standard error correction, and so the likelihood-ratio p-values strongly reject the null for value, investment, and profitability.

The next set of columns in Panel A of Table 4 relax the i.i.d. assumption and estimate the values of H and γ that maximize the likelihood of observing the historical data.[19] We find that when we allow for the possibility of time-varying means, the p-value for the hypothesis that $\mu = 0$ increases for the three portfolios and is no longer statistically significant at the 5% level. As the panel shows, when μ is restricted to be zero, the model fits the data by using larger values of H combined with positive values of γ. This combination explains the historical returns as arising from positive realizations of *persistent* expected returns that eventually dissipate. Because such persistent variation is difficult to reject empirically, it offers a plausible alternative to large μ, increasing the p-values for rejecting the $\mu = 0$ hypothesis above 5% in all cases.

[19] As discussed in Section 1, we estimate the parameters summarizing the covariance matrix of returns $\theta = [\mu, H, \sigma_r, \gamma]$ directly rather than the underlying parameters $\Omega = [\mu, \lambda, \sigma_\epsilon, \sigma_\delta, \rho,]$ because the later are not fully identified. We restrict our estimates of θ to the range for which there exists at least one possible Ω yielding the same covariance matrix of returns.

3.2.2. Restrictions on H and γ

Next, we use maximum likelihood estimates to assess the plausibility of the various assumptions about H and γ that we made in our OLS analyses in Section 3. For each assumption, we re-estimate the model by 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. We also use likelihood ratios to test whether we can reject the hypothesis that $\mu = 0$ given the restrictions on H and γ.

Panel B of Table 4 presents the results. For value, investment, and profitability, the likelihood ratio tests cannot reject *any* of the restrictions on the structure of time-variation in expected returns, including the i.i.d. hypothesis and the possibility of extremely-persistent and economically large shocks ($H = 10$ years, $\gamma = 5\%$). Consistent with the evidence in Table 3, Panel B of Table 4 also shows that these alternatives have material affects on our inferences about $\mu = 0$, with p-values for $\mu = 0$ going from below 2% in the i.i.d. case to above 10% in some specifications for all three characteristic-sorted portfolios.

While the results for value, investment, and profitability suggest that the data offer very little guidance about the magnitude of persistent variations in expected returns, the results for the size portfolio show that this is not always the case. Consistent with the reduced-form evidence in Table 2, our estimates for the size portfolio suggest strong autocorrelation in returns with a half life between one and two years. As a result, Panel B of Table 4 shows the maximum likelihood estimator strongly rejects both the i.i.d. restriction and the restriction that $H = 10$. The contrast between the results for the size portfolio and the other characteristic-sorted portfolios shows that these tests can have power, but the return patterns for value, investment, and profitability portfolios are particularly inconclusive.

4. Bayesian Analysis

The analysis in the previous section shows that the historical data are consistent with a variety of substantially different return generating processes. A natural next step is to ask how investors with different prior views about these alternatives interpret the historical data in making their investment decisions. We examine this normative question in this section. Specifically, we pursue a Bayesian analysis that specifies prior likelihoods of different model parameterizations and uses the observed data to calculate posterior likelihoods. We also compute posteriors for moments that are likely to affect investors' portfolio decisions, such as the near-term and long-term Sharpe ratios for each portfolio.

4.1. Prior beliefs

We specify prior beliefs over a transformation of the parameters that is focused on Sharpe ratios rather than expected returns because Sharpe ratios have a clearer connection to asset pricing theory and economic intuition.[20] Specifically, we specify priors over μ_{sr}, the unconditional Sharpe ratio of portfolio returns; H, the half-life of shocks to expected returns; σ_r, the unconditional standard deviation of portfolio returns; σ_{sr}, the standard deviation of conditional Sharpe ratios; and ρ, the correlation between unexpected returns shocks to expected returns. We tabulate μ_{sr}, H, σ_r, and σ_{sr} in annualized units. These annualized parameters map from the underlying model parameters Ω as follows, where $N = 4$ is the number of periods per year:

$$\mu_{sr} = \frac{\mu}{\sigma_\epsilon}\sqrt{N}, \qquad H = \frac{\log(0.5)}{\log(\lambda)}\frac{1}{N}, \qquad \sigma_r^2 = \left(\frac{\lambda^2\sigma_\delta^2}{1-\lambda^2} + \sigma_\epsilon^2\right)N,$$

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

[20] 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 or the amount of financial leverage.

We consider a variety of priors on μ_{sr} and H, summarized in Panel A of Table 5. For μ_{sr}, we first consider normal prior distributions centered at -0.4, 0, and 0.4, all with a standard deviation of 0.4. These priors can be interpreted as beliefs that the portfolio has unconditional Sharpe ratios that are likely to be negative, likely to be near-zero, or likely to be positive.[21] For the value portfolio, these three investors can be viewed as having growth, neutral, or value inclinations. We also examine an uninformative prior where any μ_{sr} between -2 and 2 is equally likely.[22]

To illustrate the effect that prior beliefs about H have on Bayesian inferences about expected returns, we consider three dogmatic priors and one agnostic prior. The dogmatic priors assert that $H = 0$ (making returns i.i.d.), $H = 2.5$ years, or $H = 5$ years with certainty. The agnostic prior, by contrast, views H as unknown and uniformly distributed between 0 and 10 years.

We consider uniform priors over wide ranges for the remaining parameters. The prior for the volatility of annual returns, σ_r, is uniformly distributed between 10% and 20%.[23] The prior for the standard deviation of conditional Sharpe ratios, σ_{sr}, is uniformly distributed between 0 to 1.[24] The correlation between unexpected returns and shocks to expected returns, ρ, is likely to be negative given the inverse relation between prices and expected returns. In contrast to 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 $[-0.5, 0]$.

[21] We use ± 0.4 as the center for our Sharpe ratio distributions to roughly match the US equity market's estimated Sharpe ratio.

[22] Uniform distributions over wider supports give nearly-identical results because the data strongly reject unconditional Sharpe ratios above 2 or below -2 for for the portfolios we study when H is less than 20 years.

[23] This is the only prior that would need to shift as the leverage applied to a portfolio changes.

[24] In the most extreme case, this implies that conditional Sharpe ratios occasionally deviate from their unconditional means by as much as 2. While this extreme case is implausible given active asset management chasing and reducing large Sharpe ratio opportunities, as we emphasize below the true conditional Sharpe ratio may not be observable, and Sharpe ratios that are achievable by conditioning only on past returns will be much less extreme.

To provide further intuition for the set of priors we study, Panel B of Table 5 presents the distribution of economically-intuitive moments implied by the priors specified in Panel A. We calculate these moments by simulating 50,000 draws from each prior and calculating the value of each moment implied by each parameter draw. The first set of columns shows that μ, the unconditional expected return, has a prior mean that is about -5.3%, zero, or 5.3%, depending on the prior specification. The middle set of columns show that the prior mean values of one-lag return autocorrelation, γ, are positive in all cases, indicating that the positive effect due to persistence of expected returns typically outweighs the negative effect due to $\rho < 0$, though the 95% confidence intervals do include negative values.

The parameter σ_{sr} governs the volatility of the true conditional Sharpe ratio of returns, which is not observable to econometricians and may not be perfectly observable to investors in practice. If investors instead have to use past returns, combined with beliefs about model parameters, to forecast future returns, the relevant economic magnitude is the volatility of Sharpe ratios conditional on past returns. We compute this volatility in the steady state, conditioning on an infinite history of past returns to forecast next-period returns. These forecasts remain imperfect in the steady state because the current μ_t is unobservable and can only by imperfectly filtered from past return observations.

The last set of columns in Panel B in Table 5 show that these conditional Sharpe ratios can vary substantially as the priors about H increase and allow for more persistence in returns. However, despite the relatively wide range of priors for σ_{sr}, the ability to predict portfolio returns using past returns is somewhat limited, with the prior mean of σ(Cond. Sharpe) around half the prior mean of σ_{sr}.

4.2. Posteriors for unconditional expected returns and Sharpe ratios

We estimate the posterior distribution for each of 16 priors (four for H and μ_{sr} each) and each of four characteristic-based portfolios, making 64 prior-data pairs. We generate

50,000 posterior draws for each prior-data pair using the algorithm detailed in Appendix C. We compute the posterior for the full parameter vector Ω. However, as discussed above, we cannot separately identify all five parameters in Ω in our frequentist estimations, which manifests in extremely wide posteriors for ρ and σ_δ in the Bayesian analysis. For each posterior we therefore compute the four identified components of θ, namely μ, σ_r, H, and γ. Also, instead of γ, we present σ(Cond. Sharpe), which is easier to interpret.

Tables 6 through 9 and Figures 1 through 4 report the results for each of the four characteristic-sorted portfolios. As these results show, different priors result in substantially different inferences about unconditional expected returns μ and Sharpe ratios μ_{sr}. For instance, Bayesian investors' posteriors for μ and μ_{sr} are wider as their priors for H increase, making the possibility of zero or negative μ much more likely. The intuition is the same as for the frequentist analysis above: the data do not strongly reject the possibility that unconditional expected returns are zero or negative and that the historical performance is explained by persistent but dissipating positive shocks to conditional expected returns.

The Bayesian analysis also produces an insight that is distinct from the frequentist analysis: the extent to which priors about μ_{sr} affect posteriors about μ and μ_{sr} depends on the investor's prior about H. Bayesian investors who believe $H = 0$ largely agree about μ and μ_{sr} despite large differences in priors. On the other hand, Bayesian investors whose priors are that H is, or might be, large have substantial differences in their posterior beliefs about μ and μ_{sr} despite observing 56 years of data. For example, means of posteriors about the value portfolio's μ_{sr} are clustered between 0.25 and 0.33 for the $H = 0$ prior, but vary from 0.14 to 0.30 for the $H \sim U(0, 10)$ prior.

The reason priors about H matter for posteriors about μ is that Bayesian investors 'shrink' observed in-sample averages towards the mean of their prior, and the extent of this adjustment depends on H. If the investor believes H equals zero and thus returns are i.i.d., the data are more informative about unconditional expected returns and thus the posterior hews closer to the in-sample average. If the investor believes H is or may be larger than zero, the data are less informative and so the posterior depends more on their prior. As a result, the posterior Sharpe ratios vary more across priors for μ_{sr} (rows in Figures 1 through 4) when priors for H are larger.

Tables 6 through 9 also show that the posterior distributions for H and σ(Cond. Sharpe) differ very little from the corresponding prior distributions for the value, investment, and profitability portfolios, which is consistent with the evidence in Tables 2 and 4 that the data offer little guidance on the autocorrelation structure for these portfolios. The size portfolio, by contrast, has posterior distributions for H that are farther from zero than the prior and have means below the prior mean of five years, indicating the data push Bayesian investors towards lower H.[25] Furthermore, the mean posterior σ(Cond. Sharpe) is much higher for size than the other portfolios, and unlike the other portfolios its confidence interval excludes zero, meaning the evidence tilts in favor of positive autocorrelation. Still, posteriors for the size portfolio's H and σ(Cond. Sharpe) are quite wide, leaving room for many potential interpretations of the data.

Just as priors about H affect posteriors about μ_{sr}, priors about μ_{sr} also affect posteriors about H. Investors with bearish priors on μ_{sr} are more amenable to interpreting the observed positive Sharpe ratios as arising from extremely persistent (high H) variations in conditional Sharpe ratios, leading them to tilt their H posteriors higher than investors with positive priors on μ_{sr} that are more comfortable with large unconditional Sharpe ratios explaining the observed data. Tables 6 and 9 show this is indeed the case: for the portfolios with substantial average returns in our sample (value, investment, and profitability), posteriors about H have higher means for investors with bearish priors about μ_{sr}. This effect is small quantitatively, however, because there is so little power in measuring H.

[25]The prior for H is uniformly distributed between 0 and 10, meaning the 95% confidence interval is $[0.25, 9.75]$.

4.3. Posteriors for conditional Sharpe ratios

In addition to forming inferences about unconditional expected returns and Sharpe ratios, the Bayesian analysis allows us to compute posterior distributions of conditional Sharpe ratios through each period in our historical sample. These conditional Sharpe ratio distributions are computed using the posterior distributions for model parameters based on the full sample of data, and conditioning on the full sample of returns.[26] Figure 5 plots the time-series of posterior means for the four characteristic-sorted portfolios for the 'agnostic' prior specification with $\mu_{sr} \sim U(-2, 2)$ and $H \sim U(0, 10)$. As the figure shows, the variations in conditional Sharpe ratios are economically substantial but plausible, generally varying between 0 and 0.8 on an annualized basis. The exception is size, which, despite having estimates of unconditional Sharpe ratios near zero, has conditional Sharpe ratios varying from -0.75 to above one. These fluctuations arise due to a combination of the large variations in rolling average returns and the posterior beliefs tilted towards strong persistence in Table 9.

Finally, we compute posterior distributions of forward-looking conditional Sharpe ratios for the quarter immediately following the end of our sample period (Q1 of 2020). These posteriors differ from the unconditional posteriors because they use nearby trends in portfolio returns to extrapolate to future performance. When $\gamma > 0$, the extrapolation is positive, meaning that near-term expected returns are higher (lower) than long-term expected returns when recent returns are higher (lower) than the full-sample average. When $\gamma < 0$, the extrapolation is negative, meaning near-term expectations reverse the recent trends.

