Time-Series Efficient Factors

December 2019

Abstract

Factors in prominent asset pricing models are positively serially correlated. Momentum strategies profit by timing these factors; they pick up factor "inefficiencies." We show that, rather than augmenting a model with the momentum factor, each factor can instead be made time-series efficient. Time-series efficient factors earn significantly higher Sharpe ratios than the original factors; they typically contain all the information found in the original factors; and an asset pricing model with time-series efficient factors, such as an efficient Fama-French five-factor model, prices momentum. Time-series efficient factors also explain more of the covariance structure of returns; they therefore appear to align more closely with the true SDF.

JEL classification: G11, G12, G40

Keywords: Factors; asset pricing models; anomalies; momentum

1 Introduction

Most factors are positively autocorrelated: if a factor such as value or profitability has done well, this good performance typically persists.[1] Momentum stems from these autocorrelations: stocks with high past returns load on factors that have done well and those with low past returns load on factors that have done poorly (Ehsani and Linnainmaa, 2019). An investor who trades momentum therefore indirectly times factors based on their past returns; and an asset pricer who wants to describe the cross section of returns needs to add a momentum factor to the model to capture the predictable time variation in the factor premiums.

In this paper we show that it is not necessary to augment an asset pricing model with a momentum factor. The momentum factor serves a purpose only when the original factors are "inefficient," that is, when they are time-series predictable. Instead of adding the momentum factor, we can replace the original factors with their efficient counterparts. A *time-series efficient factor* exploits the autocorrelation in factor returns; its weight on the original factor varies to minimize variance while maintaining expected returns. We show that time-series efficient factors earn higher Sharpe ratios than the original factors and that they typically contain all the information found in the original factors. Moreover, a time-series efficient version of an asset pricing model captures momentum without requiring a distinct momentum factor: Fama and French (2015) five-factor model, for example, does not price momentum, but a model with efficient versions of the same five factors—a time-series efficient five-factor model—does. Time-series efficient factor models push momentum back into the model's individual factors.

We use the Ferson and Siegel (2001) procedure to generate time-series efficient factors. This procedure finds the mean-variance efficient portfolio when an investor has conditioning information. Because factors are autocorrelated, the conditioning information in our context is the factor's past

[1] See, for example, McLean and Pontiff (2016), Arnott et al. (2019), Ehsani and Linnainmaa (2019), and Gupta and Kelly (2019).

return. The optimal weight is a function of three constant parameters and one variable. The constants are the factor's unconditional mean, variance, and autocorrelation. The only variable is the factor's past return, which the investor uses to smooth expected returns; doing so improves the Sharpe ratio. A surprising feature of the optimal weights is that they are nonmonotonic in past returns (Ferson and Siegel, 2001). A factor that has done exceedingly well has a very high expected premium. However, if so, the investor can lower the weight to reduce the strategy's riskiness; the investor earns a higher Sharpe ratio by swapping a high mean for lower volatility.

A useful feature of the Ferson-Siegel procedure is that an investor can estimate, ex-ante, the scope and statistical significance of potential efficiency gains: given beliefs about the factor's unconditional moments, the investor can compute a test statistic for the expected improvement in its Sharpe ratio. The standard value factor, HML, for example, has an annualized mean of 3.9% with a standard deviation of 9.7%. An investor who holds this factor with a constant weight of one therefore earns a Sharpe ratio of 0.40. However, because this factor is significantly autocorrelated—the first-order autocorrelation is 0.17 (t-value = 4.46)—an investor can vary the investment on HML to preserve its premium while taking less risk. We estimate that implementing the Ferson and Siegel program, designed here to exhaust the first-order correlation of HML, increases the Sharpe ratio of HML by 0.32 units. We also predict this improvement to be statistically significant with an expected z-value of 2.54. After implementing the problem, we can compare these predictions to realized improvements: HML's actual Sharpe ratio increases by 0.22 units, and this improvement is significant with a z-value of 2.29. In a five-factor model regression, this efficient factor's alpha is 1.6% (t-value = 3.32). By contrast, in an *efficient* five-factor model regression, the original HML's alpha is 0.7% (t-value = 0.84). That is, the efficient HML contains all the information in the standard HML and more. The time-series efficient HML no longer carries an autocorrelation that can be harvested because our procedure transforms its autocorrelation into a higher Sharpe ratio.

A comparison between the standard and time-series efficient five-factor models illustrates the connection between factor efficiency and momentum. Carhart's (1997) UMD factor earns a monthly CAPM alpha of 73 basis points (t-value = 4.48) from 1963 through 2018, and this factor's five-factor model alpha is 72 basis points (t-value = 4.44). This is also the finding of Fama and French (2016): "No combination of the factors in [the five-factor model] explains average returns on portfolios formed on momentum." UMD's alpha in the *efficient* five-factor model, by contrast, falls by two-thirds and has a t-value of 1.60. The reason for this improvement is that the predictable time variations in the factor premiums in the Fama-French model alone are responsible for a significant share of UMD's profits; therefore, when we push this predictability back into the factors, UMD turns largely redundant.

An investor does not need to know the factors' precise unconditional moments to attain significant time-series improvements. We show that even large estimation errors in means, variances, and autocorrelations reduce efficient factors' Sharpe ratios relatively little. An investor who trades under the weak assumption that factors are at least weakly autocorrelated gains significantly. At the same time, we show that the average factor's moments are quite stable over time; investors can learn enough about autocorrelations from just a small sliver of data. Consistent with these findings, we find that *real-time* implementations of efficient factors perform very similar to the factors that use the original factors' full-sample moments.

Our results are not confined to the Fama-French five-factor model. We examine other factor models—including the four-factor models of Novy-Marx (2013), Hou et al. (2015), and Stambaugh and Yuan (2017)—and find the same pattern: time-series efficient factors dominate standard factors. Time-series efficiency is not the only method for improving factors. Cohen and Polk (1998), Asness et al. (2000), and Novy-Marx (2013) show that industry-neutral factors typically outperform the standard factors, and Daniel et al. (2019) show that loadings-hedged factors have this same property.

We show that time-series efficiency complements these other methods; an investor can improve already improved factors by conditioning out the autocorrelations. An industry-hedged HML, for example, has a Sharpe ratio of 0.79, and a time-series efficient version of this factor has a Sharpe ratio of 0.94.

Time-series efficient factors appear to align more closely with the true stochastic discount factor. We follow Kozak et al. (2018) and extract principal components from the 15 anomalies used in their study. The first principal component explains more of the variation in factor returns than other components. This component can also be viewed as the aggregate risk premium: it increases together with most factors and has a time-series correlation of 0.84 with an equal-weighted portfolio of all 15 anomalies. We find that the efficient five-factor model explains more of the variation in the first PC than the standard five-factor model. Moreover, when the efficient and original factors are included jointly, the first PC loads positively on all efficient factors and negatively on original factors, suggesting that the efficient factors better align with the aggregate risk premium. In fact, the standard factors contain little information about this PC when we already condition on the information found in the efficient factors. These results are striking considering that the size, value, asset growth, and investment anomalies—that closely relate to the original factors SMB, HML, RMW, and CMA—directly feed into the first PC, yet they become insignificant in presence of their efficient versions. Time-series efficient factors therefore seem to be more "systematic" in the cross section of returns than their standard counterparts.

Why might efficient factors have this property? Suppose that we have an unobserved factor F_t for which we attempt to create a mimicking portfolio. If we create a factor $\hat{F}_t$, such as HML, by sorting on characteristics, this factor's loading against F_t may vary over time: $\hat{F}_t = \beta_t F_t + \hat{\varepsilon}_t$, where $\hat{\varepsilon}_t$ is noise unrelated to F_t. That is, there is no reason why, for example, a portfolio sort with fixed 30th and 70th percentile breakpoints would yield a constant correlation against F_t. The correlation between $\hat{F}_t$ and F_t may be low because β_t is volatile. If factor $\hat{F}_t$'s risk premium varies over time, we cannot say why: it could be that the risk premium associated with F_t, λ_t, varies—or it could be that the premium is constant but that β_t varies: $\hat{\lambda}_t = \beta_t \lambda$. The autocorrelations found in almost all factors indicate that some of the variation in $\hat{\lambda}_t$ is predictable. And the fact that the efficient factors appear more "systematic" than the standard factors appears to suggest that they more closely track the true latent factors. That is, when we move from the standard factors to efficient factors, we might be stabilizing the factors' betas against the true factors.

Our results are important from the viewpoint of investing. Because the momentum factor's five-factor model alpha is statistically significant, an investor who trades that model's five factors—the market, size, value, profitability, and investment factors—can earn a higher Sharpe ratio by trading also momentum (Huberman and Kandel, 1987). Our point is that the investor can "trade momentum" in three distinct ways. The first method is for the investor to trade momentum in the cross section of stocks returns (Jegadeesh and Titman, 1993). This strategy, however, is just a noisy version of factor momentum, and so the investor might as well trade "factor momentum" in addition to the five factors: invest in the factors that have done well and shun those that have done poorly (Ehsani and Linnainmaa, 2019). The third method, based on our results, is that investors can distribute factor momentum back into the factors; by trading efficient factors, the investor need not be concerned with a distinct momentum factor.

2 Efficient factors with conditioning information

We use the Ferson and Siegel (2001) framework to construct time-series efficient factors. The problem that Ferson and Siegel examine relates to mean-variance efficient portfolios: what are the weights of the optimal portfolio when an investor is endowed with some conditioning information? An investor might, for example, have a signal that indicates that an asset's return distribution lies to the right relative to its unconditional distribution. The investor's problem is to use the signal to find the mean-variance efficient portfolio, that is, the portfolio with the lowest volatility for any level of expected return.

We first describe the general framework of Ferson and Siegel (2001) under the case of one risky asset and one risk-free asset. We then apply this framework to factors with past returns as the conditioning information. The Ferson-Siegel analysis starts from a single risky asset with a return of

$$\tilde{R} = \mu(\tilde{S}) + \tilde{\epsilon}, \tag{1}$$

in which $\tilde{R}$ is the excess return on the risky asset relative to the risk-free rate, $\tilde{S}$ is the information in the predictor variable (signal), $\mu(\tilde{S})$ is the expected excess return *conditional* on the signal, and $\tilde{\epsilon}$ is the random noise net of the signal with a mean of zero and a variance of $\sigma^2_\epsilon(\tilde{S})$.

The efficient strategy invests $x(\tilde{S})$ in the risky asset and the remaining, $1 - x(\tilde{S})$, in the risk-free asset that earns a zero excess return. This strategy's unconditional expected excess return and variance are

$$\mu_p = \mathbb{E}\big[x(\tilde{S}) \cdot \mu(\tilde{S})\big], \tag{2}$$

$$\sigma^2_p = \mathbb{E}\Big[x^2(\tilde{S}) \cdot \big(\mu^2(\tilde{S}) + \sigma^2_\epsilon(\tilde{S})\big)\Big] - \mu^2_p. \tag{3}$$

Ferson and Siegel (2001) show that, for a given conditional expectation μ_p, the portfolio that minimizes σ^2_p invests $x(\tilde{S})$ in the risky asset:

$$x(\tilde{S}) = \frac{\mu_p}{\zeta} \cdot \frac{\mu(\tilde{S})}{\mu(\tilde{S})^2 + \sigma^2_\epsilon(\tilde{S})}, \tag{4}$$

in which ζ is a constant equal to

$$\zeta = \mathbb{E}\left[\frac{\mu^2(\tilde{S})}{\mu^2(\tilde{S}) + \sigma^2_\epsilon(\tilde{S})}\right]. \tag{5}$$

This weighting program produces a unique mean-variance efficient portfolio (Ferson and Siegel, 2001). That is, no other portfolio obtains the same unconditional return at a lower unconditional variance.

In this paper, we assume that the signal ($\tilde{S}$) is a function of past realized returns. We describe an asset's (or: a factor's) return as following an autoregressive process,

$$\tilde{R}_t = \mu + \rho \tilde{R}_{t-1} + \epsilon_t, \tag{6}$$

with an unconditional mean and variance of $\mu/(1-\rho)$ and $\sigma^2_\epsilon/(1-\rho^2)$, respectively. We assume that the agent knows how to predict the value of the next period's excess return. If the agent has the right model, the conditional expected excess return is $\mu(\tilde{S}) = \mu + \rho\tilde{R}_{t-1}$. Under these assumptions, the R^2 of a regression of expected returns on the signal (past returns) is equal to

$$R^2 = \frac{\sigma^2_{\mu(\tilde{S})}}{\sigma^2_{\mu(\tilde{S})} + \sigma^2_\epsilon} = \rho^2. \tag{7}$$

Using equations (2) and (4), the investor's optimal weight on the risky asset in this setting is

$$x(S_t) = \frac{\mu_p}{\zeta}\left(\frac{\mu(S_t)}{\mu(S_t)^2 + \sigma^2_\epsilon}\right), \tag{8}$$

where the constant ζ and the conditional expected return are equal to

$$\zeta = \frac{SR^2 + \rho^2}{SR^2 + 1}, \tag{9}$$

$$\mu(S_t) = \mu_p\,(1-\rho) + \rho\, r_{t-1}. \tag{10}$$

In this formulation SR is the risky asset's unconditional Sharpe ratio. We define **time-series efficient factor** as the portfolio that places the weight $x(S_t)$ from equation (8) on the original factor. A time-series efficient HML, for example, would be the return on a portfolio that optimally times HML given, in this derivation, its month $t-1$ return. In Appendix A.3 we derive the optimal weight for an efficient factor that conditions on the factor's average return over the prior n months. In our empirical analyses we use both month $t-1$ return and the average return from month $t-12$ to $t-1$ as the conditioning information.