Panel B of Figures 1 through 4 present the mean and 95% confidence intervals for posterior beliefs about the 2020 conditional Sharpe ratios of each portfolio. As the first rows

[26]Note that even with the full sample of data we do not observe conditional expected returns in each period and instead need to estimate them using realized returns and model parameters. These results should be interpreted as a measure of full-sample economic significance (such as an R^2) rather than feasible Sharpe ratios that inform real-time trading strategies.

shows, conditional and unconditional Sharpe ratios are the same when $H = 0$. When $H > 0$, however, both value and investment have smaller conditional Sharpe ratios than unconditional. For the value portfolio, which had particularly poor recent performance as discussed above, the pessimistic or neutral Bayesian investors believe conditional Sharpe ratios in 2020 are centered near zero and could even be quite negative.

Because the profitability portfolio performed better in recent years than earlier in the sample, we find the opposite effect in Figure 3: posteriors about 2020 Sharpe ratios are *higher* than posteriors about the unconditional Sharpe ratio. As with the bearish view of value and investment, the bullish view for profitability is stronger for larger values of H than smaller ones.

4.4. Out-of-sample forecasts

Our final analyses repeat the Bayesian estimation procedure at the end of each calendar year using only past data to form posterior beliefs about model parameters and the distribution of next-year returns. This expanding-window out-of-sample approach allows us to quantify how different prior about persistence and average returns affect Bayesian learning through the sample and lead to persistent disagreements between investors.

For each calendar year y in our sample, 1963–2019, we calculate posterior beliefs about model parameters $\tilde{\Omega}_y$ based on quarterly data through y given each of the prior beliefs described in Section 4.1.[27] While most out-of-sample forecasting exercises wait for a 'training' period to avoid over-fitting extremely short samples, we start immediately in 1963 after observing only two quarterly observations because Bayesian investors rely on their prior distributions – and do not 'over-fit' – when given so little data.

From each of the 50,000 posterior draws from $\tilde{\Omega}_y$, we compute the posterior distribution

[27]We repeat the exercise once per year instead of each quarter to save computational time. 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 either model parameters or returns.

for next-year returns, which are normally distributed, and draw 10 random observations. Combined, we get 500,000 draws from the posterior distribution of next-year returns $\tilde{r}_{y+1}$. Given these draws, we compute out-of-sample portfolio weights w_y^* allocated to the (zero-cost) characteristic-sorted portfolio for an investor with CRRA utility:

$$w_y^* = \arg\max_w \mathbb{E}_y \left(\frac{(1 + r_{f,y} + w \cdot \tilde{r}_{y+1})^{1-\gamma}}{1-\gamma} \right), \tag{15}$$

where $r_{f,y}$ is the risk-free rate as of year y.

To avoid adding covariance parameters to our Bayesian estimation, we assume that the investor optimizes only over the amount allocated to a single characteristic-sorted portfolio and not all four portfolios simultaneously. Our portfolio weights should therefore be interpreted as economically meaningful summaries of the posterior return distribution rather than practical portfolio weights for an optimal multi-asset portfolio.

Figure 6 illustrates the main results of the out-of-sample analyses by plotting differences in portfolio weights as a function of prior beliefs.[28] The solid lines show how learning about the unconditional expected return μ affects portfolio choices by comparing weights chosen by an investor with a bullish prior ($\mu_{sr} \sim N(0.4, 0.4)$) to those resulting from a neutral prior ($\mu_{sr} \sim N(0.0, 0.4)$), where both investors believe $H = 0$. While the bullish investor initially puts much more weight in each characteristic-sorted portfolio, this difference gradually dissipates over time as more data becomes available and the investors place less weight on their priors.

The dotted lines in Figure 6 show how differences in H lead to differences in the aggressiveness of market timing by comparing weights chosen by an investors allowing for variations in conditional expected return ($H \sim U(0, 10)$) to those to those resulting from a prior that returns are independent ($H = 0$), where both investors have neutral beliefs about μ. Because the $H \sim U(0, 10)$ investor uses recent past returns to 'time' the characteristic-sorted

[28]Plots illustrating the portfolio weights over time for all sixteen priors we study are in the Appendix.

portfolio, while the $H = 0$ investor does not, there are large variations in differences between these investors' portfolio weights. Further, these differences are not dissipating over time because, as discussed above, the data offer very little guidance on the true value of H.

We also compute the returns of the out-of-sample portfolio chosen by each investor we study, and tabulate the realized Sharpe ratio and realized certainty equivalent of these return series in Appendix Table 3, in each case including the 95% confidence interval for the historical average of these realized returns. We find that out-of-sample performance is very imprecisely estimated, with the 95% confidence intervals including the full range of means we find across different priors. With this in mind, differences in realized performance across priors about μ_{sr} (more bullish performed better) and H (inconsistent across portfolios) should not be interpreted as conclusive.

5. Conclusion

Time-variation in expected returns is an overlooked but important source of uncertainty about the long-term performance of characteristic-based portfolios. The uncertainty is both direct – the data offer little guidance on the magnitude and persistence of expected return shocks – and indirect – the fact that there may be persistent variations in conditional expected returns causes the data to be much less informative about unconditional expected returns. Because the data are less informative, investors with different priors can have substantially different posterior views despite observing close to 60 years of return data.

Our results have implications for practitioners and academics interested in characteristic-sorted portfolios. Financial institutions now offer a multitude of relatively passive investment products, such as ETFs and mutual funds, that aim to exploit the long-term links between returns and characteristics that have been identified in academic research. At the same time, there exist active hedge funds that attempt to time the variations in characteristic return premia that are described in the more recent literature. Our Bayesian analysis offers guidance for both groups as to how much they should use recent performance to guide their decisions, as well as the near- and medium-term performance outlooks for these portfolios.

Time-variation in expected returns could also offer a potential explanation for the ever-increasing number of characteristics or factors that are shown to predict returns (McLean and Pontiff (2016), Harvey, Liu, and Zhu (2016)). The ease with which finance research uncovers such anomalies raises doubts about whether the documented return patterns point to genuine return premia. We offer a complementary explanation: historical returns may reflect economic or behavioral conditions that persist for relatively long periods of time but not indefinitely.

References

Altı, Aydoğan and Sheridan Titman, 2019, A dynamic model of characteristic-based return predictability, *The 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.

Asness, Cliff, John Liew, Lasse Heje Pedersen, and Ashwin Thapar, 2021, Deep value, *The Journal of Portfolio Management* 47, 11–40.

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

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

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

Brown, Philip, Allan W Kleidon, and Terry A Marsh, 1983, New evidence on the nature of size-related anomalies in stock prices, *Journal of Financial Economics* 12, 33–56.

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

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.

Ehsani, Sina and Juhani T Linnainmaa, 2021, Factor momentum and the momentum factor, *The Journal of Finance* Forthcoming.

Eisfeldt, Andrea L, Edward Kim, and Dimitris Papanikolaou, 2022, Intangible value, *Critical Finance Review* Forthcoming.

Fama, Eugene F and Kenneth R French, 1992, The cross-section of expected stock returns, *The 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, *The Review of Asset Pricing Studies* 11, 105–121.

Ferson, Wayne E, Sergei Sarkissian, and Timothy T Simin, 2003, Spurious regressions in financial economics?, *The 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, 2021, The fundamental-to-market ratio and the value premium decline, SSRN working paper 3573444 .

Guidolin, Massimo and Allan Timmermann, 2008, Size and value anomalies under regime shifts, *Journal of Financial Econometrics* 6, 1–48.

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, *The Review of Financial Studies* 29, 5–68.

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

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

Kandel, Shmuel and Robert F Stambaugh, 1996, On the predictability of stock returns: an asset-allocation perspective, *The 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.

Lewellen, Jonathan, 2002, Momentum and autocorrelation in stock returns, *The 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?, *The 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, *The Journal of Finance* 55, 179–223.

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

Pástor, L'uboš and Robert F Stambaugh, 2012, Are stocks really less volatile in the long run?, *The 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* Forthcoming.

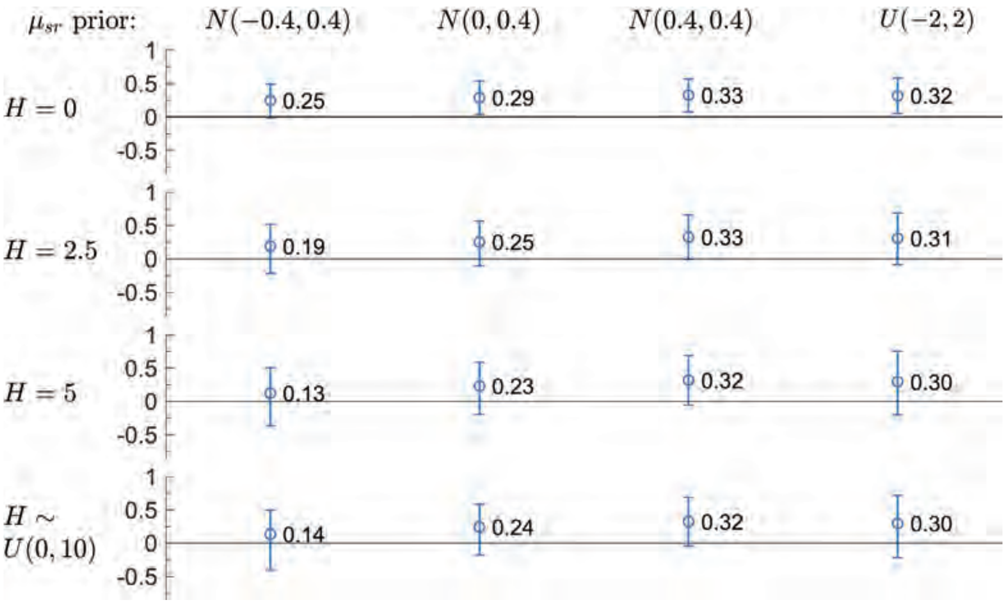
Wachter, Jessica A and Missaka Warusawitharana, 2009, Predictable returns and asset allocation: Should a skeptical investor time the market?, *Journal of Econometrics* 148, 162–178.

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

Figure 1: Posterior Sharpe Ratios for Value Portfolio

This figure presents posterior beliefs for the Sharpe ratios of a portfolio formed by sorting stocks on value, as defined in the header of Table 1. Panel A presents beliefs about the unconditional Sharpe ratio, Panel B the conditional Sharpe ratio for Q1 of 2020. Each panel shows posteriors for twelve possible priors, combining three possible priors for the unconditional Sharpe ratio, μ_{sr}, with four possible priors for the half-life of shocks to conditional Sharpe ratio, H. All Sharpe ratios and H are annualized. Our sample consists of 226 quarterly observations from Q3 1963 through Q4 2019.

Panel A: Unconditional Sharpe Ratio



Panel B: Sharpe Ratio for 2020 Conditional on 1963–2019 Returns

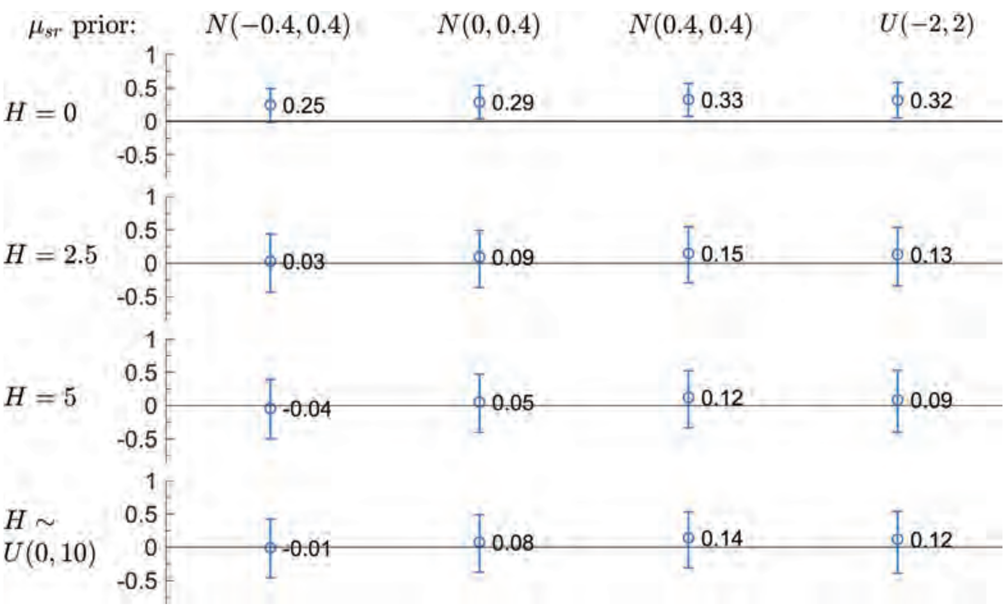
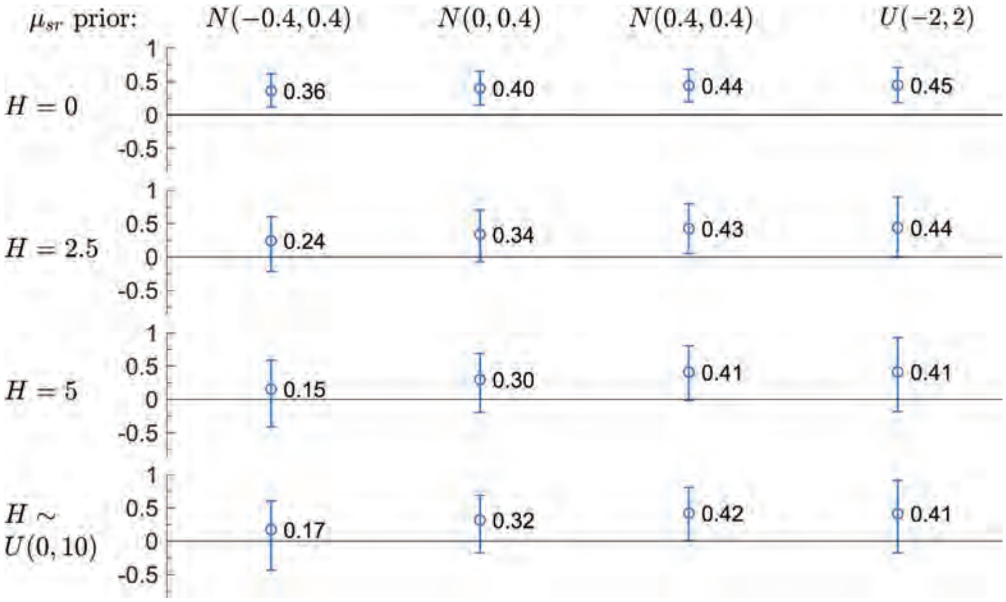


Figure 2: Posterior Sharpe Ratios for Investment Portfolio

This figure presents posterior beliefs for the Sharpe ratios of a portfolio formed by sorting stocks on investment, as defined in the header of Table 1. Panel A presents beliefs about the unconditional Sharpe ratio, Panel B the conditional Sharpe ratio for Q1 of 2020. Each panel shows posteriors for twelve possible priors, combining three possible priors for the unconditional Sharpe ratio, μ_{sr}, with four possible priors for the half-life of shocks to conditional Sharpe ratio, H. All Sharpe ratios and H are annualized. Our sample consists of 226 quarterly observations from Q3 1963 through Q4 2019.

Panel A: Unconditional Sharpe Ratio



Panel B: Sharpe Ratio for 2020 Conditional on 1963–2019 Returns

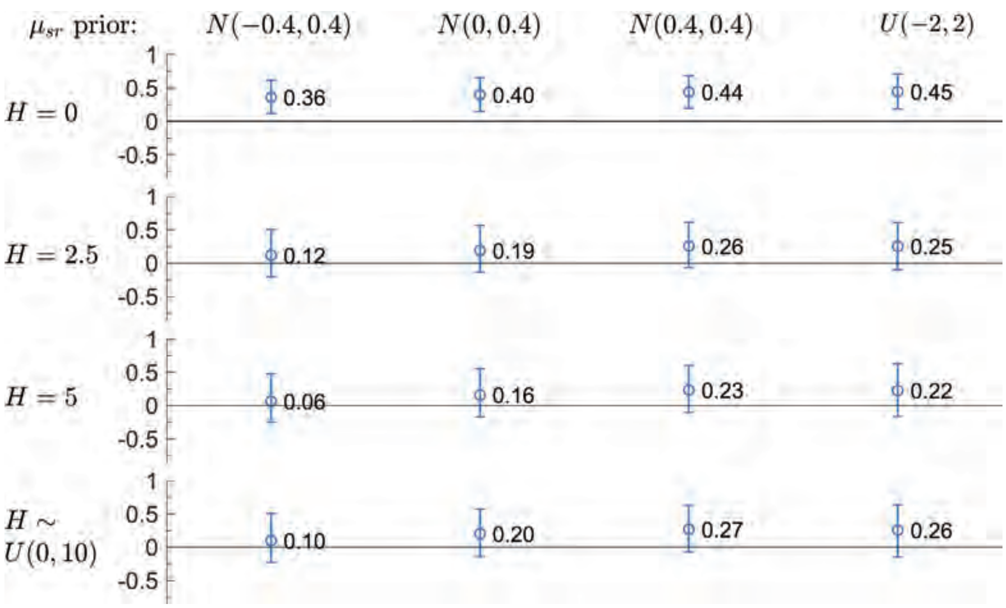
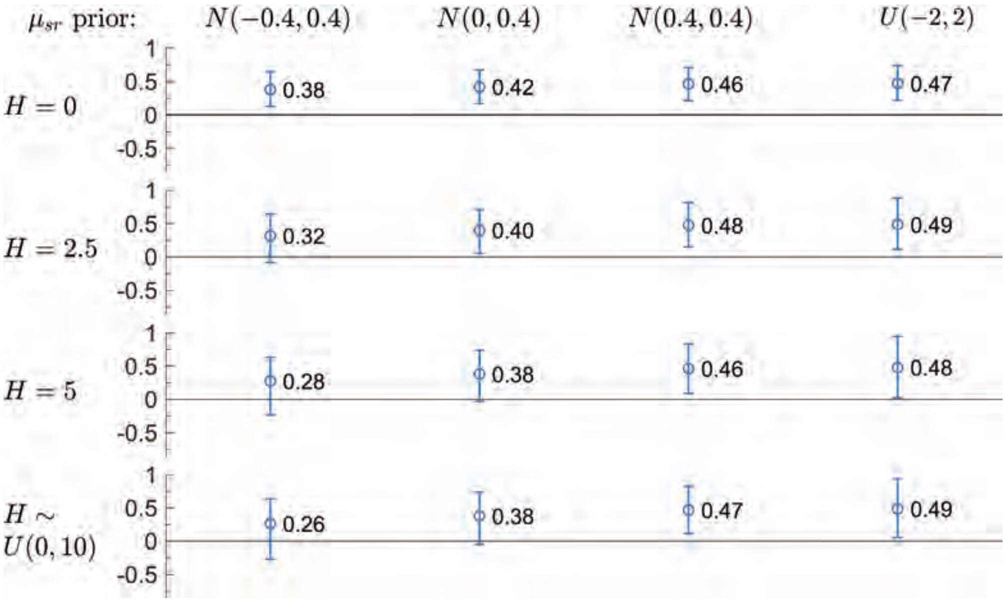


Figure 3: Posterior Sharpe Ratios for Profitability Portfolio

This figure presents posterior beliefs for the Sharpe ratios of a portfolio formed by sorting stocks on profitability, as defined in the header of Table 1. Panel A presents beliefs about the unconditional Sharpe ratio, Panel B the conditional Sharpe ratio for Q1 of 2020. Each panel shows posteriors for twelve possible priors, combining three possible priors for the unconditional Sharpe ratio, μ_{sr}, with four possible priors for the half-life of shocks to conditional Sharpe ratio, H. All Sharpe ratios and H are annualized. Our sample consists of 226 quarterly observations from Q3 1963 through Q4 2019.

Panel A: Unconditional Sharpe Ratio



Panel B: Sharpe Ratio for 2020 Conditional on 1963–2019 Returns

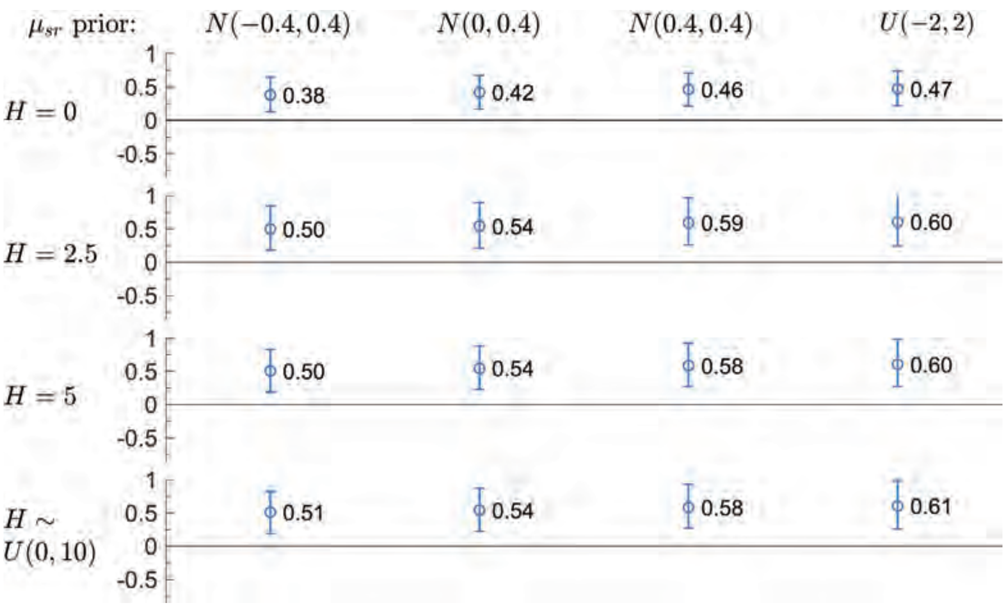
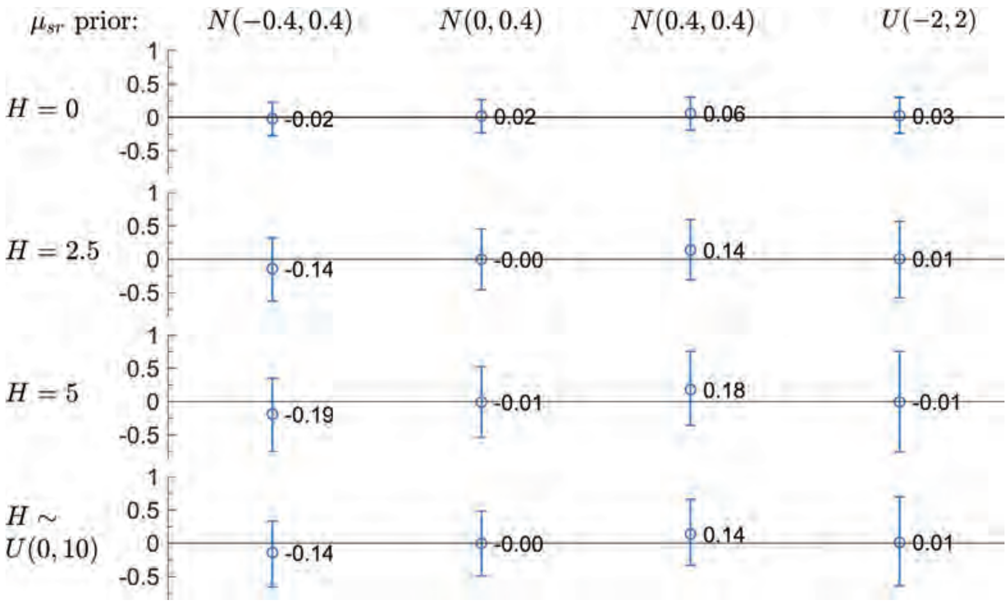


Figure 4: Posterior Sharpe Ratios for Size Portfolio

This figure presents posterior beliefs for the Sharpe ratios of a portfolio formed by sorting stocks on size, as defined in the header of Table 1. Panel A presents beliefs about the unconditional Sharpe ratio, Panel B the conditional Sharpe ratio for Q1 of 2020. Each panel shows posteriors for twelve possible priors, combining three possible priors for the unconditional Sharpe ratio, μ_{sr}, with four possible priors for the half-life of shocks to conditional Sharpe ratio, H. All Sharpe ratios and H are annualized. Our sample consists of 226 quarterly observations from Q3 1963 through Q4 2019.

Panel A: Unconditional Sharpe Ratio



Panel B: Sharpe Ratio for 2020 Conditional on 1963–2019 Returns

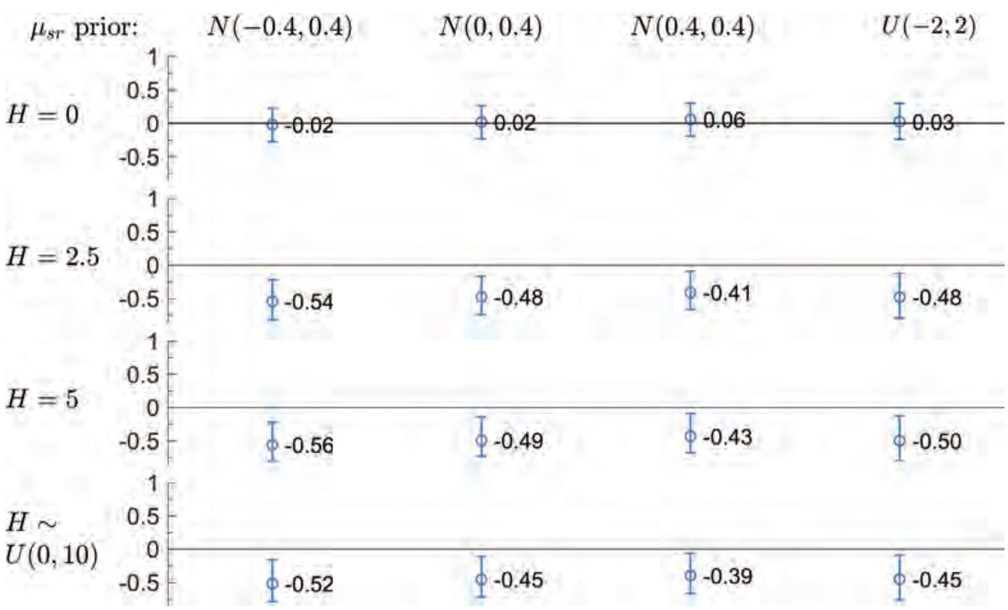
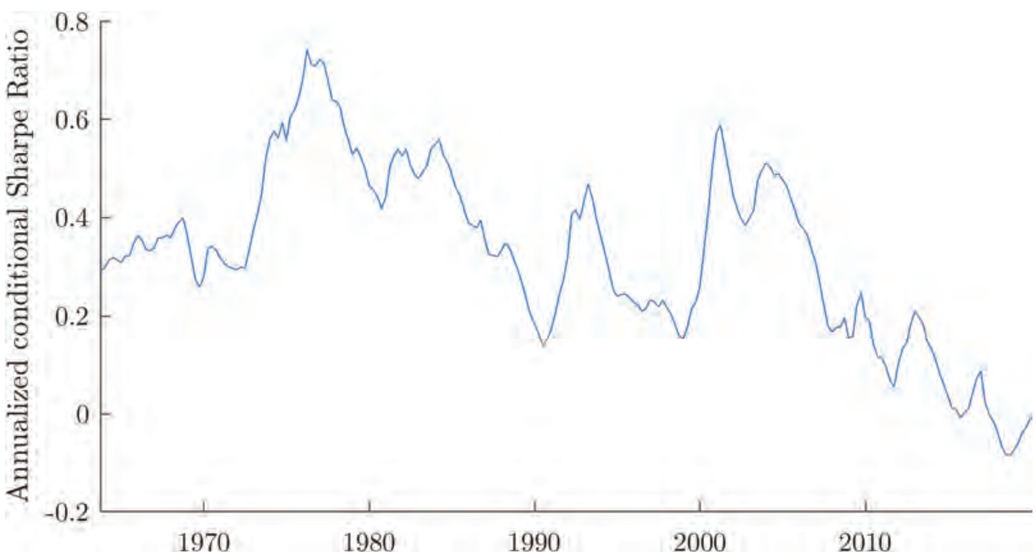


Figure 5: Posteriors on Conditional Sharpe Ratios Across Time

This figure presents posterior beliefs for Sharpe ratios of value, investment, profitability, and size portfolios, as defined in the header of Table 1. Posterior distributions are formed using the prior that $\mu_{sr} \sim N(0, 0.4)$ and $H \sim U(0, 10)$) and the full sample of data. For each possible parameterization, we compute the expected value of conditional expected returns in each quarter of the sample using the full sample of realized returns. We then plot the average of this quantity across posterior draws for the value portfolio in Panel A, investment in Panel B, profitability in Panel C, and size in Panel D. Our sample consists of 226 quarterly observations from Q3 1963 through Q4 2019.