Kozak et al. (2019) suggest that reduced-form factor models do not adequately describe the cross section of stock returns, but that the first few principal components extracted from the universe of factors does. Let us call the first such principal component the dominant factor. We now extend our framework to let a factor's return in month t to depend not only its own past return, but also on the dominant factor's past return. That is, we assume that factors are positively autocorrelated and positively cross-serially correlated with the dominant factor,

$$\tilde{R_{f,t}} = \mu + \rho \tilde{R}_{f,t-1} + \rho' \tilde{R}'_{t-1} + \epsilon_t, \tag{11}$$

where $\tilde{R}'_{t-1}$ is the dominant factor's past return. The signal is now a function of both the factor's own past return and the dominant factor's past return: $\mu(S_t) = \mu + \rho \tilde{R}_{f,t-1} + \rho' \tilde{R}'_{t-1}$. In Appendix A.4 we show that the optimal weight in this extended model is,

$$\begin{aligned} x(S_t) &= \frac{\mu_p}{\zeta} \left(\frac{\mu(S_t)}{\mu(S_t)^2 + \sigma_\epsilon^2} \right), && (12) \\ \zeta &= \frac{SR^2 + \rho^2 + {\rho'}^2}{SR^2 + 1}, && (13) \\ \mu(S_t) &= \mu_p (1-\rho) + \rho \tilde{R}_{f,t-1} + \rho' \tilde{R}'_{t-1}. && (14) \end{aligned}$$

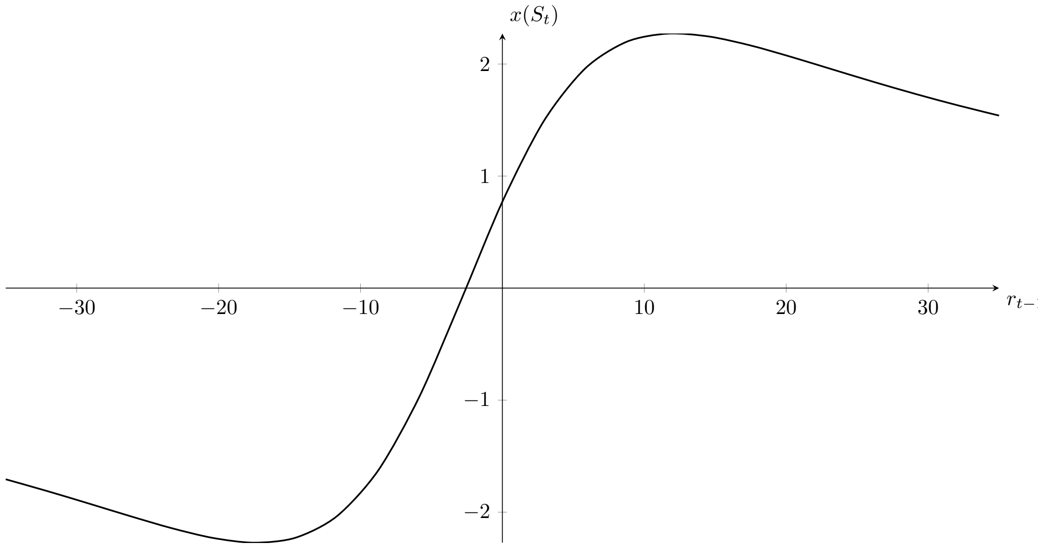

Figure 1: **The optimal weight function.** An investor constructs a time-series efficient factor by predicting month t factor returns with month $t-1$ returns. The optimal weight depends on the factor's mean, standard deviation, and first-order autocorrelation. We compute the average values of these parameters using factor data from popular factor models (see text for details). This figure plots the optimal weight invested in the factor, $x(S_t) = x(r_{t-1})$, as a function of month $t-1$ return. The returns on the x-axis are in percentage points; the weights on the y-axis in decimals, that is, a value of 1.0 indicates a weight of 100%.

2.1 An example of the optimal weight function

What is the optimal investment policy in this setting, that is, how much should an investor allocate to a factor when its return has been close to zero or when it has been very high? To illustrate the optimal policy, we consider the case in which the signal is the factor's return in the prior month. The optimal weight depends on factor means, standard deviations, their ratio and, in this formulation, first-order autocorrelations. For the factors in our sample, the average full-sample monthly estimates of these parameters are approximately:

Mean:	0.35%
Standard deviation:	2.00%
First-order autocorrelation:	0.12

From equation (9), ζ, evaluated at these parameter values, equals $\zeta = 0.044$. The weight function for the "average" factor from equation (8) is then,

$$x(S_t) = x(r_{t-1}) = \frac{0.35}{0.044}\left(\frac{0.35(1-0.12)+0.12\times r_{t-1}}{\left[0.35(1-0.12)+0.12\times r_{t-1}\right]^2+\left[(1-0.12^2)\cdot 2^2\right]}\right). \tag{15}$$

The optimal weight $x(S_t) = x(r_{t-1})$ is now only a function of past returns. We plot this function in Figure 1. Because the factor's unconditional premium is positive, an investor does not switch between positive and negative weights at zero; it is only if the factor has lost more than 2.6% over the prior month that the investor would find it prudent to short the factor.

A counterintuitive result that is apparent in Figure 1 is that the weighting program is not monotone in signal: although the optimal weight initially increases in past return, the optimal strategy begins to scale back as the past return becomes very high. Similarly, on the short side, the investor begins to scale back the negative exposure when the factor's past return is very low. The reason for the nonmonotone behavior lies in the investor's objective of minimizing variance for a given average return. If the value of the signal is very high, the investor *could* make an aggressive bet to earn a high expected return; but because the objective is, in effect, to smooth returns, the investor can afford to invest less in the risky asset to lower the strategy's riskiness. Ferson and Siegel (2001, p. 973) describe this intuition as follows:

> "By using the nonmonotone portfolio weight function, one could either attain a higher unconditional expected return for the same standard deviation, or, alternatively, attain a lower standard deviation for the same expected return. The extra expected return that might be achieved by buying aggressively when the signal is high leads to additional risk. In other words, if the objective is to get the smallest unconditional variance for a given average return, then a signal that the expected return is unusually high presents

an opportunity to reduce risk by purchasing a smaller amount of the risky asset, while maintaining the portfolio average return."

Ferson and Siegel (2001) draw a parallel between the optimal policy function and robust estimation; how an agent should make decisions when he or she does not fully trust the model (Hansen and Sargent, 2008). Bekaert and Liu (2004) confirm this conjecture; they show that the Ferson and Siegel (2001) solution is robust to misspecifications in the conditional moments.

Ferson and Siegel (2001) show that the resulting portfolio earns the target expected return of μ_p with a minimized variance of

$$\sigma_p^2 = \mu_p^2 \left(\frac{1}{\zeta} - 1\right). \tag{16}$$

We can use this formula to compute, in the foregoing example, the improvement in the Sharpe ratio that an investor can *expect* to attain by switching from the "standard factor" (that is, holding the weight constant at $x = 1$) to the time-series efficient factor. Let us first set μ_p equal to the factor's *unconditional* expected return, $\mu_p = 0.35\%$. Because the constant ζ from equation (9) equals $\zeta = 0.044$, the average factor's minimized variance from equation (16) is 2.68. This value corresponds to a standard deviation of 1.64% which, in turn, implies that the time-series efficient factor's monthly Sharpe ratio is $\frac{0.35\%}{1.64\%} = 0.214$. We thus expect the monthly Sharpe ratio for the average factor to increase from 0.175 to 0.214, an improvement of 22%.

If we create a time-series efficient factor that has the same unconditional mean as the original factor, the improvement in the factor's Sharpe ratio results purely from the lower variance. The expected improvement in Sharpe ratio can be expressed as the ratio of the efficient factor's Sharpe ratio (SR^*) to that of the original factor (SR),[2]

$$\frac{SR^*}{SR} = \frac{\zeta}{1-\zeta} = \sqrt{\frac{1 + \left(\frac{\rho}{SR}\right)^2}{1-\rho^2}}. \tag{17}$$

Equation (17) has several implications. First, moving from the standard factor to the time-series efficient factor, the improvement in the factor's Sharpe ratio depends on the ratio of the autocorrelation coefficient to its original Sharpe ratio. When ρ is small, returns become unpredictable, and the ratio in equation (17) converges to one. Because time-series efficiency is about exhausting autocorrelations, the gains evaporate when the time-series predictability vanishes. The potential improvements are the largest for factors that are more autocorrelated, that have lower Sharpe ratios, or that display both of these properties at the same time.

Second, equation (17) suggests that the improvements in Sharpe ratios may depend on the frequency of rebalancing. Suppose, for example, that factor autocorrelations at monthly and daily frequencies are similar; then, because daily Sharpe ratio is $1/\sqrt{21}$ of the monthly Sharpe ratio, the improvements in daily Sharpe ratio would be higher. That is, ignoring transaction costs, high-frequency signals can yield greater improvements in factor efficiency than low-frequency signals; whether this happens in the data depends on the behavior of the autocorrelations when we move across frequencies.

Third, in addition to being able to compute the expected improvement in a factor's Sharpe ratio given the factor's properties—its mean, standard deviation, and autocorrelation—we can also compute the expected z-statistic for the difference in the efficient and original factors' Sharpe ratios. Following Jobson and Korkie (1981), with the correction from Memmel (2003), the test statistic

[2] We derive equation (17) in Appendix A.1 under the assumption that returns follow an AR(1) process, as in equation (6); that the first-order autocorrelation is not too large, and that returns are homoskedastic.

for the expected difference in Sharpe ratios is

$$
\begin{aligned}
z &= \frac{\sigma_o \mu_e - \sigma_e \mu_o}{\sqrt{\theta}}, && (18) \\
\text{where } \theta &= \frac{1}{T}\Big(2\sigma_e^2 \sigma_o^2 - 2\sigma_e \sigma_o \sigma_{e,o} + \frac{1}{2}\mu_e^2 \sigma_o^2 + \frac{1}{2}\mu_o^2 \sigma_e^2 - \frac{\mu_e \mu_o}{\sigma_e \sigma_o}\sigma_{e,o}^2\Big). && (19)
\end{aligned}
$$

Here, σ_e and σ_o are the standard deviations of the efficient and original factors, respectively, μs represent the mean returns, and $\sigma_{e,o}^2$ is the squared covariance between the efficient and original factors. In the foregoing example, we set the efficient factor's mean equal to that of the original factor, $\mu_e = \mu_o$, but the two means can differ. Because the variance of the efficient factor can be computed from equation (16), the only additional that we need to compute the value of the z-statistic from equations (18) and (19) is the covariance between the efficient and original factors. We derive an approximation for this covariance in Appendix A.2.

The predicted z-value—the expected statistical significance of the potential boost in a factor's Sharpe ratio (Jobson and Korkie, 1981)—is a useful tool because Type I and II errors and test power are intrinsic elements of hypothesis testing. Consider, for example, the task of trying to improve a factor that earns a large Sharpe ratio and that is weakly autocorrelated. If we implement the program, we may fail to reject the null hypothesis of no improvement; but only because the large amount of noise in factor returns masks the economically small improvement. The predicted z-statistic takes this lack of power into account; it gives us an idea of what might lie ahead.

3 Standard and efficient five-factor model

3.1 Unconditional returns and time-series predictability

We start with the Fama and French (2015) five-factor model at monthly frequency to illustrate the properties of time-series efficient factors. After presenting the main results, we construct time-series efficient factors using factors drawn from other popular asset pricing models and also construct them for the five factors of the Fama-French model at daily frequency.

The factors in the Fama-French model are the market (MKTRF), size (SMB), value (HML), profitability (RMW), and investment (CMA) factors. The four non-market factors are constructed by sorting stocks at the end of each June into six portfolios by size and (i) book-to-market (HML), (ii) operating profitability (RMW), and (iii) total asset growth (CMA). The breakpoint for size is the NYSE median and the breakpoints for the other predictors are the 30th and 70th percentiles. Fama and French compute value-weighted returns for the resulting portfolios and hold these portfolios for a year until the next rebalancing date.

The value, profitability, and investment factors take long positions in the two high portfolios (small, high profitability, or low investment) with equal weights and short positions in the two low portfolios (big, low profitability, or high investment) also with equal weights. The size factor is an equal-weighted average of three different long-short size factors. Fama and French (2015) first compute the average return for the three small portfolios minus the average return for the three large portfolios using the six portfolios that they create to construct the value factor. This size strategy is the SMB factor from the original Fama and French (1993) three-factor model. They then construct similar size factors using the portfolios underneath the profitability and investment factors; by doing so, they create two slightly different size strategies. The SMB size factor in the five-factor model is the average of these three alternative size strategies.

In Panel A of Table 1 we show the average returns and standard deviations for the factors in the Fama-French five-factor model. The premiums on these factors range from 2.9 percent per year (SMB) to 6.2 percent per year (MKTRF), and the t-values from 2.04 (SMB) to 3.67 (CMA). Much of our analysis is about factor mean-variance efficiency, a concept which relates to factors' Sharpe ratios. The Sharpe ratios, which are proportional to the t-values, range from 0.27 (SMB) to 0.49

Table 1: Fama-French five-factor model: Average returns, standard deviations, and predictability

Panel A reports annualized means, standard deviations, Sharpe ratios, and t-values associated with the average returns for the five factors of the Fama and French (2015) model. Panel B assigns factors into terciles based on month $t-1$ returns and reports average month t returns (and t-values) for these terciles. "Average" at the bottom of the table is computed by assigning all factors first into terciles and then computing the average returns for each tercile. A tercile in month t is empty if no factor gets assigned into it. At least one factor is in the lowest tercile in 586 months, in the middle tercile in 539 months, and in the highest tercile in 588 months. Autocorrelations, reported in the rightmost column, are between month $t-1$ and t returns. Panel C is similar to Panel B except that it assigns factors into portfolios based on prior one-year returns (from month $t-12$ to month $t-1$) and computes the autocorrelations between the prior one-year and month t returns. The data are monthly factor returns from July 1963 through December 2018.