Panel A: Value



Panel B: Investment

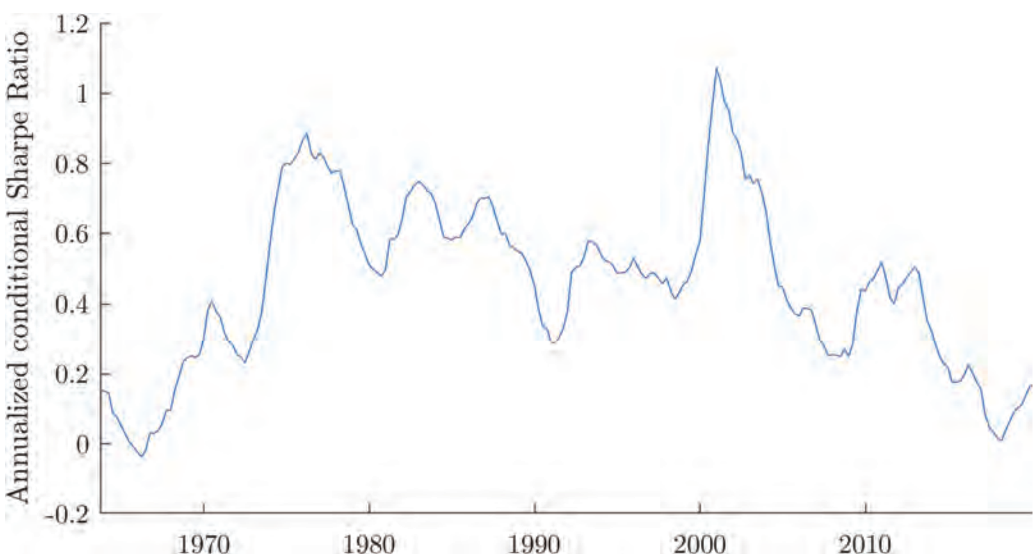
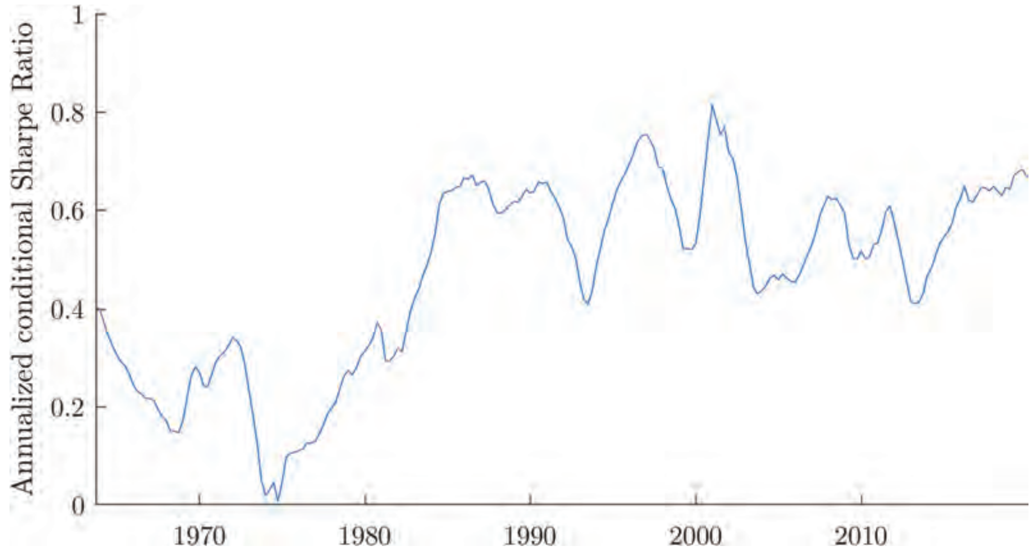


Figure 5: Posteriors on Conditional Sharpe Ratios Across Time (Continued)

Panel C: Profitability



Panel D: Size

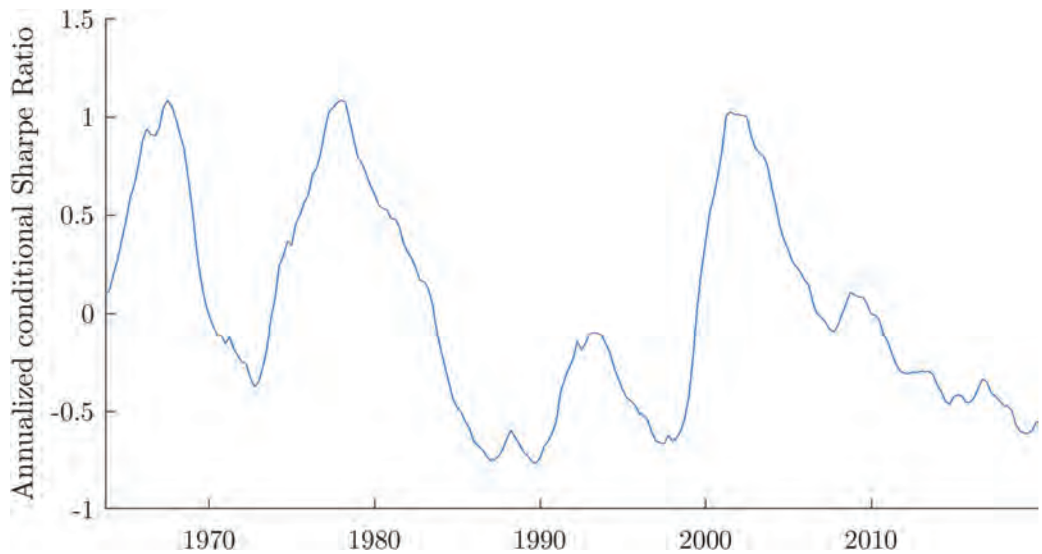
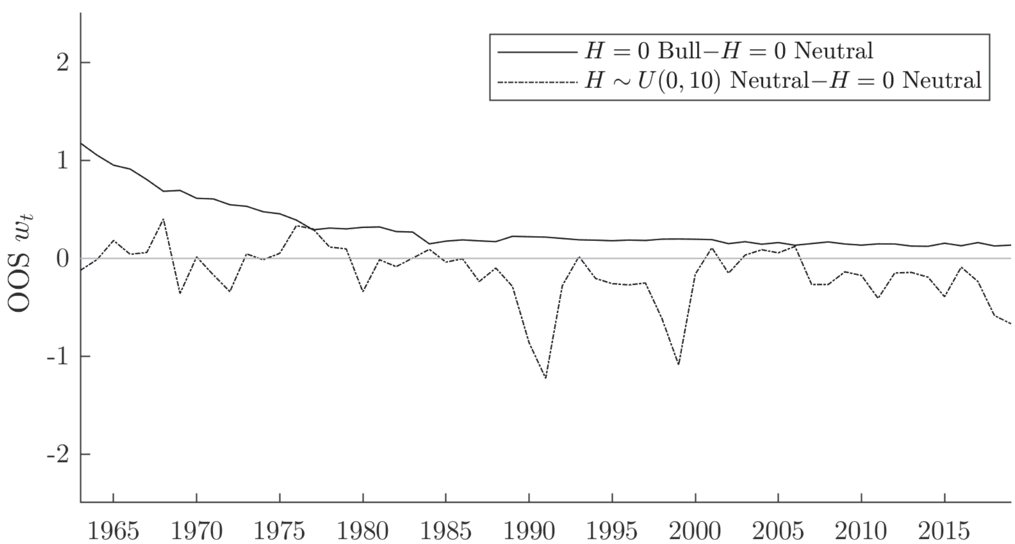


Figure 6: Out-of-Sample Portfolio Choices

This figure presents differences in portfolio weights w_t chosen by Bayesian investors with different priors who calculate posteriors using only past data. The solid line shows the difference in w_t chosen by an investor with priors that the half-lives of expected return shocks equal zero ($H = 0$) and that the unconditional Sharpe ratios has distribution $\mu_{sr} \sim N(0.4, 0.4)$ ('Bull'), and an investor with priors $H = 0$ and $\mu_{sr} \sim N(0, 0.4)$ ('Neutral'). The dotted line shows the difference in weights chosen by an investor with priors $H \sim U(0, 10)$ and $\mu_{sr} \sim N(0, 0.4)$, and an investor with priors $H = 0$ and $\mu_{sr} \sim N(0, 0.4)$. For each prior and characteristic-sorted portfolio, we compute the posterior distribution of next-year returns and the corresponding optimal portfolio weights in the risk-free asset and characteristic-sorted portfolio for an investor with CRRA utility and relative risk aversion of 2.

Panel A: Value



Panel B: Investment

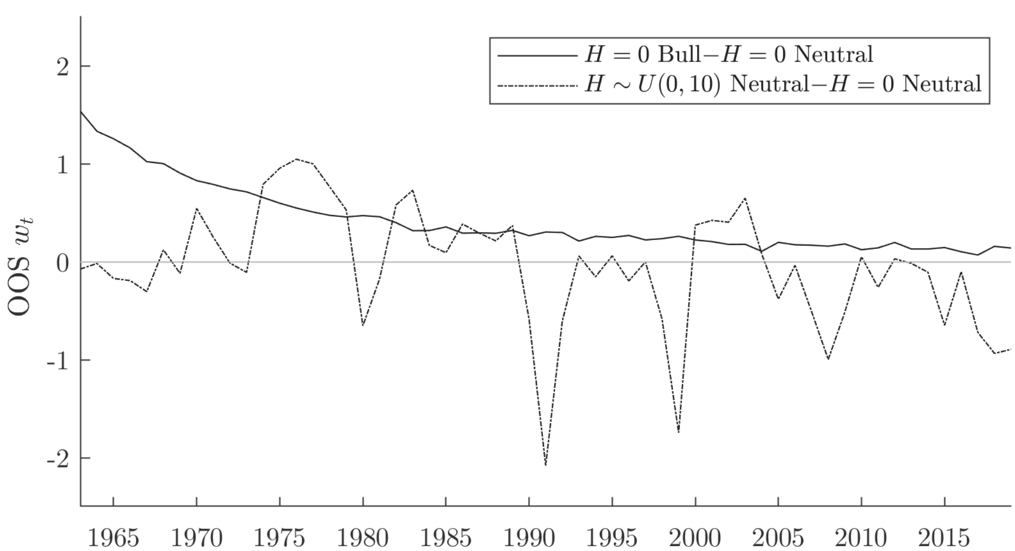
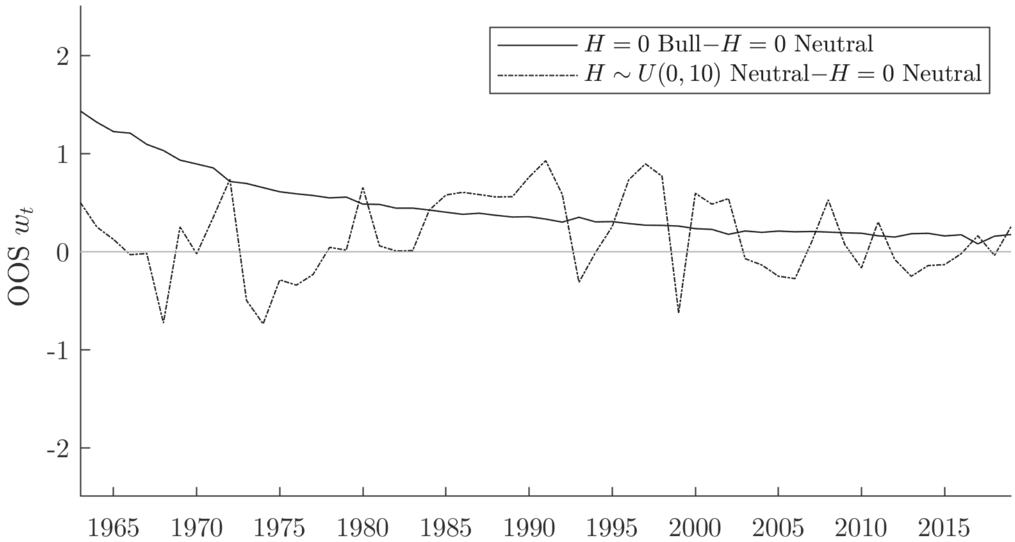


Figure 6: Out-of-Sample Portfolio Choices (Continued)

Panel C: Profitability



Panel D: Size

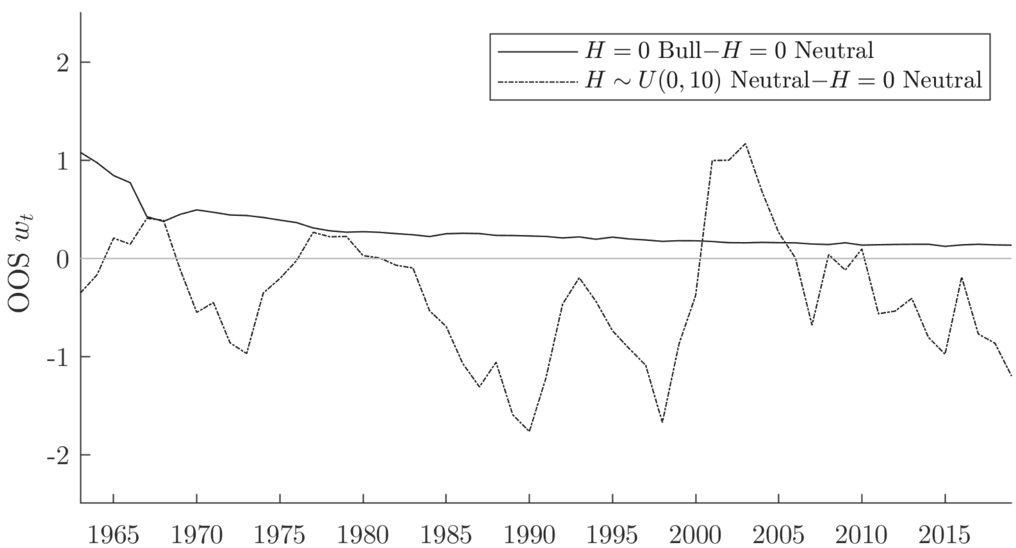


Table 1: Historical Performance of Characteristic-Sorted Portfolios

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. We present the mean annualized quarterly return and annualized Sharpe ratio in the full sample, two subsamples, and for the difference between the subsamples. t-statistics based on iid re-sampling of the calendar quarters in our sample are in parenthesis. Our sample consists of 226 quarterly observations from Q3 1963 through Q4 2019.

	Mean (annualized %)				Sharpe Ratio (annualized)			
	All	1963–1991	1992–2019	Diff	All	1963–1991	1992–2019	Diff
Value	4.37	5.52	3.21	-2.30	0.32	0.42	0.22	-0.20
	(2.39)	(2.14)	(1.23)	-(0.63)	(2.37)	(2.23)	(1.15)	-(0.75)
Investment	4.72	4.25	5.20	0.95	0.44	0.44	0.45	0.02
	(3.34)	(2.14)	(2.58)	(0.34)	(3.31)	(2.30)	(2.35)	(0.06)
Profitability	4.91	3.51	6.33	2.82	0.47	0.36	0.57	0.20
	(3.51)	(1.77)	(3.21)	(1.01)	(3.48)	(1.89)	(2.96)	(0.75)
Size	0.35	1.66	-0.99	-2.66	0.02	0.10	-0.07	-0.17
	(0.17)	(0.59)	-(0.35)	-(0.66)	(0.17)	(0.54)	-(0.37)	-(0.64)

Table 2: Autocorrelations of Characteristic-Sorted Portfolios

This table presents statistics summarizing the autocorrelation of value, investment, profitability, and size portfolios, as defined in the header of Table 1. For each characteristic-sorted portfolio, we run overlapping time-series regressions of quarterly returns r_t on a constant and rolling averages of past quarterly returns from months $t-L$ through $t-1$:

$$r_t = a + b_L \left(\frac{1}{L} \sum_{l=1}^{L} r_{t-l} \right) + \epsilon_t, \tag{16}$$

Panel A presents average coefficients b_L in 226-quarter samples simulated under various parameterizations of the model. The parameter γ is the first-order autocorrelation of returns implied by each parameterization, and H is the half-life of shocks to expected returns (in years). Panel B show estimates for the market-neutral portfolios. Standard errors based on iid re-sampling of the calendar months in our sample are in parenthesis. The joint significance row presents the fraction of i.i.d. simulations for which the sum of the $\hat{b}$ across the four portfolios exceeds the sum in observed data. Our sample consists of 226 quarterly observations from Q3 1963 through Q4 2019.

Panel A: Model Simulations

Parameterization			Avg. Coefficients [95% Conf. Interval]		
H	γ	σ(Cond. Sharpe)	b_4	b_{20}	b_{40}
IID		0.00	-0.03	-0.20	-0.49
			[-0.32,0.22]	[-1.08,0.38]	[-2.14,0.45]
2.5	2.5%	0.12	0.04	0.01	-0.21
			[-0.24,0.30]	[-0.79,0.53]	[-1.65,0.61]
2.5	5%	0.22	0.11	0.13	-0.05
			[-0.19,0.36]	[-0.62,0.61]	[-1.38,0.69]
2.5	10%	0.38	0.22	0.30	0.12
			[-0.08,0.48]	[-0.36,0.71]	[-1.02,0.78]
5	2.5%	0.15	0.04	0.02	-0.15
			[-0.25,0.29]	[-0.79,0.56]	[-1.61,0.66]
5	5%	0.26	0.10	0.17	0.03
			[-0.20,0.37]	[-0.61,0.66]	[-1.31,0.76]
10	2.5%	0.12	0.09	0.02	-0.00
			[-0.27,0.28]	[-0.86,0.56]	[-1.71,0.68]
10	5%	0.20	0.15	0.07	0.14
			[-0.23,0.35]	[-0.68,0.67]	[-1.41,0.80]

Panel B: Estimates for historical data

	b_4	b_{20}	b_{40}
Value	0.19	0.05	0.24
Investment	0.21	-0.29	-0.73
Profitability	0.21	-0.24	0.16
Size	0.47	0.20	0.34
Pooled	0.30	0.06	0.18
iid p-value	0.0%	24.5%	4.7%
Pooled (without size)	0.20	-0.13	-0.01
iid p-value	0.1%	45.1%	15.2%

Table 3: OLS 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, along with t-statistics calculated using a variety of approaches. The 'unadjusted t-stat' row gives the typical OLS standard error, which assumes returns are independently distributed. The next three rows give standard errors calculated using Newey and West (1987) with 10, 20, or 40 quarterly lags. The remaining rows give OLS standard errors corrected for autocorrelation using Equation (13) with different values of H, the half-life of shocks to expected returns (in years), and γ, the first-order autocorrelation of realized returns. Our sample consists of 226 quarterly observations from Q3 1963 through Q4 2019.

	Value	Investment	Profitability	Size
$\hat{\mu}$	4.37	4.72	4.91	0.35
Unadjusted t-stat	(2.39)	(3.34)	(3.54)	(0.17)
Newey-West t-stat (10 lags)	(2.35)	(2.90)	(3.33)	(0.12)
Newey-West t-stat (20 lags)	(2.35)	(3.04)	(3.71)	(0.11)
Newey-West t-stat (40 lags)	(2.41)	(3.23)	(3.48)	(0.12)
Model t-stat ($H = 2.5$, $\gamma = 2.5\%$)	(1.83)	(2.56)	(2.71)	(0.13)
Model t-stat ($H = 2.5$, $\gamma = 5\%$)	(1.54)	(2.15)	(2.28)	(0.11)
Model t-stat ($H = 2.5$, $\gamma = 10\%$)	(1.23)	(1.71)	(1.81)	(0.09)
Model t-stat ($H = 5$, $\gamma = 2.5\%$)	(1.58)	(2.21)	(2.34)	(0.11)
Model t-stat ($H = 5$, $\gamma = 5\%$)	(1.26)	(1.77)	(1.87)	(0.09)
Model t-stat ($H = 10$, $\gamma = 2.5\%$)	(1.34)	(1.87)	(1.98)	(0.10)
Model t-stat ($H = 10$, $\gamma = 5\%$)	(1.03)	(1.44)	(1.53)	(0.07)

Table 4: Maximum Likelihood Hypothesis Tests

This table presents parameter estimates and hypothesis tests based on maximum-likelihood estimates of our model for value, investment, profitability, and size portfolios, as defined in the header of Table 1. Panel A presents estimates of μ, the unconditional expected return; σ_r, the unconditional standard deviation of returns; H, the half-life of time-variations in conditional mean returns; and γ, the first-order autocorrelation of factor returns. μ, σ_r, and H are all annualized. The rows labelled $\mu = 0$ re-estimate the model with μ restricted to zero, and give a 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. For each restriction, we present the estimated γ, and likelihood-ratio p-values for both the H and γ restrictions (Rest. p-value) and $\mu = 0$ given the H and γ restrictions. Our sample consists of 226 quarterly observations from Q3 1963 through Q4 2019.

Panel A: Tests for $\mu = 0$

	IID			Time-varying means				
	μ (%)	σ_r (%)	p-value (%)	μ (%)	σ_r (%)	H (years)	γ (%)	p-value (%)
Value	4.37	13.76		4.60	13.75	14.87	-0.58	
$\mu = 0$	0.00	13.93	1.8%	0.00	13.90	20.00	2.73	10.2%
Investment	4.72	10.62		5.19	10.61	9.38	-0.92	
$\mu = 0$	0.00	10.88	0.1%	0.00	10.82	14.66	5.25	5.7%
Profitability	4.91	10.42		4.96	10.42	20.00	-0.43	
$\mu = 0$	0.00	10.70	0.0%	0.00	10.62	20.00	4.46	5.7%
Size	0.35	15.02		0.31	15.02	1.33	15.44	
$\mu = 0$	0.00	15.03	86.3%	0.00	15.04	2.16	14.84	37.8%

Panel B: Restrictions on H and γ

Restrictions	H (years):		2.5			5			10		
	γ (%):	IID	-	2.5	5	-	2.5	5	-	2.5	5
Value	Rest. p-value (%)	45.2	45.2	60.7	40.1	46.0	51.4	31.2	73.9	46.1	28.4
	$\mu = 0$ p-value (%)	1.8	7.1	8.1	15.4	9.5	14.9	27.6	9.6	24.5	42.0
Investment	Rest. p-value (%)	30.3	41.3	71.5	65.6	30.3	51.6	40.0	95.1	39.1	28.6
	$\mu = 0$ p-value (%)	0.1	6.3	1.5	4.5	8.9	4.5	13.0	5.5	11.2	27.4
Profitability	Rest. p-value (%)	54.8	55.8	68.8	44.8	55.9	61.5	35.5	54.8	58.1	34.5
	$\mu = 0$ p-value (%)	0.0	1.1	0.6	2.1	2.0	1.9	5.6	4.2	4.9	12.3
Size	Rest. p-value (%)	0.0	29.2	2.2	8.3	5.3	1.0	2.6	1.2	0.5	0.8
	$\mu = 0$ p-value (%)	86.3	100.0	92.2	97.5	100.0	93.1	100.0	100.0	92.6	96.8

Table 5: Priors

This table presents summary statistics for the variety of prior distributions we use for our Bayesian estimates. Panel A lists the possible priors we consider for μ_{sr}, the annualized unconditional Sharpe ratio; H, the half life of variations in conditional Sharpe ratios in years; σ_r, annualized unconditional return volatility; σ_{sr}, the annualized volatility of conditional Sharpe ratios; and ρ, the correlation between shocks to realized returns and conditional Sharpe ratios. N indicates a normal distribution, U indicates a uniform distribution, and a number indicates a dogmatic prior that the parameter equals a that value. Panel B presents the means and 95% confidence intervals for prior distributions of μ, the annualized unconditional expected return; γ, the first-order autocorrelation of returns; and σ(Cond. Sharpe), the standard deviation across time of Sharpe ratios conditional on all past realized returns. The conditional Sharpe Ratio distribution is based on the joint return distributions implied by each model parameterization and not observed data.