Panel A: Factor means and standard deviations

	Factor				
	MKTRF	SMB	HML	RMW	CMA
Mean	6.17	2.87	3.91	2.98	3.41
Standard deviation	15.22	10.48	9.70	7.47	6.92
Sharpe ratio	0.41	0.27	0.40	0.40	0.49
t-value	3.02	2.04	3.01	2.97	3.67

Panel B: Time-series predictability using month $t-1$ returns

	Month $t-1$ return tercile				Auto-
Factor	Low	Middle	High	H−L	correlation
MKTRF	0.19	0.61	0.74	0.55	0.07
	(0.52)	(2.26)	(3.32)	(1.27)	(1.82)
SMB	−0.07	−0.02	0.81	0.88	0.06
	(−0.35)	(−0.10)	(3.74)	(2.94)	(1.55)
HML	−0.11	0.19	0.90	1.02	0.17
	(−0.61)	(1.21)	(4.33)	(3.64)	(4.46)
RMW	−0.15	0.11	0.80	0.95	0.16
	(−0.92)	(1.14)	(5.21)	(4.18)	(4.13)
CMA	−0.11	0.25	0.72	0.83	0.12
	(−0.89)	(2.08)	(4.87)	(4.26)	(3.16)
Average	−0.07	0.30	0.77	0.83	
	(−0.55)	(3.65)	(7.55)	(5.30)	

(CMA).

An investor who holds one of the factors of the Fama-French five-factor model earns the returns characterized by the estimates in Panel A. Factors' average returns, however, are predictable in the time series. Ehsani and Linnainmaa (2019) show that most factors are positively autocorrelated:

Panel C: Time-series predictability using prior one-year returns

Factor	Prior one-year return tercile: Low	Middle	High	H−L	Auto-correlation
MKTRF	0.24	0.96	0.34	0.09	0.02
	(0.67)	(4.09)	(1.24)	(0.21)	(0.53)
SMB	−0.20	0.29	0.63	0.84	0.09
	(−1.09)	(1.56)	(2.76)	(2.82)	(2.34)
HML	0.12	0.35	0.51	0.39	0.08
	(0.56)	(2.05)	(2.76)	(1.43)	(2.19)
RMW	0.02	0.29	0.44	0.42	0.08
	(0.11)	(2.83)	(3.05)	(1.83)	(2.16)
CMA	0.06	0.28	0.52	0.47	0.08
	(0.44)	(2.25)	(3.48)	(2.37)	(2.08)
Average	−0.06	0.43	0.59	0.65	
	(−0.50)	(5.31)	(5.12)	(3.99)	

the average factor's return is typically significantly higher, both statistically and economically, after a year of gains than losses. For example, if big stocks have outperformed small stocks from month $t-12$ to month $t-1$, which means that the SMB factor has lost money, then SMB, on average, typically continues to lose money in month t.

We illustrate this time-series predictability in Panel B by assigning factors into terciles by their month $t-1$ returns. We compute factors' average returns and t-values associated with those averages. The estimates show, for example, that when the size factor's return in month $t-1$ is in the lowest tercile, then its average return in month t is -7 basis points. If, on the other hand, its return in month $t-1$ is in the highest tercile, this average is 81 basis points. The resulting 88-basis point difference is significant with a t-value of 2.94. Differences for four of the five factors in the Fama-French model are statistically significant at the 5% level; although the difference for the market factor is 55 basis points, this difference has a t-value of just 1.27. We also report monthly first-order autocorrelations, $\rho(r^f_{t-1}, r^f_t)$, for the factors in the rightmost column; these estimates range from 0.06 for the size factor to 0.17 for the value factor.

At the bottom of Panel B we measure the amount of predictability in the average factor of the

Fama-French model. We first sort the five factors independently into terciles using their month $t-1$ returns and then compute the average returns for each tercile in month t. In one month, one of the factors might be in the lowest tercile, three of the factors in the middle tercile, and the last factor in the highest tercile. The average factor's return, when in the lowest tercile, is -7 basis points; but when in the highest tercile, it is 77 basis points. The high-minus-low difference of 83 basis points is statistically highly significant with a t-value of 5.30.

The computations in Panel B show that factors' month t returns are significantly predictable using month $t-1$ returns. This computation does not preclude the possibility that other variables, such as value spread (Cohen et al., 2003), factor volatility (Moreira and Muir, 2017)[3], or characteristics of net stock issuers (Greenwood and Hanson, 2012), also predict factor returns or Sharpe ratios. To illustrate, in Panel C of Table 1 we follow Ehsani and Linnainmaa (2019) and use prior *one-year* returns (from month $t-12$ to $t-1$) to predict month t returns. The predictability estimates are smaller at this horizon but still typically statistically significant; for the five-factor model's average factor, the average return difference between the high and low tercile factors is 65 basis points per month. This difference is significant with a t-value of 3.99.

3.2 Time-series efficient Fama-French five-factor model

We define a time-series efficient factor as a dynamic version of the standard factor; the conditioning information is either the factor's prior performance, or both the factor's and the dominant factor's prior performance. Because factors are positively autocorrelated, the efficient factor places a higher weight on the standard factor when the factor's own prior return (or the dominant factor's prior return) is high; and it shorts the factor when the factor has lost so much money that its conditional risk premium turns negative. Figure 1 illustrates how the efficient factor's weight

[3]See, also, Cederburg et al. (2019) and Liu et al. (2019).

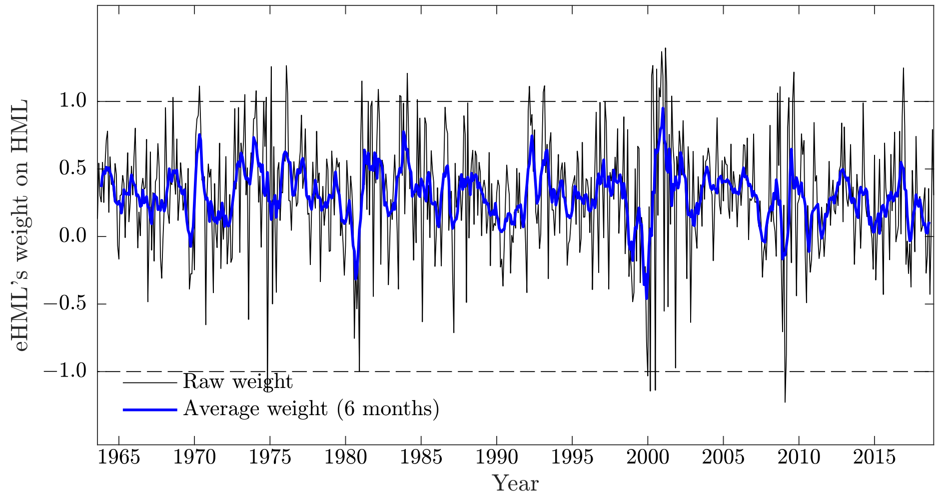

Figure 2: **Efficient HML's weight on HML.** This figure shows time-series efficient HML's weight on the standard HML. A time-series efficient factor uses past-return information as conditioning information to predict the standard factor's return. In this figure, the conditioning information is HML's month $t-1$ return. The investor is assumed to know the unconditional mean, volatility, and first-order autocorrelation of the standard HML. The black line is the actual optimal weight that varies from month to month based on r_{t-1}, $x(\tilde{S})$ from equation (8) with r_{t-1} as the signal; the blue line smooths these weights by computing the average over six-month window around each month. The efficient factor is set to have the same expected premium as the standard factor. The data are monthly factor returns from August 1963 through December 2018.

depends on the factor's prior one-month return.

Figure 2 shows how the efficient value factor changes its weight on the standard HML over the sample period from August 1963 through December 2018. In this figure and in Table 2, we compute the optimal weight from equation (8) using the full-sample estimates for the standard factor's average return, volatility, and first-order serial correlation. We make this assumption, which implies that the trader has perfect information about the full sample properties of the Fama-French factors, only for illustration purposes. In the subsequent tests we let the trader assume that all factors share the same parameters and, therefore, uses the same weight function for all of them. In Section 3.3 we measure the sensitivity of efficient factor's Sharpe ratio to estimation uncertainty in the parameter values.

Because the optimal weight in this analysis derives from HML's month $t-1$ return, the optimal weight often varies considerably from month to month. The efficient factor's average weight is 0.30.

Although the optimal weight could exceed 100% (leverage) or fall below 0% (short), it does so infrequently because of the nonlinear weighting program; the weight falls out of this range in less than 25% of the months. In our implementation of time-series efficient factors, we further bound all weights: we set negative weights to 0% and those above one to 100%. This bounding can be viewed as being about the trade-off between maintaining maximum correlation or producing orthogonal portfolios with respect to the original factor.[4] The $[0, 1]$ bounds ensure that the efficient factor maintains a high correlation with the original factor and has, at most, the same amount of leverage as the original factor. A trader whose objective is to increase alpha with respect to the original factor may let weights to vary without constraints (Ferson, 2019).

Figure 2 shows that the persistence in HML produces persistence in the optimal weight. The blue line is the efficient factor's average weight on HML; in month t, we compute this average over a window from month $t-6$ to month $t+6$. In the late 1990s, for example, when HML underperformed for a relative long period of time, the optimal factor's weight on HML first fell to zero and then became and remained negative.

The purpose of our implementation of the Ferson-Siegel procedure is to improve factors' Sharpe ratios. If factor returns are time-series predictable, an investor can earn the same mean but with lower risk by using past returns to time the factor. In Table 2 we take the five factors of the Fama-French model and compare the standard factors to two versions of time-series efficient factors. The first set of time-series efficient factors use, as in Section 2, month $t - 1$ returns as conditioning information. The second set of time-series efficient factors maintain the assumption that factor returns follow AR(1) processes but takes the factor's average return from month $t - 12$ to $t - 1$ as the signal. In Appendix A.3 we derive the optimal weight of the efficient factor under this assumption. Here, we assume the trader has correct estimates of factor means, variances, and

[4]See Ferson (2019) for a detailed description of optimal orthogonal portfolios, maximum correlation portfolios, and minimum-variance efficient portfolios.

Table 2: Time-series efficient five-factor model: Improvements in Sharpe Ratios

This table compares Sharpe ratios of the original factors of the Fama-French five-factor model to their time-series efficient counterparts. A time-series efficient factor uses either month $t-1$ return or the average return from month $t-12$ to $t-1$ as conditioning information to predict the original factor's month t premium. The efficient factor targets the same expected premium as the original factor. They are constructed using the original factor's unconditional means, volatilities, and first-order autocorrelations. Given beliefs about the original factor's moments, an investor can compute predicted Sharpe ratio for the efficient factor and the predicted z-value for the improvement in the Sharpe ratio from equations (16) and (18). We display these predictions as definition "Predicted efficient factor" when month $t-1$ is used as conditioning information. "Realized efficient factors" are created using either month $t-1$ return or the average month $t-12$ to $t-1$ return as conditioning information; the Sharpe ratios and test statistics are computed from the actual factor returns. The data are monthly factor returns from August 1963 through December 2018.

Factor definition	Statistic	Factor MKTRF	SMB	HML	RMW	CMA
Original factor	Sharpe ratio	0.41	0.27	0.40	0.40	0.49
Predicted efficient factor	Sharpe ratio	0.48	0.34	0.73	0.69	0.65
(Signal: r_{t-1})	Δ Sharpe ratio	0.07	0.07	0.32	0.29	0.16
	z-value	0.99	0.83	2.54	2.34	1.74
	p-value	0.32	0.41	0.01	0.02	0.08
Realized efficient factor	Sharpe ratio	0.52	0.38	0.62	0.64	0.62
(Signal: r_{t-1})	Δ Sharpe ratio	0.12	0.10	0.22	0.24	0.12
	z-value	1.62	1.39	2.29	2.30	1.67
	p-value	0.11	0.16	0.02	0.02	0.09
Realized efficient factor	Sharpe ratio	0.50	0.35	0.48	0.48	0.52
(Signal: $r_{t-1,t-12}$)	Δ Sharpe ratio	0.10	0.07	0.08	0.08	0.02
	z-value	1.51	1.84	1.66	1.35	0.99
	p-value	0.13	0.07	0.10	0.18	0.32

serial correlations at the {1,1} and {12,1} horizons.

We report the Sharpe ratios of the original factors on the first line of Table 2 for reference. As discussed in Section 2, one of the benefits of the efficiency framework is that an investor can compute, ex ante, the expected efficiency gain that can be attained by timing the factor. Using Equation (17) and (18), we compute the expected Sharpe ratio of the time-series efficient factor (using month $t-1$ return) and the expected z-value for the difference between the efficient and original factors' Sharpe ratios. These predicted improvements are based on the original factor's mean, volatility, and first-order autocorrelation. The estimates show that an investor would expect to obtain economically significant gains by timing all five factors of the Fama-French model. The Sharpe ratios of the value (HML) and profitability (RMW) factors are expected to improve the most because they are more autocorrelated than the others (see Panel A of Table 1); the predicted improvements in these factors' Sharpe ratios are 0.32 and 0.29.

Table 2 shows that all efficient factors indeed outperform the original factors when we use month $t-1$ returns as conditioning information. The realized improvements in Sharpe ratios range from 0.10 (for SMB) to 0.24 (to RMW). The z-values for improvements completely align with predictions: ex-ante, we expect efficient versions of HML and RMW to deliver Sharpe ratio improvements that are significant at the 5% level; we expect CMA to realize an improvement that is significant at the 10% level; and we expect MKTRF and SMB to earn improvements that "fail to reject the null of any improvement" at the 10% level. All five predictions happen in data.

That we fail to statistically detect an effect for MKTRF and SMB should not be interpreted as the program not producing the intended effect; it can mean that the tests on improvements for MKTRF and SMB should be assessed at a different power. We did not expect the latter two to produce, at the 10% level, statistically significant improvements since the start. Indeed, they both have met expectations.