Panel A: Priors on Transformed Parameters

H	μ_{sr}	σ_r (%)	σ_{sr}	ρ
0	$N(-0.4, 0.4)$	$U(10, 20)$	$U(0, 1)$	$U(-0.5, 0)$
2.5	$N(0, 0.4)$			
5	$N(0.4, 0.4)$			
$U(0, 10)$	$U(-2, 2)$			

Panel B: Moments of Priors

		μ (%)		γ (%)		σ(Cond. Sharpe)	
H	μ_{sr}	mean	95% CI	mean	95% CI	mean	95% CI
0	$N(-0.4, 0.4)$	-5.30	[-15.90,5.16]	-	-	-	-
0	$N(0, 0.4)$	-0.01	[-10.47,10.42]	-	-	-	-
0	$N(0.4, 0.4)$	5.32	[-5.10,16.05]	-	-	-	-
0	$U(-2, 2)$	-0.02	[-25.84,25.70]	-	-	-	-
2.5	$N(-0.4, 0.4)$	-5.34	[-16.07,5.07]	4.81	[-0.58,15.44]	0.19	[0.00,0.52]
2.5	$N(0, 0.4)$	-0.04	[-10.42,10.37]	4.76	[-0.59,15.41]	0.19	[0.00,0.52]
2.5	$N(0.4, 0.4)$	5.29	[-5.10,15.98]	4.78	[-0.58,15.43]	0.19	[0.00,0.52]
2.5	$U(-2, 2)$	-0.05	[-25.88,25.72]	4.76	[-0.58,15.45]	0.19	[0.00,0.52]
5	$N(-0.4, 0.4)$	-5.30	[-15.94,5.11]	5.58	[-0.25,16.49]	0.25	[0.00,0.61]
5	$N(0, 0.4)$	0.01	[-10.47,10.47]	5.60	[-0.26,16.47]	0.25	[0.00,0.61]
5	$N(0.4, 0.4)$	5.23	[-5.18,15.94]	5.61	[-0.26,16.54]	0.25	[0.00,0.61]
5	$U(-2, 2)$	0.00	[-25.75,25.93]	5.56	[-0.25,16.41]	0.25	[0.00,0.61]
$U(0, 10)$	$N(-0.4, 0.4)$	-5.32	[-16.03,4.99]	4.91	[-1.91,16.37]	0.23	[0.00,0.63]
$U(0, 10)$	$N(0, 0.4)$	-0.02	[-10.48,10.38]	4.90	[-1.79,16.37]	0.23	[0.00,0.63]
$U(0, 10)$	$N(0.4, 0.4)$	5.29	[-5.14,15.98]	4.92	[-1.76,16.40]	0.23	[0.00,0.63]
$U(0, 10)$	$U(-2, 2)$	-0.01	[-25.81,25.82]	4.90	[-1.83,16.38]	0.23	[0.00,0.63]

Table 6: Posteriors for Value Portfolio

This table presents summary statistics for the posterior distributions of our model's parameters given data for the value portfolio, as described in the header of Table 1. μ is the unconditional mean return, in annualized percent. σ_r is the unconditional volatility of returns, in annualized percent. H is the half life of variations in conditional Sharpe ratios, in years. γ is the first-order autocorrelation in realized returns implied by the model's parameters, in percent. The twelve possible priors are summarized in Table 5. We present the mean and 95% confidence intervals for each parameter and prior. We estimate the model using an M.C.M.C. procedure described in the Appendix. Our sample consists of 226 quarterly observations from Q3 1963 through Q4 2019.

Prior		μ		σ_r		H		σ(Cond. Sharpe)	
H	μ_{sr}	mean	95% CI	mean	95% CI	mean	95% CI	mean	95% CI
0	$N(-0.4, 0.4)$	3.41	[-0.10,6.73]	12.69	[12.69,14.99]	-	-	-	-
0	$N(0, 0.4)$	3.94	[0.55,7.37]	12.62	[12.62,14.95]	-	-	-	-
0	$N(0.4, 0.4)$	4.47	[1.04,7.89]	12.58	[12.58,14.98]	-	-	-	-
0	$U(-2, 2)$	4.35	[0.72,8.02]	12.56	[12.56,14.98]	-	-	-	-
2.5	$N(-0.4, 0.4)$	2.59	[-2.95,6.92]	12.74	[12.74,15.30]	-	-	0.14	[0.00,0.46]
2.5	$N(0, 0.4)$	3.45	[-1.40,7.70]	12.70	[12.70,15.25]	-	-	0.13	[0.00,0.45]
2.5	$N(0.4, 0.4)$	4.42	[-0.12,8.84]	12.69	[12.69,15.26]	-	-	0.13	[0.00,0.44]
2.5	$U(-2, 2)$	4.18	[-1.09,9.19]	12.67	[12.67,15.27]	-	-	0.13	[0.00,0.45]
5	$N(-0.4, 0.4)$	1.72	[-4.93,6.81]	12.74	[12.74,15.36]	-	-	0.19	[0.00,0.56]
5	$N(0, 0.4)$	3.16	[-2.65,7.89]	12.72	[12.72,15.32]	-	-	0.16	[0.00,0.54]
5	$N(0.4, 0.4)$	4.33	[-0.75,9.26]	12.69	[12.69,15.32]	-	-	0.15	[0.00,0.52]
5	$U(-2, 2)$	4.03	[-2.68,10.05]	12.71	[12.71,15.33]	-	-	0.17	[0.00,0.54]
$U(0, 10)$	$N(-0.4, 0.4)$	1.86	[-5.33,6.78]	12.78	[12.78,15.38]	5.12	[0.35,9.77]	0.18	[0.00,0.58]
$U(0, 10)$	$N(0, 0.4)$	3.28	[-2.42,7.86]	12.70	[12.70,15.32]	4.84	[0.34,9.75]	0.16	[0.00,0.54]
$U(0, 10)$	$N(0.4, 0.4)$	4.38	[-0.62,9.21]	12.67	[12.67,15.30]	4.78	[0.30,9.73]	0.15	[0.00,0.53]
$U(0, 10)$	$U(-2, 2)$	4.01	[-2.90,9.64]	12.67	[12.67,15.31]	4.80	[0.26,9.74]	0.15	[0.00,0.56]

Table 7: Posteriors for Investment Portfolio

This table presents summary statistics for the posterior distributions of our model's parameters given data for the investment portfolio, as described in the header of Table 1. μ is the unconditional mean return, in annualized percent. σ_r is the unconditional volatility of returns, in annualized percent. H is the half life of variations in conditional Sharpe ratios, in years. γ is the first-order autocorrelation in realized returns implied by the model's parameters, in percent. The twelve possible priors are summarized in Table 5. We present the mean and 95% confidence intervals for each parameter and prior. We estimate the model using an M.C.M.C. procedure described in the Appendix. Our sample consists of 226 quarterly observations from Q3 1963 through Q4 2019.

Prior		μ		σ_r		H		σ(Cond. Sharpe)	
H	μ_{sr}	mean	95% CI	mean	95% CI	mean	95% CI	mean	95% CI
0	$N(-0.4, 0.4)$	3.83	[1.27,6.56]	9.76	[9.76,11.61]	-	-	-	-
0	$N(0, 0.4)$	4.24	[1.60,6.92]	9.71	[9.71,11.59]	-	-	-	-
0	$N(0.4, 0.4)$	4.66	[2.06,7.26]	9.73	[9.73,11.56]	-	-	-	-
0	$U(-2, 2)$	4.74	[2.01,7.51]	9.75	[9.75,11.55]	-	-	-	-
2.5	$N(-0.4, 0.4)$	2.51	[-2.25,6.26]	9.87	[9.87,11.84]	-	-	0.22	[0.00,0.51]
2.5	$N(0, 0.4)$	3.49	[-0.76,7.15]	9.81	[9.81,11.83]	-	-	0.20	[0.00,0.50]
2.5	$N(0.4, 0.4)$	4.41	[0.47,8.11]	9.80	[9.80,11.81]	-	-	0.19	[0.00,0.50]
2.5	$U(-2, 2)$	4.52	[-0.02,9.27]	9.82	[9.82,11.79]	-	-	0.20	[0.00,0.50]
5	$N(-0.4, 0.4)$	1.62	[-4.25,6.07]	9.90	[9.90,11.91]	-	-	0.27	[0.00,0.60]
5	$N(0, 0.4)$	3.11	[-2.03,7.15]	9.83	[9.83,11.87]	-	-	0.23	[0.00,0.58]
5	$N(0.4, 0.4)$	4.27	[-0.20,8.30]	9.81	[9.81,11.83]	-	-	0.21	[0.00,0.58]
5	$U(-2, 2)$	4.28	[-1.88,9.44]	9.83	[9.83,11.87]	-	-	0.23	[0.00,0.59]
$U(0, 10)$	$N(-0.4, 0.4)$	1.81	[-4.54,6.29]	9.86	[9.86,11.92]	5.15	[0.59,9.76]	0.26	[0.00,0.61]
$U(0, 10)$	$N(0, 0.4)$	3.29	[-1.81,7.11]	9.84	[9.84,11.83]	4.55	[0.36,9.66]	0.22	[0.00,0.59]
$U(0, 10)$	$N(0.4, 0.4)$	4.34	[0.04,8.23]	9.79	[9.79,11.84]	4.49	[0.35,9.66]	0.20	[0.00,0.58]
$U(0, 10)$	$U(-2, 2)$	4.25	[-1.84,9.27]	9.81	[9.81,11.83]	4.55	[0.38,9.76]	0.22	[0.00,0.59]

Table 8: Posteriors for Profitability Portfolio

This table presents summary statistics for the posterior distributions of our model's parameters given data for the profitability portfolio, as described in the header of Table 1. μ is the unconditional mean return, in annualized percent. σ_r is the unconditional volatility of returns, in annualized percent. H is the half life of variations in conditional Sharpe ratios, in years. γ is the first-order autocorrelation in realized returns implied by the model's parameters, in percent. The twelve possible priors are summarized in Table 5. We present the mean and 95% confidence intervals for each parameter and prior. We estimate the model using an M.C.M.C. procedure described in the Appendix. Our sample consists of 226 quarterly observations from Q3 1963 through Q4 2019.

Prior		μ		σ_r		H		σ(Cond. Sharpe)	
H	μ_{sr}	mean	95% CI	mean	95% CI	mean	95% CI	mean	95% CI
0	$N(-0.4, 0.4)$	3.98	[1.36,6.70]	9.63	[9.63,11.39]	-	-	-	-
0	$N(0, 0.4)$	4.37	[1.72,6.94]	9.57	[9.57,11.38]	-	-	-	-
0	$N(0.4, 0.4)$	4.81	[2.28,7.40]	9.53	[9.53,11.33]	-	-	-	-
0	$U(-2, 2)$	4.92	[2.27,7.59]	9.54	[9.54,11.36]	-	-	-	-
2.5	$N(-0.4, 0.4)$	3.25	[-0.88,6.65]	9.66	[9.66,11.60]	-	-	0.15	[0.00,0.47]
2.5	$N(0, 0.4)$	4.12	[0.53,7.27]	9.61	[9.61,11.56]	-	-	0.13	[0.00,0.45]
2.5	$N(0.4, 0.4)$	4.88	[1.53,8.22]	9.61	[9.61,11.55]	-	-	0.13	[0.00,0.45]
2.5	$U(-2, 2)$	4.98	[1.18,8.90]	9.63	[9.63,11.56]	-	-	0.13	[0.00,0.46]
5	$N(-0.4, 0.4)$	2.87	[-2.28,6.49]	9.71	[9.71,11.63]	-	-	0.18	[0.00,0.55]
5	$N(0, 0.4)$	3.92	[-0.32,7.51]	9.65	[9.65,11.61]	-	-	0.16	[0.00,0.53]
5	$N(0.4, 0.4)$	4.77	[0.92,8.48]	9.61	[9.61,11.60]	-	-	0.15	[0.00,0.52]
5	$U(-2, 2)$	4.93	[0.23,9.65]	9.62	[9.62,11.61]	-	-	0.16	[0.00,0.54]
$U(0, 10)$	$N(-0.4, 0.4)$	2.71	[-2.78,6.52]	9.67	[9.67,11.67]	5.09	[0.38,9.82]	0.19	[0.00,0.58]
$U(0, 10)$	$N(0, 0.4)$	3.89	[-0.51,7.51]	9.62	[9.62,11.64]	4.77	[0.37,9.75]	0.16	[0.00,0.54]
$U(0, 10)$	$N(0.4, 0.4)$	4.79	[1.19,8.42]	9.60	[9.60,11.57]	4.61	[0.25,9.72]	0.14	[0.00,0.51]
$U(0, 10)$	$U(-2, 2)$	4.96	[0.56,9.47]	9.60	[9.60,11.56]	4.66	[0.44,9.70]	0.15	[0.00,0.53]

Table 9: Posteriors for Size Portfolio

This table presents summary statistics for the posterior distributions of our model's parameters given data for the size portfolio, as described in the header of Table 1. μ is the unconditional mean return, in annualized percent. σ_r is the unconditional volatility of returns, in annualized percent. H is the half life of variations in conditional Sharpe ratios, in years. γ is the first-order autocorrelation in realized returns implied by the model's parameters, in percent. The twelve possible priors are summarized in Table 5. We present the mean and 95% confidence intervals for each parameter and prior. We estimate the model using an M.C.M.C. procedure described in the Appendix. Our sample consists of 226 quarterly observations from Q3 1963 through Q4 2019.

Prior		μ		σ_r		H		σ(Cond. Sharpe)	
H	μ_{sr}	mean	95% CI	mean	95% CI	mean	95% CI	mean	95% CI
0	$N(-0.4, 0.4)$	-1.07	[-4.63,2.46]	13.78	[13.78,16.33]	-	-	-	-
0	$N(0, 0.4)$	0.26	[-3.23,3.76]	13.78	[13.78,16.36]	-	-	-	-
0	$N(0.4, 0.4)$	1.55	[-1.75,4.99]	13.77	[13.77,16.36]	-	-	-	-
0	$U(-2, 2)$	0.38	[-3.59,4.48]	13.77	[13.77,16.33]	-	-	-	-
2.5	$N(-0.4, 0.4)$	-3.15	[-8.57,2.02]	13.81	[13.81,16.55]	-	-	0.42	[0.19,0.56]
2.5	$N(0, 0.4)$	-0.02	[-5.07,5.06]	13.78	[13.78,16.56]	-	-	0.41	[0.17,0.56]
2.5	$N(0.4, 0.4)$	3.21	[-1.91,8.39]	13.78	[13.78,16.59]	-	-	0.41	[0.19,0.57]
2.5	$U(-2, 2)$	0.09	[-7.95,7.92]	13.81	[13.81,16.59]	-	-	0.42	[0.18,0.57]
5	$N(-0.4, 0.4)$	-3.90	[-9.83,1.92]	13.97	[13.97,16.81]	-	-	0.49	[0.19,0.65]
5	$N(0, 0.4)$	-0.05	[-5.76,5.63]	13.96	[13.96,16.78]	-	-	0.47	[0.14,0.64]
5	$N(0.4, 0.4)$	3.83	[-1.94,9.68]	13.95	[13.95,16.76]	-	-	0.48	[0.17,0.64]
5	$U(-2, 2)$	-0.08	[-10.61,10.56]	13.96	[13.96,16.79]	-	-	0.48	[0.18,0.64]
$U(0, 10)$	$N(-0.4, 0.4)$	-3.33	[-9.13,2.05]	13.77	[13.77,16.65]	3.57	[0.80,9.10]	0.42	[0.15,0.65]
$U(0, 10)$	$N(0, 0.4)$	-0.10	[-5.26,5.09]	13.75	[13.75,16.66]	3.25	[0.82,8.94]	0.40	[0.14,0.64]
$U(0, 10)$	$N(0.4, 0.4)$	3.29	[-1.90,9.08]	13.79	[13.79,16.71]	3.51	[0.84,9.11]	0.42	[0.15,0.64]
$U(0, 10)$	$U(-2, 2)$	0.18	[-8.91,9.75]	13.76	[13.76,16.66]	3.59	[0.84,9.06]	0.42	[0.15,0.65]

Appendix A. Autocorrelations in Monthly Returns

As described in Section 2.1, we analyze a sample of quarterly returns for characteristic-based portfolios rather than the monthly returns typically studied in the literature. We do so because monthly returns exhibit strong first-order autocorrelations that may be caused by lead-lag effects, under-reaction, or some other transitory source of persistence. While an interesting topic on its own, this form of autocorrelation is not the topic of this paper, namely slow-moving but persistent variations in expected returns. These two phenomenon are likely to co-exist, and so 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 does not exhibit a strong first-order autocorrelation (see Table 2) because the two- and three-month autocorrelation 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 also estimate μ in a GLS regression that uses the covariance matrix Σ implied by H and γ to adjust both the standard error and the point estimate $\hat{\mu}$. 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 somewhat redundant because they 'over-sample' the same epoch of conditional expected returns. As illustrated by the first panel of Appendix Figure 2, GLS therefore overweights observations at the beginning and end of the sample. This effect is larger for higher H, and reverses when γ is negative.