At the bottom of the table we examine the performance of efficient factors that condition on the factor's average return from month $t-12$ to $t-1$. Although the improvements are not as large as those at the monthly frequency—consistent with the results on time-series predictability in Panels B and C of Table 1—all five factors still see improvements in their Sharpe ratios. The improvements in the Sharpe ratios of the size (SMB) and value (HML) factors are statistically significant at the 10% level.

Table 2 shows that the efficient factors, constructed using either month $t-1$ returns or the

Table 3: Time-series efficient five-factor model: Spanning tests

This table reports regressions in which the dependent variable is one of the factors of the efficient (Panels A, C, E, and G) or standard (Panels B, D, F, and H) five-factor model and the explanatory variables are all factors of the other model. We do not report the slope estimates on the five right-hand-side factors. The efficient factors are constructed as in Table 2; each time-series efficient factor invests in the standard factor with a weight that is a function of either month $t-1$ returns (Panels A and B) or the average returns from month $t-12$ to $t-1$ (Panels C and D). The Gibbons et al. (1989) tests at the bottom of each panel test the null hypothesis that the alphas of the five factors are jointly zero. The data are monthly factor returns from August 1963 through December 2018.

Panel A: Efficient factors regressed on standard factors, signal: r_{t-1}

	Factor				
	MKTRF	SMB	HML	RMW	CMA
Alpha	2.20	1.58	2.09	1.59	0.90
	(2.40)	(2.67)	(3.90)	(3.76)	(2.64)
	F-value	p-value			
GRS test: αs jointly zero	5.38	0.01%			

Panel B: Standard factors regressed on efficient factors, signal: r_{t-1}

Explanatory	Factor				
variable	MKTRF	SMB	HML	RMW	CMA
Alpha	4.07	0.37	0.10	0.70	0.67
	(2.91)	(0.45)	(0.13)	(1.17)	(1.54)
	F-value	p-value			
GRS test: αs jointly zero	2.74	1.86%			

Panel C: Efficient factors regressed on standard factors, signal: r_{t-1} and r'_{t-1}

	Factor				
	MKTRF	SMB	HML	RMW	CMA
Alpha	2.28	1.81	2.08	1.51	0.80
	(2.57)	(3.11)	(3.98)	(3.67)	(2.36)
	F-value	p-value			
GRS test: αs jointly zero	6.12	0.00%			

Panel D: Standard factors regressed on efficient factors, signal: r_{t-1} and r'_{t-1}

Explanatory	Factor				
variable	MKTRF	SMB	HML	RMW	CMA
Alpha	3.63	0.00	0.01	0.75	0.64
	(2.52)	(0.00)	(0.01)	(1.23)	(1.39)
	F-value	p-value			
GRS test: αs jointly zero	2.50	2.96%			

Panel E: Efficient factors regressed on standard factors, signal: $r_{t-1,t-12}$

Explanatory variable	Factor				
	MKTRF	SMB	HML	RMW	CMA
Alpha	1.65	1.19	1.36	1.06	0.13
	(1.86)	(2.61)	(2.85)	(2.68)	(0.74)
	F-value	p-value			
GRS test: αs jointly zero	3.25	0.65%			

Panel F: Standard factors regressed on efficient factors, signal: $r_{t-1,t-12}$

Explanatory variable	Factor				
	MKTRF	SMB	HML	RMW	CMA
Alpha	2.02	−0.46	0.32	0.62	0.27
	(1.89)	(−0.85)	(0.57)	(1.19)	(1.36)
	F-value	p-value			
GRS test: αs jointly zero	1.67	14.00%			

Panel G: Efficient factors regressed on standard factors, signal: $r_{t-1,t-12}$ and $r'_{t-1,t-12}$

Explanatory variable	Factor				
	MKTRF	SMB	HML	RMW	CMA
Alpha	1.58	1.12	1.48	1.15	0.17
	(1.86)	(2.53)	(3.10)	(2.94)	(0.92)
	F-value	p-value			
GRS test: αs jointly zero	3.61	0.31%			

Panel H: Standard factors regressed on efficient factors, signal: $r_{t-1,t-12}$ and $r'_{t-1,t-12}$

Explanatory variable	Factor				
	MKTRF	SMB	HML	RMW	CMA
Alpha	1.69	−0.42	0.25	0.43	0.25
	(1.63)	(−0.79)	(0.45)	(0.85)	(1.20)
	F-value	p-value			
GRS test: αs jointly zero	1.31	25.67%			

average return from month $t-12$ to $t-1$, earn higher Sharpe ratios than the original factors when the trader has correct information about each factor's full-sample properties. In all further tests, we assume that the trader does not have perfect information; instead, the trader approximates and uses the same weighting function for all factors,

$$x(S_t) = x(r_{t-1}) = \frac{0.35}{0.044}\left(\frac{0.35(1-0.12) + 0.12 \times r_{t-1}}{\left[0.35(1-0.12) + 0.12 \times r_{t-1}\right]^2 + \left[(1-0.12^2) \cdot 2^2\right]}\right). \tag{20}$$

This weighting function assumes that all means, standard deviations, and autocorrelations are 0.35%, 2.00%, and 0.12. In addition to this specification, we also consider the three alternative signals described above: (1) the signal is the average return over the prior year (see Appendix A.3); (2) the signal combines the factor's own prior month with the dominant factor's prior month return; and (4) the signal combines the factor's own prior one-year return with the dominant factor's prior one-year return (see Appendix A.4)

In Table 3 we measure the incremental information contents of the two sets of factors relative to each other. In Panel A, C, E, and G we regress efficient factors against the five standard factors; in Panels B, D, F, and H we regress the standard factors against the efficient factors. The four panels correspond to our four signal schemes.

The alphas from these regressions have two interpretations. First, following Barillas and Shanken (2017), an insignificant alpha would indicate that the left-hand side factor would not improve the asset pricing model represented by the right-hand side factors. For example, when the standard SMB's alpha is insignificant in Panel B's regression against the efficient five-factor model, this finding suggests that the efficient five-factor model could not be improved by augmenting it with the standard SMB factor. Second, following Huberman and Kandel (1987), the alpha also measures a factor's worth from an investment viewpoint. An insignificant alpha indicates that an investor who already optimally trades the right-hand side factors to maximize the portfolio's Sharpe ratio would not benefit by adding the left-hand side factor to the investment opportunity set. Again, in Panel B, the SMB factor's insignificant alpha indicates that an investor who trades the efficient five-factor model would not have earned a statistically significantly higher Sharpe ratio over the sample period by trading the standard SMB factor as well.

The estimates in Table 3 show that the time-series efficient factors typically contain all the information in the standard factors—and more. Panel A shows, consistent with Table 2, that all time-series efficient factors are incrementally informative about future returns even when controlling for the five-factor model. The Gibbons et al. (1989) test for the joint significance of the alphas returns a p-value of 0.0001. By contrast, Panel B shows that, except for the market factor (MKTRF), the standard factors are not incrementally informative about future returns when controlling for the efficient five-factor model. Nevertheless, because of the significance of the market factor (t-value of 2.91), the GRS test rejects the null that all alphas are jointly zero with a p-value of 1.86%.

Panels C and D of Table 3 show the results for time-series efficient factors constructed using both the factor's own and the dominant factor's past return. In regressions of the time-series efficient factors against the standard factors, alphas are on average larger than those of Panel A, as reflected in a higher GRS F-value of 6.12. This indicates that the investor can expand the mean-variance boundary by using both individual factor returns and aggregate factor returns jointly for portfolio construction. Panel D, by contrast, shows that only MKTRF is informative when the regressions are reversed. The GRS test for the joint significance of standard factor gives a p-value of 2.96%.

Panels E and F of Table 3, which construct the time-series efficient factors using last year's returns, more starkly delineate between the standard and time-series efficient factors. In regressions of the time-series efficient factors against the standard factors, all alphas are positive, those of the size (SMB), value (HML), and profitability (RMW) factors are statistically significant at the 1% level, and the GRS test shows that these efficient factors contain information not found in the standard factors. This test returns a p-value of 0.65%. Panel F, by contrast, shows that *all* standard factors are uninformative when the regressions are reversed; by defining the time-series efficient factors using prior one-year returns as the conditioning information, none of the standard factors have any information incremental to that found in the time-series efficient five-factor model. The GRS test for the joint significance of the alphas returns a p-value of 0.14. When we form efficient factors using both the factor's own and the dominant factor's past returns (Panels G and H), the alphas on the original factors tend even closer to zero. None of the alphas in Panel H are significant at conventional levels and the p-value from the GRS test is 26%. Put differently, an investor trading the time-series efficient factors would not have gained anything by trading the standard factors.

In short, our results suggest the Fama-French factors are, at times, inefficient: when their expected returns are low, they carry too much volatility. We can detect this unpriced volatility ex-ante; using equation (17), we estimate that more than 20% of the time-series volatility of these portfolios is not priced. The weight function dynamically adjusts exposure to minimize the unpriced volatility which results in factors with lower risk but the same expected return, that is, "time-series efficient factors." The results in Tables 2 and 3 confirm that all factors benefit from hedging out the unpriced risk in factor returns. Every individual factor earns a higher Sharpe ratio than its standard version, and the resulting efficient five-factor model spans the standard five-factor model. This improvement is also reflected in ex-post tangency portfolios of the original and efficient factors. The squared Sharpe ratio of the five-factor model is 1.16; the squared Sharpe ratios of the efficient five-factor models that use either the factor's prior one-month or prior one-year returns are 1.41 and 1.31; and the squared Sharpe ratios of the models that use both the factor's own and the dominant factor's prior one-month prior one-year returns are 1.51 and 1.38.

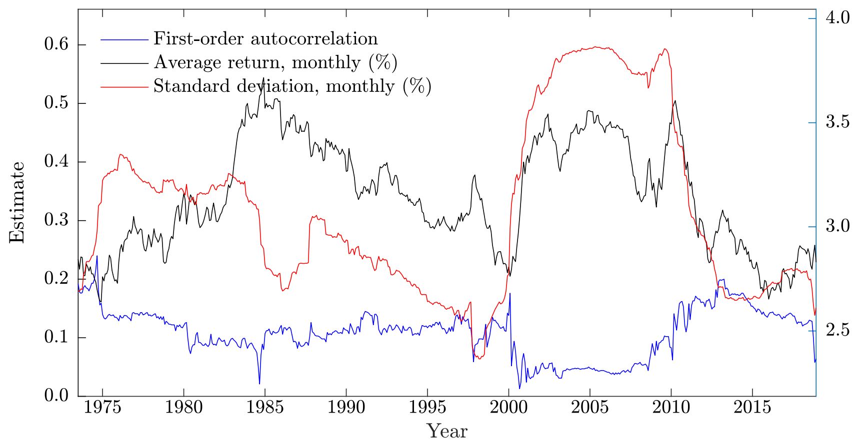

Figure 3: **Average returns, first-order autocorrelations, and standard deviations of the Fama and French (2015) factors.** This figure shows 10-year rolling average estimates of average returns, first-order autocorrelation, and standard deviations for the five factors of the Fama and French (2015) model. We estimate these parameters separately for each factor and then average over the five factors. Average returns and standard deviations are reported in percentages per month. The data are monthly factor returns from July 1963 through December 2018.

3.3 Efficient factors and the sensitivity to estimation errors

An investor or an econometrician constructing time-series efficient factors needs estimates of standard factors' means, standard deviations, and autocorrelations. In Tables 2 and 3 we use the full-sample estimates for these moments to construct the efficient factors. Important questions about efficient factors relate to implementability and lookahead bias: can investors obtain meaningful efficiency gains from the Ferson-Siegel procedure without the benefit of hindsight? In Figures 3 and 4 and Table 4 we show that any estimation errors are largely inconsequential: the parameters are relatively stable in the data; even if an investor makes a large error in the parameter estimates, he or she still improves the Sharpe ratio substantially; and even if an investor naïvely estimates the parameters from historical data, the improvements in Sharpe ratios are similar to those in Tables 2.

In Figure 3 we show the average monthly returns, average standard deviations of monthly returns, and average first-order autocorrelations for the five factors of the Fama-French model. At the end of every month starting in June 1973, we compute these moments for all five factors using ten years of historical data. We then take the averages of these moments over the five factors. Figure 3 plots the resulting rolling averages.

Although factors' autocorrelations, for example, vary over time, they remain positive over any ten-year segment in the data. Up to 1997, the average factor's first-order autocorrelation was typically above 0.1; and after a period of lower (but positive) autocorrelations, they again rose above 0.1 in 2010. The estimates in Figure 3 suggest that investor, in real time, could have made reasonable inferences about the required moments using historical data alone.

Figure 3 does not address the question of how large estimation errors would be acceptable. That is, if an investor estimates an autocorrelation of 0.2 when the true autocorrelation is just 0.10, would the investor, in fact, be worse off by holding the efficient factor? First, any estimation errors in the original factor's mean are largely inconsequential. The problem that the efficient factors solve is isomorphic to the mean-variance analysis; if an investor assumes that the standard factor's mean is lower than what it truly is, then the investor merely searches for a factor with a lower mean—but one that still lies on the efficient frontier. That is, this "wrong" factor will still have the lowest variance for its mean; although an investor could increase the Sharpe ratio by moving to the right on the efficient frontier, those gains would be relatively small.

Estimation errors in the other two parameters, first-order autocorrelation and standard deviation, might, nevertheless, be more important. Figure 4 shows that they are not. In this figure we assume that the investor uses the same parameters for all five factors of the five-factor model to generate their efficient counterparts. We compute these estimates using the full sample of returns for all factors. Using these same estimates for all factors, for simplicity, the average efficient factor earns a Sharpe ratio of 0.64; this is an increase of 0.11 over the Sharpe ratio of the average standard factor. With these full-sample estimates on hand, we then vary the estimates of the first-order autocorrelations (Panel A) and standard deviations (Panel B), construct new time-series efficient

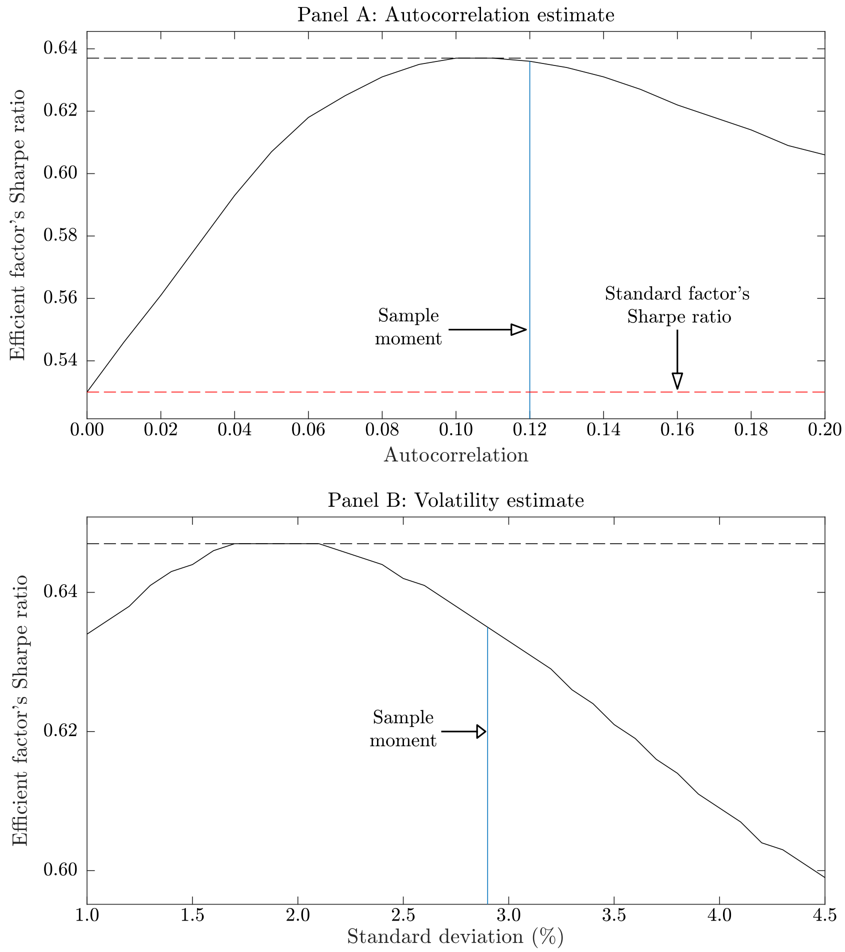

Figure 4: **The sensitivity of the efficient factor's Sharpe ratio to estimation errors in the inputs.** An investor constructs a time-series efficient factor by predicting month t factor returns with month $t-1$ returns. The optimal weight depends on the factor's mean, standard deviation, and first-order autocorrelation. We compute these moments using the full sample from July 1963 through December 2018 for the five-factors of the Fama and French (2015) model. The average original factor's Sharpe ratio is 0.53 and, using these same sample moments to generate efficient versions of all five factors, the average efficient factor's Sharpe ratio is 0.64. Panel A shows the efficient factor's Sharpe ratio as a function of the first-order autocorrelation estimate: what if, instead of setting $\hat{\rho} = 0.12$, which is the average across the five factors, we construct the efficient factor using some other value? Panel B is similar to Panel A except that it varies the estimate of the factor's standard deviation; the average volatility across the five factors is 2.9% per month.

factors, and compute the Sharpe ratios of the resulting not-quite-as-efficient factors.

Figure 4 shows that even large estimation errors in the inputs do not have a meaningful effect on the average efficient factor's Sharpe ratios. In terms of autocorrelations, Panel A shows that the efficiency gains would evaporate only if the investor would believe that the factors are *not* positively

Table 4: Real-time time-series efficient factors

This table compares Sharpe ratios of standard factors in the Fama-French five-factor model to time-series efficient versions of these factors. A time-series efficient factor uses month $t-1$ return as conditioning information to predict the original factor's premium. The efficient factor is set to have the same expected premium as the standard factor. This table is similar to Table 2 except that in this table the factors' means, volatilities, and first-order autocorrelations are estimated using a backward-looking ten-year rolling window. Because of the use of the ten-year window, the sample begins in July 1973 and ends in December 2018.

Factor definition	Statistic	Factor: MKTRF	SMB	HML	RMW	CMA
Original factor	Average return	6.17	2.87	3.91	2.98	3.41
	Standard deviation	15.22	10.48	9.70	7.47	6.92
	Sharpe ratio	0.41	0.27	0.40	0.40	0.49
Efficient factor	Average return	4.19	2.06	4.34	3.22	2.80
(Signal: r_{t-1})	Standard deviation	9.73	7.24	6.83	5.14	5.15
	Sharpe ratio	0.43	0.29	0.64	0.63	0.54
	Δ Sharpe ratio	0.02	0.01	0.23	0.23	0.05
	z-value	0.26	0.14	3.03	2.74	0.77
	p-value	0.80	0.89	0.00	0.01	0.44

autocorrelated. If the investor were to believe that the autocorrelation is far higher than what it is in the data, that would not matter; it would just mean that the investor would hold a bit more of the factor after it has earned positive returns and be more quick to short the factor when its past return is negative. But, by behaving so, the investor does not lose much relative to the optimal solution computed using the average in-sample correlation of approximately 0.1.

Panel B shows that the same is true for volatility. In fact, in terms of realized Sharpe ratios, an investor would have been better off by operating under the assumption that the average factor is less volatile than what it actually was; and even if the investor had overestimated the riskiness of the factors by more than 50%, the investor would have still created a set of efficient factors with an average Sharpe ratio of 0.6

Figures 3 and 4 suggest, first, that investor could have estimated the necessary parameters from the data and, second, even if the investor had faced significant estimation errors, he or she would still have observed economically significant improvements over the original factors. In Table 4 we consider a naïve investor who, at the end of each month, estimates the required parameters—unconditional means, standard deviations, and first-order autocorrelations—using past ten years of monthly data. This analysis plausibly sets a lower bound for the efficiency gains that an investor might have attained in real-time by using past return information to time the factors.

Similar to Table 2, we show the Sharpe ratios of the original five factors of the Fama-French model and, in this case, the efficient *real-time* counterparts of these factors. As before, and consistent with Figures 3 and 4, the Sharpe ratios of all efficient factors exceed those of the original factors. The gains are small for the market and size factors but economically and statistically highly significant for the value and profitability factors. The Sharpe ratios of the latter two factors both increase by 0.23. These increases are almost the same as those reported in Table 2, 0.22 and 0.24, even though in that case the investor was endowed with the full-sample estimates of required parameters.

4 Momentum and the standard and efficient Fama and French (2015) five-factor model

Jegadeesh and Titman (1993) show that a strategy that buys stocks that have earned higher returns relative to other stocks over the past year continue to outperform stocks that have earned relatively low returns. Asness et al. (2013) show that this momentum effect is not specific to just the cross section of equities; the same effect exists also in, for example, fixed income, currency, and commodity markets. Ehsani and Linnainmaa (2019) suggest that momentum is inextricably linked to factors; they find that a factor momentum strategy—a strategy that is long factors that have performed well and short those that have done poorly—explains all forms of individual stock momentum. Their explanation is that cross-sectional momentum strategies implicitly time factors using their past returns: a group of stocks with high past returns, for example, must have factor loadings such that the factors' high and low returns transmit into stock returns. A strategy that buys winning stocks and shorts losing stocks, in effect, makes the bet that factors are autocorrelated.

Time-series efficient factors redefine standard factors to push factor momentum back into the individual factors. For example, using the average month $t-12$ to $t-1$ return as the conditioning information, the average Fama-French factor's one-year autocorrelation, $\rho(r_t, r_{t-1,t-12})$, decreases from 0.063 to -0.005. If cross-sectional momentum strategies indeed profit from the time-series predictability found in the original factors, then the time-series efficient factors should explain at least a part of momentum profits; the point of the efficient factors is that they already try to exhaust such momentum profits.

In Table 5 we examine the connection between momentum and the standard and time-series efficient factors. In the first three regressions the dependent variable is the return on Carhart's (1997) momentum factor, UMD. Similar to, e.g., HML, this factor sorts stocks into six portfolios by the market value of equity (using the NYSE median as the breakpoint) and stocks' prior one-year returns skipping a month (using the 30th and 70th NYSE percentiles as the breakpoints). The first column shows that this factor earns a CAPM alpha of 73 basis points (t-value of 4.48) over the 1964 through 2018 sample period. Fama and French (2016) find, similar to their three-factor factor, that no combination of the factors of the five-factor model can explain the momentum profits. Indeed, our second regression shows that UMD's five-factor model alpha is 72 basis points (t-value of 4.44).

In the third regression we explain momentum profits using the time-series efficient five-factor model. The alpha decreases by two-thirds to 25 basis points per month, and this estimate associates with a t-value of just 1.60. The efficient five-factor model's ability to explain momentum is perhaps surprising; according to the Ehsani and Linnainmaa (2019) mechanism, momentum is the aggregation of the autocorrelations found in *all* factors. Therefore, even if that mechanism

Table 5: Momentum and the standard versus efficient Fama and French (2015) five-factor model

This table reports estimates from time-series regressions that measure the association between momentum and the standard and efficient versions of the Fama and French (2015) five-factor model. In the first three models the dependent variable is Carhart's (1997) momentum factor, which is constructed by sorting stocks into six portfolios by firm size and their prior one-year return, skipping a month. In the other three models, the dependent variable is a time-series momentum strategy that trades the five factors of the Fama-French model. This strategy is long those factors in month t that earned a positive average return from month $t-12$ to $t-1$, and short the factors that lost money. The asset pricing model is the Sharpe (1964)-Lintner (1965) CAPM, the standard Fama-French model, and an efficient version of this model. A time-series efficient factor uses month $t-1$ return as conditioning information to predict the standard factor's return. The efficient factor targets the same expected premium as the standard factor; the weight is based on the original factors' unconditional means, volatilities, and first-order autocorrelations. The sample period begins in August 1964 and ends in December 2018.

	Dependent variable					
	UMD			FF5 Factor Momentum		
Independent		Fama-French FF5			Fama-French FF5	
variable	CAPM	Standard	Efficient	CAPM	Standard	Efficient
Alpha	0.73	0.72	0.25	0.30	0.24	−0.01
	(4.48)	(4.44)	(1.60)	(4.52)	(3.63)	(−0.25)
MKTRF	−0.12	−0.13	0.19	0.02	0.02	0.20
	(−3.33)	(−3.24)	(3.96)	(1.04)	(1.01)	(12.07)
SMB		0.07	0.22		0.10	0.19
		(1.16)	(3.52)		(4.45)	(8.56)
HML		−0.51	−0.14		−0.04	0.18
		(−6.60)	(−1.53)		(−1.33)	(5.57)
RMW		0.24	0.92		−0.02	0.29
		(3.09)	(9.55)		(−0.50)	(8.53)
CMA		0.35	0.35		0.17	0.18
		(3.05)	(2.90)		(3.56)	(4.27)
N	665	665	665	665	665	665
Adjusted R^2	1.5%	8.3%	13.3%	0.0%	5.1%	35.4%

holds in the data, there is no a priori reason for why the momentums present in the five factors of the Fama-French model should exhaust the majority of the cross-sectional momentum profits. Nevertheless, in the data, they do.

The other three columns in Table 5 have the return on the five-factor model *factor momentum* strategy as the dependent variable. This strategy takes long and short positions in the five factors based on their average returns from month $t-12$ to $t-1$. Similar to the UMD factor, this strategy rebalances monthly. If, for example, market has lost money but the other four factors are up, then this time-series factor momentum strategy is short the market and long the other four factors. Because all factors are zero-investment strategies, also the time-series factor momentum strategy is self-financing. A strategy that rotates the five factors of the Fama-French model is profitable: it earns a CAPM alpha 30 basis points per month (t-value = 4.52). Moreover, because this strategy times factors based on their past returns, five-factor model itself cannot explain these profits; its five-factor model alpha is 24 basis points (t-value = 3.63). The part of the alpha that the five-factor model explains relates to the time-series momentum strategy's net long investment in factors (Goyal and Jegadeesh, 2017). The factor momentum strategy is more exposed to factors with higher means, and the five-factor model regression removes this "static" component of the trading profits.

Because the time-series factor momentum strategy relies on the same feature as time-series efficient factors, this strategy's alpha in the *time-series efficient* five-factor model is close to zero: -1 basis point per month (t-value = -0.25). Taking the investment perspective, this result implies that an investor who trades the time-series efficient factors would find the factor momentum strategy redundant; the time-series efficient factors already reap the same predictable time variation in factor premiums that the factor momentum strategy targets.

Five-factor model's ability to price momentum is important from the model building perspective. Starting with Fama and French (1996), the dual objective of new asset pricing models has been, first, to capture cross-sectional variation in asset returns and, second, to do so by presenting the most parsimonious model. The ideal is to have a model that does not include a separate factor for each new anomaly discovered in the data. The spanning results in Table 5 represent a step towards parsimony: instead of augmenting the five-factor model with a distinct momentum factor—either with UMD or, as in Ehsani and Linnainmaa (2019), with a time-series factor momentum strategy—the original factors can be replaced with their time-series efficient counterparts to capture *most* of the momentum profits. The results in Table 5 do not suggest that the time-series efficient five-factor model captures *all* forms of momentum. It almost certainly does not: there are additional factors unrelated to the five-factor model, those factors autocorrelate as well, and the efficient five-factor model, by definition, cannot price strategies that trade momentum in those omitted factors. At the same time, the same argument about time-series efficiency applies: instead of creating a separate momentum factor, we can instead put momentum back into the original factors.