Appendix Table B shows the GLS point estimates and t-statistics given by our model, along with the comparable OLS statistics, for a variety of assumptions about H and γ. We find that the GLS correction to $\hat{\mu}$ discussed above has minimal impact on the point estimates for the profitability portfolio, but somewhat decreases point estimates for value, investment, and size. The reason for this decrease is that the latter three portfolios had unusually low returns at the beginning and/or ends of the sample, which GLS infers as being more independent from the observations in the middle. Overall, the main conclusions from the OLS analysis, that t-statistics can be half as large for reasonable values of H and γ, remain unchanged when using GLS.

Appendix C. Sampling Bayesian Posteriors

We draw samples of $N = 50,000$ observations from the posterior distribution of model parameters Ω^{post} 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 $i = 2 \ldots N$:
   (a) Evaluate the conditional likelihood of the data D given the ith draw from the prior parameters as well as the $i-1$st draw from the posterior parameters:

$$\mathcal{L}^{\text{propose}} = \mathcal{L}(D|\Omega = \Omega_i^{\text{prior}}), \tag{17}$$

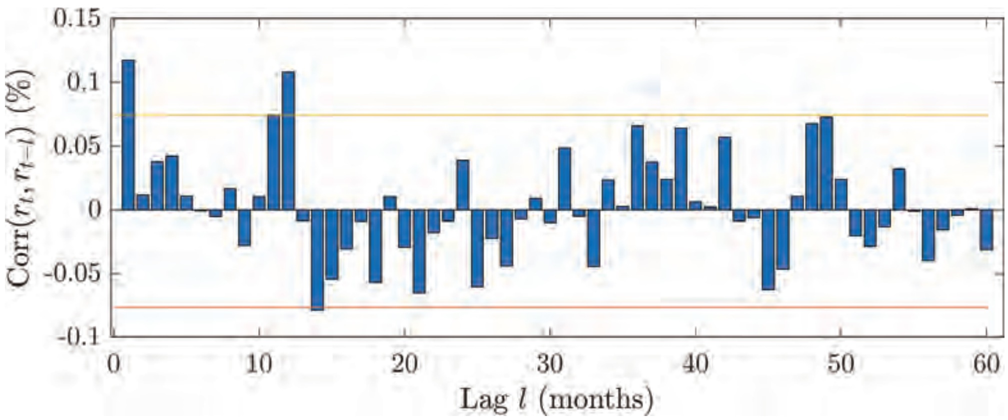
$$\mathcal{L}^{\text{previous}} = \mathcal{L}(D|\Omega = \Omega_{i-1}^{\text{posterior}}). \tag{18}$$

   (b) If $\mathcal{L}^{\text{propose}} \geq \mathcal{L}^{\text{previous}}$, accept $\Omega_i^{\text{posterior}} = \Omega_i^{\text{prior}}$.
   (c) If $\mathcal{L}^{\text{propose}} \leq \mathcal{L}^{\text{previous}}$, accept $\Omega_i^{\text{posterior}} = \Omega_i^{\text{prior}}$ with probability $\frac{\mathcal{L}^{\text{propose}}}{\mathcal{L}^{\text{previous}}}$, and otherwise retain $\Omega_i^{\text{posterior}} = \Omega_{i-1}^{\text{prior}}$.

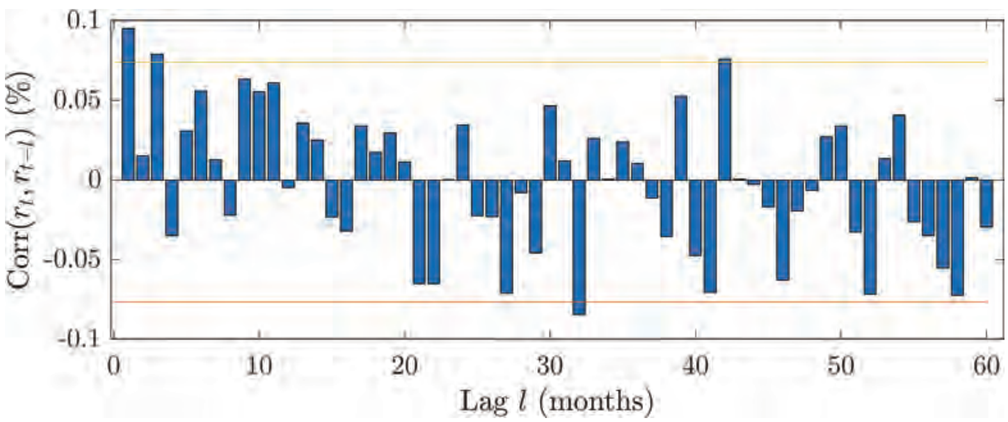
Appendix Figure 1: Monthly Autocorrelograms

This figure presents autocorrelations of monthly returns for value, investment, profitability, and size portfolios, as defined in the header of Table 1, as well 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. Our sample consists of 678 monthly observations from Q3 1963 through Q4 2019.

Panel A: Value

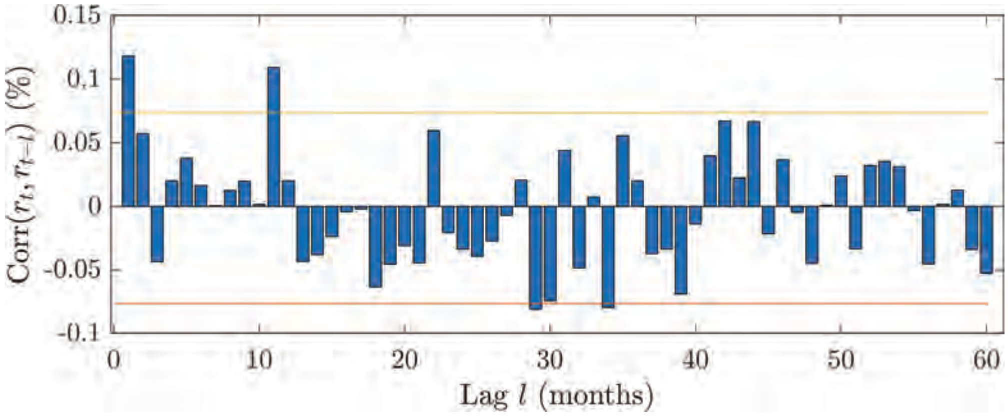


Panel B: Investment

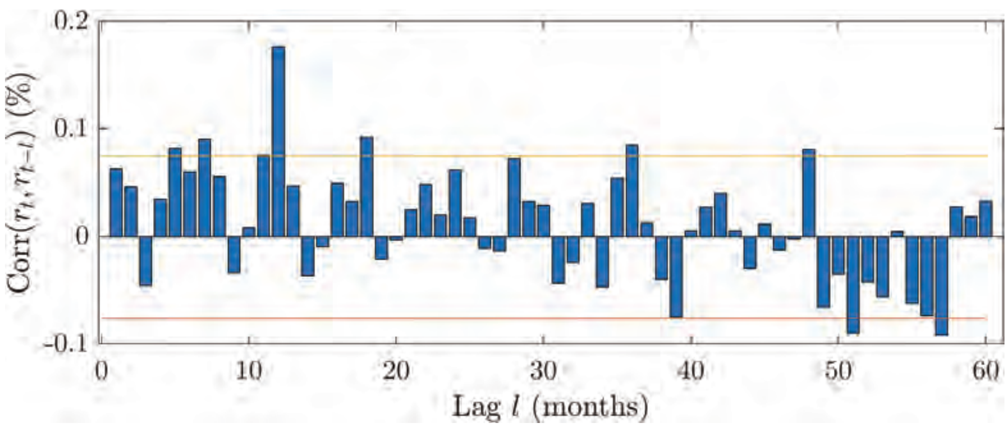


Appendix Figure 1: Monthly Autocorrelograms (continued)

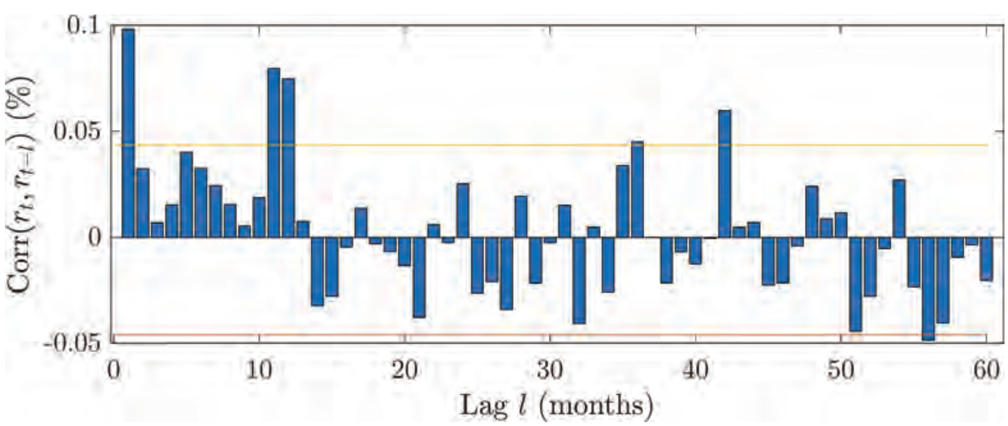
Panel C: Profitability



Panel D: Size

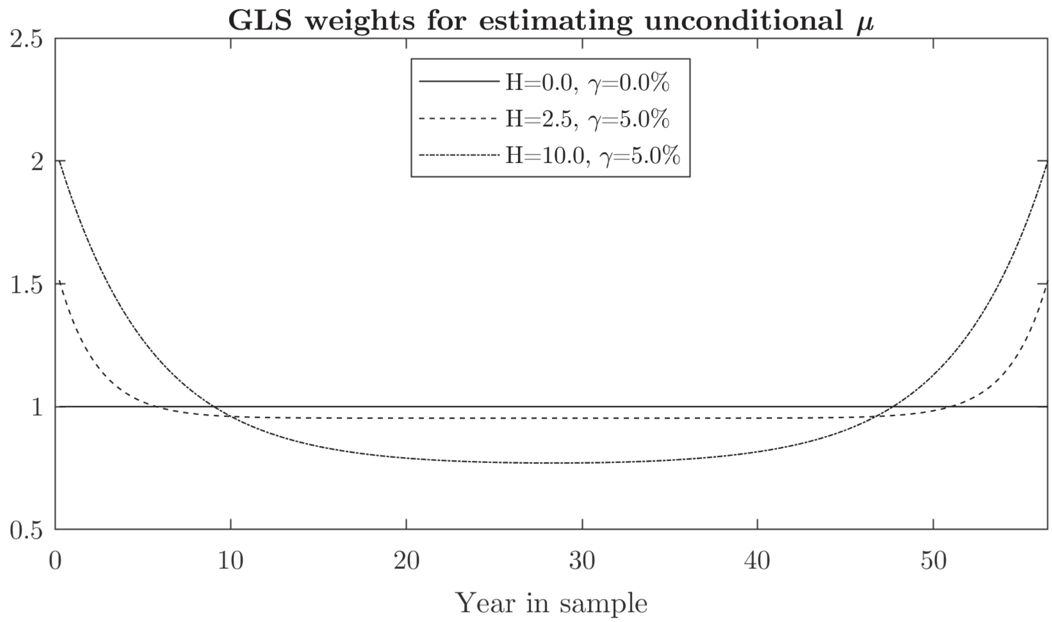


Panel E: Pooled



Appendix Figure 2: GLS Influence Functions and Return Autocorrelations

This figure presents the weights GLS uses when calculating unconditional (Panel A) and conditional end-of-sample (Panel B) average returns under a variety of assumptions about the correlation matrix of returns. H represents the half life of shocks to expected returns, in years. γ is the first-order autocorrelation of returns driven by persistence shocks to expected returns and correlations between realized returns and shocks to expected returns.