5 Are the results specific to the Fama and French (2015) model and monthly returns?

5.1 Other factor models

In Table 6 we show that the results on time-series efficient factors are not specific to the five-factor model. We consider the following models:

1. **The Daniel et al. (2019)** hedged-factor model. The factors in this model are the same as those in Fama and French (2015). Daniel et al. (2019) extend the characteristics-versus-covariances result of Daniel and Titman (1997): holding characteristics fixed, factor loadings do not meaningfully predict the cross section of stock returns. They construct a "hedge portfolio" for each factor by using stocks' predicted factor loadings. Their method seeks portfolios that have net-zero characteristics but that maximally correlate with the original characteristics-based factors. The five factors in their model are equal to the returns on the original Fama-French factors minus the returns on these hedge portfolios.

2. **The Stambaugh and Yuan (2017) four-factor mispricing model.** The factors in this

Table 6: Other factor models: Standard versus time-series efficient factors

This table compares Sharpe ratios of standard factors to their time-series efficient counterparts. A time-series efficient factor uses past returns as signals to remove unpriced volatility from the factor's return. The signal is, alternatively, (1) the factor's own month $t-1$ return, (2) the factor's own month $t-1$ return and the average month $t-1$ return of all other factors, (3) the factor's own prior one-year return, or (4) the factor's own prior-one year return and the average prior one-year return of all other factors. SR(1) is the Sharpe ratio for an efficient factor that uses a single signal, its own past return, and SR(2) is the Sharpe ratio for an efficient factor that uses both signals. The efficient factor is set to have the same expected premium as the standard factor; for each of the four signal types, efficient factors of all factors are created using the same function. The asset pricing models are: (1) the Daniel et al. (2019) factors are loadings-hedged versions of the Fama-French factors; (2) the Stambaugh and Yuan (2017) factors cluster 11 prominent anomalies into two clusters; (3) the Hou et al. (2015) model is similar to the Fama-French model except that it rebalances the profitability (ROE) factor monthly using quarterly Compustat data and drops the value (HML) factor; (4) industry-neutral factors are the four non-market factors from the Fama-French model; stocks are sorted into portfolios using characteristics demeaned by industry and industry exposures are hedged: if the factor invests w_i in stock i, it also invests $-w_i$ in the value-weighted industry portfolio into which i belongs; and (5) the Novy-Marx (2013) factors are industry-neutral versions of the gross profitability factor and Carhart's (1997) UMD. The data are monthly factor returns from August 1963 through December 2018, except for the Novy-Marx (2013) factors, for which the data end in December 2012.

		Standard	Signal(s): last month returns				Signal(s): last year returns			
Model	factor	SR	SR(1)	z-stat	SR(2)	z-stat	SR(1)	z-stat	SR(2)	z-stat
Daniel et al. (2019)	MKTRF	0.56	0.90	3.68	0.85	3.18	0.92	4.50	0.89	4.27
loadings-hedged model	SMB	0.41	0.46	0.69	0.47	0.84	0.46	1.50	0.47	1.59
	HML	0.54	0.69	2.60	0.70	2.63	0.59	1.80	0.59	1.99
	RMW	0.47	0.55	1.14	0.56	1.22	0.49	0.62	0.50	0.75
	CMA	0.64	0.66	0.44	0.68	0.71	0.65	0.46	0.65	0.40
Stambaugh and Yuan (2017)	MGMT	0.71	0.77	0.91	0.78	1.06	0.74	0.84	0.75	1.10
four-factor mispricing model	PERF	0.62	0.71	1.10	0.69	0.87	0.61	-0.42	0.61	-0.49
	SMB	0.54	0.59	0.67	0.59	0.58	0.59	1.07	0.59	1.12
Hou et al. (2015)	ME	0.32	0.39	0.87	0.38	0.69	0.37	1.10	0.37	1.05
q-factor model	I/A	0.70	0.76	0.88	0.78	1.09	0.73	0.81	0.74	1.16
	ROE	0.75	0.92	1.92	0.89	1.62	0.77	0.51	0.78	0.76
Industry-neutral	SMB	0.24	0.31	0.76	0.30	0.65	0.33	1.94	0.33	1.84
five-factor model	HML	0.79	0.93	2.32	0.93	2.17	0.85	1.97	0.85	1.85
	RMW	0.55	0.61	1.10	0.61	0.99	0.65	2.50	0.65	2.46
	CMA	0.78	0.86	1.70	0.87	1.64	0.81	1.67	0.81	1.74
Novy-Marx (2013)	PMU	0.68	0.73	0.98	0.73	0.78	0.70	1.07	0.71	1.50
four-factor model	UMD	0.64	0.73	0.98	0.70	0.66	0.66	0.70	0.67	0.91
	AVE.	0.58	0.68	1.34	0.68	1.26	0.64	1.33	0.64	1.41

model are the market and size factors and two clusters formed from 11 anomalies of Novy-Marx (2013). Stambaugh and Yuan (2017) interpret the first cluster, MGMT, as containing factors related to management actions and the second cluster, PERF, as containing factors related to firm performance.

3. **Hou et al. (2015) *q*-factor model.** The factors in this model are the market (MKT), size (ME), profitability (ROE), and investment (I/A). The market, size, and investment factors are the same as those in the Fama and French (2015) five-factor model except for some differences in sample restrictions and the definition of the universe of stocks. The profitability factor is different. Hou et al. (2015) construct this factor by sorting stocks into six portfolios *monthly* by firm size and the income-to-book value of equity computed from the *quarterly* Compustat database.[5] Novy-Marx (2015) suggests that this profitability factor, rather than presenting pure profitability, performs well because it captures momentum in firm fundamentals. That is, it derives its profits, in part, from the post-earnings announcement drift (Ball and Brown, 1968).

4. **Industry-neutral five-factor model.** We follow Novy-Marx (2013) and construct industry-neutral versions of the four non-market factors of the Fama-French model using a two-step process. The first step takes the 49 Fama-French industries and demeans each characteristic within the industry before sorting stocks into portfolios. The second step creates an industry hedge: if stock i's weight in a portfolio is w_i, the industry-hedge portfolio invests $-w_i$ in the value-weighted industry portfolio into which stock i belongs. Industry neutral factors are approximately balanced across industries in their long and short legs; and because of the industry hedge component, industry-wide return shocks leave factors' return largely unchanged.

5. **Novy-Marx (2013) four-factor model.** The factors in this model are the market (MKTRF), size (SMB), profitability (PMU), and momentum (UMD) factors. The profitability

[5] Hou et al. (2015) define income-to-book value of equity as the most recent quarter's income before extraordinary items to previous quarter's book value of equity. They define a firm's book value of equity as the quarterly version of Fama and French's book value of equity (Hou et al., 2015, p. 10).

factor is constructed similar to that in the five-factor model except that, instead of being defined as operating profits-to-book value of equity, Novy-Marx defines the signal as gross profits-to-total assets. The size, profitability, and momentum factors are industry-neutral; the two-step procedure is the same as that described above.

We create time-series efficient versions of those factors that are unique to each model. For example, although the Stambaugh and Yuan (2017) and Hou et al. (2015) include the market factor, this factor was already in the five-factor model, and we therefore do not list it in Table 6.

Similar to Table 2, we construct time-series efficient factors by using month $t-1$ returns as the only conditioning information, and compare the Sharpe ratios of the original and efficient factors. Table 6 shows that the Ferson-Siegel procedure consistently improves factor efficiency: the Sharpe ratios of all 17 factors improve. Although some improvements are economically and statistically modest, this consistency across all factors stands out. The average improvement of 0.10 units (from 0.58 to 0.68) or 17.2% is just below the 22% improved predicted by equation (17) under the assumption that the agent has perfect information on the factor's distribution properties. Instead, the results in the second column of Table 2 are from a naïve strategy that uses the function in equation (15) for all factors. Efficient factors constructed using other signals, like those that we construct using last month returns, all earn higher Sharpe ratios with an exception of PERF whose Sharpe ratio falls by 0.01 unit when we use past-year data as the signal.

The results in Table 6 show that time-series efficiency, as a method for improving factor's Sharpe ratios, is distinct from and complementary to other improvement protocols. The Daniel et al. (2019) factors, for example, are improved versions of the factors in the five-factor model: they enhance Sharpe ratios by hedging out risk that appears not to be compensated in the cross section of stock returns. Whereas the standard market factor's Sharpe ratio, for example, is 0.41 (see Table 2), the market factor in the DRMS model has a Sharpe ratio of 0.56. However, because the Daniel et al. (2019) procedure does not address time-series predictability in factor premiums, time-series efficiency generates additional improvements. The Sharpe ratio of the *time-series efficient* DRMS market factor, for example, is 0.87, and the difference in the Sharpe ratios is significant with a t-value of 3.67, the highest in Table 6. Although DRMS factors are *cross-sectionally* efficient, they are not *time-series* efficient; the gains in Table 6 result from hedging out unpriced variation in factor returns. Similarly, while Novy-Marx's industry neutral HML outperforms the standard HML—these two factors' Sharpe ratios are 0.79 and 0.40—time-series efficiency further increases this factor's Sharpe ratio to 0.94.

5.2 Daily data

In the analyses in Section 3 we constructed time-series efficient versions of the five factors of the Fama-French model at monthly frequency. We used either prior month or prior one-year returns as conditioning information. Factor returns are time-series predictable at other frequencies as well. We illustrate this result in Table 7 by constructing time-series efficient Fama-French factors at daily frequency.

Panel A of Table 7 shows that factors returns are significantly positively autocorrelated at the daily frequency. The average daily return on the market factor, for example, varies from -8 basis points to 12 basis points depending on whether this factor's return was in the top or bottom tercile the prior day. The 20-basis point difference in the averages between the top and bottom terciles is significant with a t-value of 8.96. The market factor is not an aberration. The average factor, constructed by sorting all factors independently into terciles based on their day $d-1$ returns, earns an average return of -5 points when in the bottom tercile, a return of 8 basis points when in the top decile, and the difference between the two has a t-value of 19.42. The daily autocorrelations, shown in the rightmost column, are consistent with these tercile sorts. The value, profitability,

Table 7: Daily Fama-French five-factor model: Time-series predictability and standard versus efficient factors

Panel A assigns Fama-French factors into terciles based on day $d-1$ returns and reports average day d returns (and t-values) for these terciles. "Average" at the bottom of the table is computed by assigning all factors first into terciles and then computing the average returns for each tercile. Autocorrelation, reported in the rightmost column is between day $d-1$ and d returns. Panel B compares annualized Sharpe ratios of standard time-series efficient factors. A time-series efficient factor uses day $d-1$ return as conditioning information to predict the original factor's return. The efficient factor is set to have the same expected premium as the original factor; it uses the unconditional mean, volatility, and first-order autocorrelation of the original factor. The data are daily factor returns from July 1963 through December 2018.

Panel A: Daily time-series predictability

	Day $d-1$ return tercile				Auto-
Factor	Low	Middle	High	H−L	correlation
MKTRF	−0.08	0.03	0.12	0.20	0.05
	(−4.46)	(2.62)	(8.73)	(8.96)	(6.08)
SMB	−0.03	0.00	0.05	0.08	0.04
	(−3.75)	(0.17)	(6.43)	(7.24)	(5.30)
HML	−0.06	0.01	0.10	0.16	0.12
	(−7.50)	(2.01)	(11.52)	(13.57)	(14.31)
RMW	−0.03	0.01	0.06	0.09	0.15
	(−5.80)	(3.03)	(9.54)	(11.06)	(17.77)
CMA	−0.04	0.01	0.08	0.12	0.14
	(−7.60)	(1.44)	(12.39)	(14.34)	(17.04)
Average	−0.05	0.01	0.08	0.13	
	(−10.55)	(4.26)	(17.24)	(19.42)	

Panel B: Sharpe ratios of standard and time-series efficient factors

Factor		Factor				
definition	Statistic	MKTRF	SMB	HML	RMW	CMA
Original factor	Sharpe ratio	0.40	0.24	0.51	0.56	0.59
Efficient factor	Sharpe ratio	1.09	0.60	1.12	1.08	1.26
(Signal: r_{d-1})	Δ Sharpe ratio	0.70	0.36	0.61	0.52	0.67
	z-value	6.01	3.10	5.12	4.47	5.86
	p-value	0.00	0.00	0.00	0.00	0.00

and investment factors, in particular, are significantly positively autocorrelated. These factors' first-order autocorrelations range from 0.12 to 0.15.

Panel B of Table 7 constructs time-series efficient factors by using this daily predictability in factor premiums to time each factor. Given the amount of predictability, all five efficient versions of the factors significantly outperform their standard counterparts. The annualized Sharpe ratio of the market increases from 0.40 to 1.09; that of the size factor from 0.24 to 0.60; and so forth.[6] The large improvements at the daily frequency are consistent with the prediction of equation (17): because daily and monthly autocorrelations are very similar, but the daily Sharpe ratios are much lower, the attainable gains in Sharpe ratios are significantly larger.

We caution *against* interpreting the results in Table 7 as suggesting that an investor can increase his or her portfolio's Sharpe ratio from 0.40 to 1.09 by switching from the market factor to the time-series efficient market factor. While an investor can earn the Sharpe ratio of 0.40 by following a passive, near-buy and hold[7] investment strategy, the time-series efficient market factor calls for daily rebalancing; the costs of daily rebalancing can easily negate all of the "paper efficiency gains." At the same time, as we discuss in Section 7, the efficient factors, despite any transaction costs, serve a vital function: if a researcher benchmarks *new* factors against a model consisting of inefficient factors, these new factors may seem valuable—but only because they correlate with the inefficiencies found in the original factors. The daily time-series efficient five-factor model raises the bar for new factors by increasing the squared Sharpe ratio of the ex-post optimal portfolio to 4.

[6]The Sharpe ratios of the standard factors in Panel B of Table 7 differ from those reported in Table 2 because of the sampling frequency. In Table 2 a factor's Sharpe ratio is its average monthly return divided by the standard deviation of its monthly returns. In Table 7 it is the factor's average daily return divided by the standard deviations of its *daily* returns. The Sharpe ratios differ because of compounding and, more importantly, because factor returns are autocorrelated; that is, the differences in the Sharpe ratios of the daily and monthly factors in itself indicates that the factors significantly autocorrelate at the daily frequency. A comparison of Sharpe ratios at different frequencies is akin to the variance-ratio test of Lo and MacKinlay (1988).

[7]Pedersen (2018) suggests that some caveats apply: an investor who wants to track the market needs to trade as well because the "market" changes with the entry of new and the disappearance of old companies and when existing firms issue new and repurchase old shares.

Table 8: Are time-series efficient factors more systematic?

This table reports estimates from time-series regressions in which the dependent variable is the first principal component extracted from the long and short legs of the Novy-Marx and Velikov (2015) anomalies. The first model is the standard Fama and French (2015) five-factor model; the second model is the time-series efficient five-factor model; and the third model is the union of these two models. For each model we run two regressions, an unconstrained OLS and a constrained regression (CSR) of the form $b \geq 0$. The time-series efficient factors use the prior one-year return as the conditioning information to predict the standard factor's return. The efficient factor targets to earn the same expected premium as the standard factor; the efficient factors use the full-sample means, volatilities, and first-order autocorrelations of the standard factors. The two F-values at the bottom of the table test the null hypothesis that the slopes on all standard factors or all efficient factors are jointly zero. In the first two models, these test statistics are distributed $F(5, 599)$; in the last model, both are distributed $F(5, 594)$. The data are monthly factor returns from August 1963 through December 2018.

		Model					
		(1)	(2)	(3)	(4)	(5)	(6)
Model	Factor	OLS	CSR	OLS	CSR	OLS	CSR
Standard FF5	MKTRF	-0.03	0.00			-0.08	0.00
		-2.92	0.00			-4.99	0.00
	SMB	0.29	0.29			0.05	0.04
		19.64	21.22			1.69	1.13
	HML	0.25	0.24			0.06	0.00
		11.98	11.67			1.32	0.00
	RMW	-0.05	0.00			-0.06	0.00
		-2.22	0.00			-1.50	0.00
	CMA	0.41	0.44			0.32	0.35
		13.21	15.29			3.18	3.88
Efficient FF5	MKTRF			0.01	0.01	0.09	0.01
				0.48	0.48	4.82	1.26
	SMB			0.34	0.34	0.28	0.30
				23.07	23.07	7.44	8.07
	HML			0.35	0.35	0.30	0.35
				15.88	15.88	6.39	16.38
	RMW			0.03	0.03	0.09	0.05
				1.21	1.21	1.69	1.98
	CMA			0.46	0.46	0.06	0.07
				16.29	16.29	0.57	0.72
N		605	605	605	605	605	605
adj. R^2		71.6%	71.2%	78.0%	78.0%	79.3%	78.5%
F-test: Standard		305.6				7.7	
F-test: Efficient				444.9		49.1	

6 Are time-series efficient factors more systematic?

Kozak et al. (2018) consider an economy in which factor premiums are disjointed from the underlying risks in the economy. In their model sentiment-driven investors induce mispricing, and rational arbitrageurs trade against these mispricings for profit. The point of Kozak et al. (2018) is that when the mispricings happen to align with high-eigenvalue principal components (PCs), arbitrageurs face risk: they are unwilling to trade so aggressively as to eliminate all mispricings. As a consequence, the largest factor premiums associate with the most "systematic" factors even though the premiums do not stem from the risks themselves. Kozak et al. (2018) measure how "systematic" different factors are by taking the long and short legs of the Novy-Marx and Velikov (2015) anomalies and extracting the principal components from this set. Ordered by their corresponding eigenvalues, the first PC captures most of the variation in anomaly returns; the second PC the second most; and so forth.

We follow Kozak et al. (2018) and start with the same Novy-Marx and Velikov (2015) anomaly data at the monthly frequency. We extract the first principal component and measure the extent to which this component aligns with the original and the time-series efficient five-factor models. Because all factors are designed to earn positive premiums, the first PC is (almost) equally increasing in the returns on most factors; this PC can therefore also be viewed as the average risk premium in the economy. The time-series efficient five-factor model that we consider is the one that uses month $t-12$ returns as the conditioning information. In Table 8 we report estimates from regressions in which the dependent variable is the first principal component and the model is either the standard or efficient five-factor model, or both models at the same time. We estimate these regressions both as unconstrained regressions as well as constrained regressions with the restriction that the slopes are nonnegative, $b \geq 0$.

An important "true" factor's premium should correlate positively with the aggregate factor premia. Consider an extreme case when all returns are driven by a single factor. In this economy, the correlation between the first principal component of the cross-section of returns and the factor will be close to one. This is the idea for imposing restriction on the slope coefficients.

The first column of Table 8 shows that the original five factors of the Fama-French model explain 71.6% of the time-series variation in the first principal component. The loadings on the size, value, and investment factors stand out; each of these factors is significant with a t-value above ten; and a test for the joint significance of the five factors returns an F-value of 305.6. The second column of Table 8 shows that the restricted regression explains almost as much variation in PC1 as the OLS regression. This is consistent with the conjecture that PC1 is increasing in the main factors and excluding others has little to no impact on the model: excluding MKTRF and RMW does almost nothing to the model's explanatory power.[8]

The third column of Table 8 shows that the efficient version of the same model performs at least as well: the explanatory power increases to 78.6% and both the slopes and t-values of *all* factors increase from those in the first column. The test statistic for the joint significance of the five efficient factors is now 444.9. Because all slopes in the unconstrained model are positive, the estimates from the constrained model are identical to those from the unconstrained model.

Having set the stage, the fifth column runs a horse-race regression between the two models. How informative are the standard factors of the five-factor model when we control for the efficient factors, or vice versa? Although the factors in the two models, by construction, correlate significantly, the results clearly favor the efficient five-factor model. Only one of the slopes on the original factors, the investment factor (CMA), is positive and statistically significant, and that on the market is negative and statistically significantly with a t-value of -4.88. All slopes on the time-series efficient factors, by contrast, are positive and three of them are statistically significant. The F-values of 7.7

[8] The lack of significance for MKTRF goes back to the choice of the 15 long-short anomalies used to extract PC1. The anomalies are designed to be market-neutral, thus rendering the market factor less important.

(standard factors) and 49.1 (efficient factors) also indicate that it is the efficient factors that do most of the work in the combined model.

The sixth model excludes the standard MKTRF, HML, and RMW; the only relevant factor from the standard FF5 is CMA with a coefficient of 0.35 that is significant with a t-stat of 3.88. Excluding almost all but CMA from the standard FF5 set has minimal impact on the regression's R^2. The slopes on all efficient factors remain positive and highly significant for SMB and HML, and marginally significant for RMW. A comparison of the three models' explanatory powers puts this result into perspective: adding the efficient factors on top of the standard model increases the model's explanatory power by 8.1 percentage points; but adding the standard factors on top of the efficient model increases the power by just 1.1 percentage points. Further, if we require our factors to correlate positively with aggregate factor premium, we will be left with the efficient five-factor model with a tilt towards the standard CMA.

If we interpret the first PC as being the most important part of the economy's SDF (Kozak et al., 2018), then the results in Table 8 suggest that the efficient five-factor model aligns more closely with the true SDF than the standard five-factor model.

7 Conclusions

Asset pricing factors are positively autocorrelated. This autocorrelation appears to manifest as momentum in the cross section of stock returns. Because different assets have different loadings on the underlying factors, the momentum in the factors transmits into individual stock returns. One way for an investor to exploit the time-series predictability in factor returns is to trade momentum as a separate factor. In this paper, we show that, instead of doing so, an investor (or an asset pricer) can instead redefine the factors to push the autocorrelations back into the factors themselves. We call factors that exhaust their own time-series predictability *time-series efficient factors*.

We show that time-series efficient factors outperform their original counterparts. In the Fama and French (2015) five-factor model, all efficient factors earn higher Sharpe ratios than the original factors, and the time-series efficient factors contain most—or all, depending on the definition of the conditioning information—of the information found in the original factors. This result is not specific to the five-factor model: time-series efficiency improves Sharpe ratios of the factors in all popular asset pricing models. Although some of the improvements in Sharpe ratios are modest, many are both economically and statistically highly significant. Importantly, given an investor's beliefs about the original factor's moments, an investor can compute ex ante the expected efficiency gain; if a factor's premium is not time-series predictable, no gains can be expected, and there is no need to bother. Because time-series efficiency gains and momentum profits stem from the same source, factor autocorrelations, the momentum factor becomes redundant when an investor or a builder of asset pricing models switches from the original factors to time-series efficient factors.

Our results are important from both investing and asset pricing perspectives. Because factor autocorrelations are persistent, the increases in Sharpe ratios at the monthly frequency can cover any trading costs. A reference to the cross-sectional momentum effect buttresses this point: if standard momentum strategies are profitable net of trading costs (Asness et al., 2014), then time-series efficiency must be profitable as well—the two effects stem from the same source.

Moreover, from the perspective of Cochrane's (2011) factor zoo, time-series efficient factors serve a vital purpose no matter what their transaction costs are. The time-series predictability present in the original factors implies that the original factors are inefficient; and that, without introducing additional factors, these factors can be greatly improved. The corollary to this point is that when researchers use the *original* factors as the benchmarks, they may easily mistake indirect efficiency gains for novel anomalies. An additional factor that they propose may seem to yield Sharpe ratio improvements—but only because it correlates with the inefficiencies present in the original factors.

By using time-series efficient factors in the asset pricing models, we can ensure that any proposed factor serve a purpose other than extracting such indirect efficiency gains.

References

Arnott, Rob, Mark Clements, Vitali Kalesnik, and Juhani T. Linnainmaa, 2019, Factor momentum, Working paper.

Asness, Clifford, Andrea Frazzini, Ronen Israel, and Tobias Moskowitz, 2014, Fact, fiction, and momentum investing, *Journal of Portfolio Management* 40, 75–92.

Asness, Clifford S., Tobias J. Moskowitz, and Lasse Heje Pedersen, 2013, Value and momentum everywhere, *Journal of Finance* 68, 929–985.

Asness, Clifford S., R. Burt Porter, and Ross L. Stevens, 2000, Predicting stock returns using industry-relative firm characteristics, AQR Capital Management working paper.

Ball, Ray, and Philip Brown, 1968, An empirical evaluation of accounting income numbers, *Journal of Accounting Research* 159–178.

Barillas, Francisco, and Jay Shanken, 2017, Which alpha? *Review of Financial Studies* 30, 1316–1338.

Bekaert, Geert, and Jun Liu, 2004, Conditioning information and variance bounds on pricing kernels, *The Review of Financial Studies* 17, 339–378.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Cederburg, Scott, Michael S. O'Doherty, Feifei Wang, and Xuemin Sterling Yan, 2019, On the performance of volatility-managed portfolios, *Journal of Financial Economics*, forthcoming.

Cochrane, John H, 2011, Presidential address: Discount rates, *Journal of Finance* 66, 1047–1108.

Cohen, Randolph B, Christopher Polk, and Tuomo Vuolteenaho, 2003, The value spread, *Journal of Finance* 58, 609–641.

Cohen, Randolph B., and Christopher K. Polk, 1998, The impact of industry factors in asset-pricing tests, Kellogg Graduate School of Management working paper.

Daniel, Kent, Lira Mota, Simon Rottke, and Tano Santos, 2019, The cross section of risk and return, Working paper.

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the characteristics of cross sectional variation in stock returns, *Journal of Finance* 52, 1–33.

Ehsani, Sina, and Juhani Linnainmaa, 2019, Factor momentum and the momentum factor, NBER Working Paper No. 25551.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F, and Kenneth R French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55–84.

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, 2016, Dissecting anomalies with a five-factor model, *Review of Financial Studies* 29, 69–103.

Ferson, Wayne, 2019, *Empirical Asset Pricing: Models and Methods* (MIT Press).

Ferson, Wayne E., and Andrew F. Siegel, 2001, The efficient use of conditioning information in portfolios, *Journal of Finance* 56, 967–982.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121–1152.

Goyal, Amit, and Narasimhan Jegadeesh, 2017, Cross-sectional and time-series tests of return predictability: What is the difference? *Review of Financial Studies* 31, 1784–1824.

Greenwood, Robin, and Samuel G. Hanson, 2012, Share issuance and factor timing, *Journal of*

Finance 67, 761–798.

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

Hansen, Lars Peter, and Thomas J Sargent, 2008, *Robustness* (Princeton University Press, Princeton, NJ).

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *Review of Financial Studies* 28, 650–705.

Huberman, Gur, and Shmuel Kandel, 1987, Mean-variance spanning, *Journal of Finance* 42, 873–888.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

Jobson, J. Dave, and Bob M. Korkie, 1981, Performance hypothesis testing with the Sharpe and Treynor measures, *Journal of Finance* 36, 889–908.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2018, Interpreting factor models, *Journal of Finance* 73, 1183–1223.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2019, Shrinking the cross-section, *Journal of Financial Economics* .

Lintner, John, 1965, Security prices, risk, and maximal gains from diversification, *Journal of Finance* 20, 587–615.

Liu, Fang, Xiaoxiao Tang, and Guofu Zhou, 2019, Volatility-managed portfolio: Does it really work? *Journal of Portfolio Management* 46, 38–51.

Lo, Andrew W, and A Craig MacKinlay, 1988, Stock market prices do not follow random walks: Evidence from a simple specification test, *Review of Financial Studies* 1, 41–66.

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

Memmel, Christoph, 2003, Performance hypothesis testing with the Sharpe ratio, *Finance Letters* 1, 21–23.

Moreira, Alan, and Tyler Muir, 2017, Volatility-managed portfolios, *Journal of Finance* 72, 1611–1644.

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

Novy-Marx, Robert, 2015, How can a q-theoretic model price momentum? NBER Working Paper No. 20985.