Appendix Table 1: Return Autocorrelations for Different Parameterizations

This table shows the first-order autocorrelation in returns, $\gamma = corr(r_t, r_{t-1})$, implied by various parameterizations of the model. We assume $\rho = -1$ throughout to provide an upper bound on negative autocorrelations. Panel A examines various combinations of the half-life of shocks to expected returns, H, and the time-series volatility of conditional Sharpe ratios, σ_{sr}. Panel B examines combinations of H and the ratio of the volatilities of expected and unexpected return shocks, $\sigma_\delta / \sigma_\epsilon$.

Panel A: Autocorrelation as a function of σ_{sr} and H

H	σ_{sr}							
	0	0.001	0.01	0.1	0.25	0.5	0.75	1
0.1	0.00	-0.39	-3.80	-30.22	-49.46	-35.36	14.80	87.39
0.25	0.00	-0.06	-0.55	-3.38	-1.50	13.40	37.08	66.03
0.5	0.00	-0.02	-0.16	-0.26	3.36	15.69	32.36	51.47
1	0.00	-0.01	-0.04	0.53	4.13	14.01	26.29	39.81
2.5	0.00	0.00	0.00	0.73	3.76	10.93	19.22	28.00
5	0.00	0.00	0.01	0.71	3.23	8.75	14.85	21.19
10	0.00	0.00	0.01	0.65	2.68	6.84	11.28	15.83
25	0.00	0.00	0.01	0.54	2.01	4.78	7.66	10.57

Panel B: Autocorrelation as a function of $\frac{\sigma_\delta}{\sigma_\epsilon}$ and H

H	$\frac{\sigma_\delta}{\sigma_\epsilon}$							
	0	0.001	0.01	0.1	0.25	0.5	0.75	1
0.1	0.00	-0.01	-0.12	-1.23	-3.03	-5.94	-8.74	-11.43
0.25	0.00	-0.02	-0.16	-1.45	-2.88	-3.57	-2.50	0.00
0.5	0.00	-0.01	-0.11	-0.46	1.11	7.74	17.53	29.29
1	0.00	-0.01	-0.05	1.05	6.78	21.35	38.69	57.33
2.5	0.00	0.00	0.03	3.56	14.32	35.59	58.00	80.81
5	0.00	0.00	0.10	5.55	18.67	42.13	66.01	90.01
10	0.00	0.00	0.21	7.24	21.51	45.88	70.37	94.91
25	0.00	0.01	0.41	8.71	23.51	48.30	73.12	97.94

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 are summarized by H, the half-life of shocks to expected returns (in years), and γ, the first-order autocorrelation of realized returns. We estimate unconditional expected returns using both OLS and GLS and calculate t-statistics using the model-implied correlation matrix of the regression error terms. Our sample consists of 226 quarterly observations from Q3 1963 through Q4 2019.

Panel A: Value

	H (years)	2.5			5		10	
	γ (%)	2.5	5	10	2.5	5	2.5	5
OLS	$\hat{\mu}$ (%)	4.37	4.37	4.37	4.37	4.37	4.37	4.37
	Model t-stat	(1.83)	(1.54)	(1.23)	(1.58)	(1.26)	(1.34)	(1.03)
GLS	$\hat{\mu}$ (%)	4.19	4.07	3.92	3.99	3.78	3.79	3.45
	Model t-stat	(1.75)	(1.44)	(1.10)	(1.45)	(1.10)	(1.17)	(0.83)

Panel B: Investment

	H (years)	2.5			5		10	
	γ (%)	2.5	5	10	2.5	5	2.5	5
OLS	$\hat{\mu}$ (%)	4.72	4.72	4.72	4.72	4.72	4.72	4.72
	Model t-stat	(2.56)	(2.15)	(1.71)	(2.21)	(1.77)	(1.87)	(1.44)
GLS	$\hat{\mu}$ (%)	4.52	4.41	4.30	4.28	4.05	4.01	3.62
	Model t-stat	(2.45)	(2.02)	(1.57)	(2.01)	(1.53)	(1.60)	(1.13)

Panel C: Profitability

	H (years)	2.5			5		10	
	γ (%)	2.5	5	10	2.5	5	2.5	5
OLS	$\hat{\mu}$ (%)	4.91	4.91	4.91	4.91	4.91	4.91	4.91
	Model t-stat	(2.71)	(2.28)	(1.81)	(2.34)	(1.87)	(1.98)	(1.53)
GLS	$\hat{\mu}$ (%)	4.96	5.02	5.12	4.94	5.02	4.87	4.94
	Model t-stat	(2.75)	(2.34)	(1.91)	(2.37)	(1.94)	(1.99)	(1.56)

Panel D: Size

	H (years)	2.5			5		10	
	γ (%)	2.5	5	10	2.5	5	2.5	5
OLS	$\hat{\mu}$ (%)	0.35	0.35	0.35	0.35	0.35	0.35	0.35
	Model t-stat	(0.13)	(0.11)	(0.09)	(0.11)	(0.09)	(0.10)	(0.07)
GLS	$\hat{\mu}$ (%)	0.35	0.26	0.17	0.00	0.26	0.15	0.33
	Model t-stat	(0.10)	(0.05)	(0.00)	(0.09)	(0.04)	(0.09)	(0.05)

Appendix Table 3: Performance of Bayesian OOS Timing Portfolios

This table presents realized Sharpe ratios and certainty equivalent returns for dynamic trading strategies of Bayesian investors with different priors who calculate posteriors using only past data. For each prior and characteristic-sorted portfolio, we compute the posterior distribution of next-year returns and the corresponding optimal portfolio weights in the risk-free asset and characteristic-sorted portfolio for an investor with CRRA utility and relative risk aversion of 2.

Panel A: Value

Prior		Realized Sharpe Ratio		Certainty Equivalent Return (%)	
H	μ_{sr}	Value	95% CI	Value	95% CI
0	$N(-0.4, 0.4)$	0.15	[-0.11,0.40]	0.31	[-3.00,3.66]
0	$N(0, 0.4)$	0.24	[-0.02,0.49]	1.40	[-2.65,5.57]
0	$N(0.4, 0.4)$	0.29	[0.03,0.54]	1.98	[-3.00,7.07]
0	$U(-2, 2)$	0.25	[-0.01,0.50]	1.31	[-3.59,6.30]
2.5	$N(-0.4, 0.4)$	0.03	[-0.24,0.28]	-0.88	[-3.64,1.97]
2.5	$N(0, 0.4)$	0.18	[-0.09,0.43]	0.59	[-2.91,4.13]
2.5	$N(0.4, 0.4)$	0.25	[-0.01,0.50]	1.44	[-2.96,5.90]
2.5	$U(-2, 2)$	0.20	[-0.06,0.45]	0.64	[-3.72,5.02]
5	$N(-0.4, 0.4)$	0.05	[-0.22,0.31]	-0.67	[-3.55,2.27]
5	$N(0, 0.4)$	0.18	[-0.09,0.43]	0.58	[-2.96,4.20]
5	$N(0.4, 0.4)$	0.25	[-0.01,0.50]	1.35	[-3.09,5.77]
5	$U(-2, 2)$	0.20	[-0.06,0.46]	0.66	[-3.60,5.01]
$U(0, 10)$	$N(-0.4, 0.4)$	0.06	[-0.21,0.31]	-0.58	[-3.44,2.31]
$U(0, 10)$	$N(0, 0.4)$	0.18	[-0.08,0.44]	0.69	[-2.85,4.32]
$U(0, 10)$	$N(0.4, 0.4)$	0.25	[-0.01,0.51]	1.45	[-2.97,5.99]
$U(0, 10)$	$U(-2, 2)$	0.21	[-0.05,0.46]	0.76	[-3.53,5.11]

Panel B: Investment

Prior		Realized Sharpe Ratio		Certainty Equivalent Return (%)	
H	μ_{sr}	mean	95% CI	mean	95% CI
0	$N(-0.4, 0.4)$	0.28	[0.02,0.53]	1.95	[-2.17,6.09]
0	$N(0, 0.4)$	0.35	[0.10,0.60]	3.28	[-1.04,7.72]
0	$N(0.4, 0.4)$	0.39	[0.14,0.64]	4.12	[-0.79,9.09]
0	$U(-2, 2)$	0.35	[0.10,0.60]	3.19	[-1.88,8.34]
2.5	$N(-0.4, 0.4)$	0.17	[-0.09,0.44]	0.22	[-3.96,4.41]
2.5	$N(0, 0.4)$	0.28	[0.02,0.54]	1.99	[-2.19,6.19]
2.5	$N(0.4, 0.4)$	0.34	[0.09,0.60]	3.04	[-1.57,7.73]
2.5	$U(-2, 2)$	0.30	[0.04,0.56]	2.21	[-2.62,7.10]
5	$N(-0.4, 0.4)$	0.20	[-0.06,0.47]	0.71	[-3.47,4.91]
5	$N(0, 0.4)$	0.29	[0.03,0.55]	2.14	[-2.05,6.37]
5	$N(0.4, 0.4)$	0.35	[0.09,0.61]	3.11	[-1.61,7.89]
5	$U(-2, 2)$	0.30	[0.04,0.57]	2.22	[-2.60,7.12]
$U(0, 10)$	$N(-0.4, 0.4)$	0.21	[-0.05,0.47]	0.81	[-3.28,4.92]
$U(0, 10)$	$N(0, 0.4)$	0.30	[0.04,0.56]	2.31	[-1.92,6.53]
$U(0, 10)$	$N(0.4, 0.4)$	0.35	[0.09,0.61]	3.18	[-1.58,7.94]
$U(0, 10)$	$U(-2, 2)$	0.31	[0.05,0.58]	2.43	[-2.43,7.31]

Appendix Table 3: Performance of Bayesian OOS Timing Portfolios (Continued)

Panel C: Profitability

Prior		Realized Sharpe Ratio		Certainty Equivalent Return (%)	
H	μ_{sr}	mean	95% CI	mean	95% CI
0	$N(-0.4, 0.4)$	0.33	[0.09,0.57]	2.61	[-0.20,5.49]
0	$N(0, 0.4)$	0.39	[0.14,0.63]	3.57	[0.24,6.91]
0	$N(0.4, 0.4)$	0.40	[0.15,0.66]	4.01	[-0.24,8.34]
0	$U(-2, 2)$	0.36	[0.11,0.60]	3.35	[-0.64,7.38]
2.5	$N(-0.4, 0.4)$	0.22	[-0.04,0.48]	1.14	[-2.58,4.93]
2.5	$N(0, 0.4)$	0.30	[0.04,0.56]	2.24	[-2.01,6.46]
2.5	$N(0.4, 0.4)$	0.34	[0.08,0.61]	2.89	[-2.11,7.76]
2.5	$U(-2, 2)$	0.31	[0.04,0.58]	2.24	[-2.67,7.11]
5	$N(-0.4, 0.4)$	0.25	[-0.01,0.51]	1.47	[-2.42,5.35]
5	$N(0, 0.4)$	0.31	[0.05,0.58]	2.48	[-1.92,6.82]
5	$N(0.4, 0.4)$	0.35	[0.09,0.61]	2.97	[-2.18,7.98]
5	$U(-2, 2)$	0.31	[0.05,0.58]	2.39	[-2.64,7.38]
$U(0, 10)$	$N(-0.4, 0.4)$	0.25	[-0.02,0.51]	1.47	[-2.43,5.33]
$U(0, 10)$	$N(0, 0.4)$	0.31	[0.05,0.58]	2.47	[-1.78,6.70]
$U(0, 10)$	$N(0.4, 0.4)$	0.35	[0.09,0.62]	3.08	[-1.88,8.07]
$U(0, 10)$	$U(-2, 2)$	0.32	[0.05,0.58]	2.45	[-2.50,7.35]

Panel D: Size

Prior		Realized Sharpe Ratio		Certainty Equivalent Return (%)	
H	μ_{sr}	mean	95% CI	mean	95% CI
0	$N(-0.4, 0.4)$	-0.30	[-0.54,-0.05]	-2.26	[-4.04,-0.62]
0	$N(0, 0.4)$	-0.08	[-0.35,0.18]	-1.37	[-3.55,0.84]
0	$N(0.4, 0.4)$	0.03	[-0.23,0.29]	-1.20	[-4.52,2.18]
0	$U(-2, 2)$	-0.05	[-0.32,0.21]	-2.15	[-5.42,1.06]
2.5	$N(-0.4, 0.4)$	0.24	[-0.02,0.51]	1.29	[-2.55,5.03]
2.5	$N(0, 0.4)$	0.29	[0.03,0.55]	2.11	[-1.25,5.47]
2.5	$N(0.4, 0.4)$	0.31	[0.05,0.56]	2.46	[-1.01,6.02]
2.5	$U(-2, 2)$	0.25	[-0.01,0.51]	1.58	[-2.25,5.45]
5	$N(-0.4, 0.4)$	0.18	[-0.08,0.46]	0.48	[-3.51,4.43]
5	$N(0, 0.4)$	0.23	[-0.03,0.50]	1.32	[-2.15,4.84]
5	$N(0.4, 0.4)$	0.26	[0.00,0.51]	1.71	[-1.91,5.35]
5	$U(-2, 2)$	0.21	[-0.05,0.47]	0.84	[-3.25,4.89]
$U(0, 10)$	$N(-0.4, 0.4)$	0.18	[-0.08,0.46]	0.58	[-3.11,4.25]
$U(0, 10)$	$N(0, 0.4)$	0.24	[-0.02,0.50]	1.45	[-1.82,4.67]
$U(0, 10)$	$N(0.4, 0.4)$	0.26	[0.01,0.52]	1.82	[-1.58,5.32]
$U(0, 10)$	$U(-2, 2)$	0.21	[-0.05,0.47]	0.98	[-2.80,4.79]