Novy-Marx, Robert, and Mihail Velikov, 2015, A taxonomy of anomalies and their trading costs, *Review of Financial Studies* 29, 104–147.

Pedersen, Lasse Heje, 2018, Sharpening the arithmetic of active management, *Financial Analysts Journal* 74, 21–36.

Sharpe, William F, 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.

Stambaugh, Robert F., and Yu Yuan, 2017, Mispricing factors, *Review of Financial Studies* 30, 1270–1315.

A Appendix

A.1 Deriving ζ and the optimal-to-original Sharpe ratio in equation (17)

Constant ζ is required for computing the optimal weight in the Ferson-Siegel procedure. In this appendix we derive ζ of the weight function of the portfolio when the portfolio's past return is used to forecast future returns: $\tilde{R} = \mu(\tilde{S}) + \tilde{\epsilon}$, in which $\mu(\tilde{S}) = \mu + \rho \tilde{R}_{t-1}$. The portfolio's unconditional expected return is $\mu_p = \frac{\mu}{1-\rho}$. This portfolio's Sharpe ratio is

$$SR = \frac{\mu_p}{\sqrt{\sigma^2_{\mu(\tilde{S})} + \sigma^2_{\epsilon(\tilde{S})}}}. \tag{A-1}$$

Because returns follow an AR(1) process, we can write

$$\frac{\sigma^2_{\mu(\tilde{S})}}{\sigma^2_{\mu(\tilde{S})} + \sigma^2_{\epsilon(\tilde{S})}} = \rho^2. \tag{A-2}$$

This ratio can be interpreted as the R^2 from a regression of returns on past returns. In a univariate regression R^2 is the squared the correlation coefficient, which in this case is also the AR(1) coefficient, ρ. Rearranging and assuming homoskedasticity, $\sigma^2_\epsilon(\tilde{S}) = \sigma^2_\epsilon$, we have

$$\sigma^2_\epsilon = \left(\frac{1}{\rho^2} - 1\right) \sigma^2_{\mu(\tilde{S})}. \tag{A-3}$$

We can now compute ζ as follows:

$$\begin{aligned}
\zeta &= \mathbb{E}\left[\frac{\mu^2(\tilde{S})}{\mu^2(\tilde{S})+\sigma_\epsilon^2}\right] = \mathbb{E}\left[1-\frac{\sigma_\epsilon^2}{\mu^2(\tilde{S})+\sigma_\epsilon^2}\right] = 1-\sigma_\epsilon^2\,\mathbb{E}\left[\frac{1}{\mu^2(\tilde{S})+\sigma_\epsilon^2}\right] \\
&= 1-\frac{\sigma_\epsilon^2}{\mathbb{E}\big[\mu^2(\tilde{S})+\sigma_\epsilon^2\big]} + \frac{\mathrm{var}[\mu^2(\tilde{S})+\sigma_\epsilon^2]}{\mathbb{E}[\mu^2(\tilde{S})+\sigma_\epsilon^2]^3} \approx 1-\frac{\sigma_\epsilon^2}{\mathbb{E}\big[\mu^2(\tilde{S})\big]+\sigma_\epsilon^2} \\
&= 1-\frac{\sigma_\epsilon^2}{\mu^2+\sigma_{\mu(S)}^2+\sigma_\epsilon^2} = 1-\frac{\sigma_\epsilon^2}{SR^2[\sigma_{\mu(S)}^2+\sigma_\epsilon^2]+\sigma_{\mu(S)}^2+\sigma_\epsilon^2} \\
&= \frac{SR^2+\rho^2}{SR^2+1}.
\end{aligned} \tag{A-4}$$

The approximation on the second line uses Taylor series expansion and homoskedasticity in returns, and the final result uses equation (A-3) to remove variances from the identity.[9]

We now derive equation (17). The mean and variance of the Ferson-Siegel portfolio are μ_p and $\mu_p^2(\frac{1}{\zeta}-1)$. The Sharpe ratio squared becomes:

$$SR^{*2} = \frac{\mu_p^2}{\mu_p^2(\frac{1}{\zeta}-1)} = \frac{1}{\frac{1}{\zeta}-1} = \frac{\zeta}{1-\zeta} = \frac{\frac{SR^2+\rho^2}{SR^2+1}}{1-\frac{SR^2+\rho^2}{SR^2+1}} = \frac{SR^2+\rho^2}{1-\rho^2} = \frac{SR^2(1+(\frac{\rho}{SR})^2)}{1-\rho^2}. \tag{A-5}$$

We can therefore write the ratio of the efficient factor's Sharpe ratio to the original factor's Sharpe ratio as

$$\frac{SR^*}{SR} = \sqrt{\frac{1+(\frac{\rho}{SR})^2}{1-\rho^2}}, \tag{A-6}$$

which is equation (17).

[9] Using the Taylor series expansion around $\mathbb{E}[X]$ we have:

$$\mathbb{E}\left[\frac{1}{X}\right] \approx \mathbb{E}\left[\frac{1}{\mathbb{E}[X]} - \frac{1}{\mathbb{E}[X]^2}(X-\mathbb{E}[X]) + \frac{1}{\mathbb{E}[X]^3}(X-\mathbb{E}[X])^2\right] = \frac{1}{\mathbb{E}[X]} + \frac{\mathrm{var}[X]}{\mathbb{E}[X]^3}$$

In our framework, the variance of X is the "vol of vol," which is negligible under the homoskedasticity assumption.

A.2 Covariance between the efficient and original factors

Equations (18) and (19) give the test statistic z for the expected difference in the Sharpe ratios between the efficient and original factors. In addition to the means and standard deviations, this test statistic also depends on the covariance between the efficient and original factors. We derive an approximation for this covariance in this appendix.

Omitting subscripts, we let R denote the return to the original factor and x the efficient factor's weight on the original factor. The return on the efficient factor is thus xR and the covariance that we need to compute is $\text{cov}(xR, R)$. This covariance can be expressed as:

$$\begin{aligned}
\text{cov}(xR, R) &= \mathbb{E}[xR^2] - \mathbb{E}[xR]\mathbb{E}[R] \\
&= \mathbb{E}[xR^2] - \mathbb{E}[R]\big[\text{cov}(x, R) + \mathbb{E}[x]\mathbb{E}[R]\big] \\
&= \mathbb{E}[xR^2] - \mathbb{E}[R]\text{cov}(x, R) - \mathbb{E}[R]^2\mathbb{E}[x].
\end{aligned} \tag{A-7}$$

We therefore have to compute $\mathbb{E}[R] = \mu_p$, $\mathbb{E}[x]$, $\text{cov}(x, R)$, and $\mathbb{E}[xR^2]$. We start by computing $\mathbb{E}[x]$:

$$\begin{aligned}
\mathbb{E}[x] &= \frac{\mu_p}{\zeta}\,\mathbb{E}\left[\frac{\mu(\tilde{S})}{\mu(\tilde{S})^2 + \sigma_\epsilon^2}\right] \\
&= \frac{\mu_p}{\zeta}\left[\mathbb{E}[\mu(\tilde{S})] \cdot \mathbb{E}[\frac{1}{\mu(\tilde{S})^2 + \sigma_\epsilon^2}] + \text{cov}\big(\mu(\tilde{S}), \frac{1}{\mu(\tilde{S})^2 + \sigma_\epsilon^2}\big)\right] \\
&\approx \frac{\mu_p}{\zeta}\left[\frac{\mu_p}{\mu^2 + \sigma^2_{\mu(S)} + \sigma_\epsilon^2}\right] = \frac{\mu_p}{\zeta}\left[\frac{\mu_p}{(\sigma_\mu^2 + \sigma_\epsilon^2)(1 + SR^2)}\right] \\
&= \frac{SR^2}{1 + SR^2} \cdot \frac{1}{\zeta} = \frac{SR^2}{1 + SR^2} \cdot \frac{SR^2 + 1}{SR^2 + \rho^2} = \frac{SR^2}{SR^2 + \rho^2}.
\end{aligned} \tag{A-8}$$

Equation (A-8) suggests that the efficient factor's average time-series exposure to a factor decreases as the factor's autocorrelation coefficient increases. In other words, the strategy "trusts" the signal (past returns) more than the unconditional factor's Sharpe ratio when the autocorrelation coefficient is high. By contrast, when the autocorrelation is low, the strategy prefers to invest less in the signal and more in the factor to secure the unconditional Sharpe ratio provided by the factor.

We next compute $\text{cov}(x, R)$:

$$\begin{aligned}
\text{cov}(x, R) &= \text{cov}\left(\frac{\mu_p}{\zeta}\frac{\mu(\tilde{S})}{\mu(\tilde{S})^2 + \sigma_\epsilon^2}, \mu(\tilde{S}) + \tilde{\epsilon}\right) = \text{cov}\left(\frac{\mu_p}{\zeta}\frac{\mu + \rho\tilde{R}_{t-1}}{\mu(\tilde{S})^2 + \sigma_\epsilon^2}, \mu + \rho\tilde{R}_{t-1} + \tilde{\epsilon}\right) \quad \text{(A-9)} \\
&= \frac{\mu_p}{\zeta}\text{cov}\left(\frac{\rho\tilde{R}_{t-1}}{\mu(\tilde{S})^2 + \sigma_\epsilon^2}, \rho\tilde{R}_{t-1}\right) = \frac{\mu_p}{\zeta}\rho^2(\sigma^2_{\mu(\tilde{S})} + \sigma^2_\epsilon) \cdot \mathbb{E}[\frac{1}{\mu(\tilde{S})^2 + \sigma_\epsilon^2}] \\
&= \frac{\mu_p}{\zeta}\rho^2(\sigma^2_{\mu(\tilde{S})} + \sigma^2_\epsilon) \cdot [\frac{1}{(\sigma^2_{\mu(\tilde{S})} + \sigma^2_\epsilon)(1 + SR^2)}] = \frac{\mu_p}{\zeta}\rho^2\frac{1}{(1 + SR^2)} \\
&= \mu_p \cdot \frac{SR^2 + 1}{SR^2 + \rho^2} \cdot \rho^2 \cdot \frac{1}{(1 + SR^2)} = \frac{\mu_p \cdot \rho^2}{SR^2 + \rho^2}.
\end{aligned}$$

The term $\mathbb{E}[xR^2]$ can be computed as follows:

$$\begin{aligned}
\mathbb{E}[xR^2] &= \mathbb{E}[x] \cdot \mathbb{E}[R^2] + \text{cov}(x, R^2) = \frac{SR^2}{SR^2 + \rho^2} \cdot (\mu_p^2 + \sigma^2_{\mu(\tilde{S})} + \sigma^2_{\epsilon(\tilde{S})}) + 0 \qquad \text{(A-10)} \\
&= \frac{SR^2}{SR^2 + \rho^2} \cdot \left(\mu_p^2 + \left(\frac{\mu_p}{SR}\right)^2\right) = \frac{SR^2}{SR^2 + \rho^2} \cdot \mu_p^2\left(1 + \frac{1}{SR^2}\right) \\
&= \frac{SR^2}{SR^2 + \rho^2} \cdot \mu_p^2\left(\frac{1 + SR^2}{SR^2}\right) = \mu_p^2 \cdot \frac{SR^2 + 1}{SR^2 + \rho^2}.
\end{aligned}$$

Putting it all together, we can compute $\text{cov}(xR, R)$ as follows:

$$\begin{aligned}
\text{cov}(xR, R) &= \mathbb{E}[xR^2] - \mathbb{E}[R]\text{cov}(x, R) - \mathbb{E}[R]^2\mathbb{E}[x] \qquad \text{(A-11)} \\
&= \mu_p^2 \cdot \frac{SR^2 + 1}{SR^2 + \rho^2} - \mu_p \cdot \frac{\mu_p \cdot \rho^2}{SR^2 + \rho^2} - \mu_p^2 \cdot \frac{SR^2}{SR^2 + \rho^2} \\
&= \frac{\mu_p^2(SR^2 + 1 - \rho^2 - SR^2)}{SR^2 + \rho^2} = \frac{\mu_p^2(1 - \rho^2)}{SR^2 + \rho^2}.
\end{aligned}$$

We use this covariance between the efficient and original factors to compute the z-value for the *predicted* improvement in Sharpe ratios.

A.3 Time-series efficient factor when the signal is the average of past n month returns

In Section 2 we use the Ferson-Siegel method to create efficient factors when the conditioning information is the factor's month $t-1$ return. In this appendix we extend this framework. We use the factor's average return over the past n months as the signal. We maintain the assumption that returns follow an AR(1) process,

$$\mathbb{E}[R_t|S_{t,n}] = \mathbb{E}[R_t|R_{t-n,t-1}] = \beta_0 + \beta_n \tilde{R}_{t-n,t-1}. \tag{A-12}$$

In Section 2 we used the same approach but only for the first lag; in this case, the expectation simplifies to

$$\mathbb{E}[R_t|R_{t-1}] = \beta_0 + \beta_1 \tilde{R}_{t-1} = \mu + \rho R_{t-1}. \tag{A-13}$$

Under the AR(1) assumption, a factor's return at time t can be described as a function of the other lags using a recursive expansion:

$$\begin{aligned}
\tilde{R}_t &= \mu + \rho \tilde{R}_{t-1} + \epsilon_t \qquad \text{(A-14)} \\
&= \mu + \rho(\mu + \rho \tilde{R}_{t-2} + \epsilon_{t-1}) + \epsilon_t \\
&= \mu(1+\rho) + \rho^2 \tilde{R}_{t-2} + \rho\epsilon_{t-1} + \epsilon_t \\
&= \mu(1+\rho+\rho^2) + \rho^3 \tilde{R}_{t-3} + \rho^2\epsilon_{t-2} + \rho\epsilon_{t-1} + \epsilon_t \\
&= \cdots = \mu \sum_{k=0}^{n-1} \rho^k + \rho^n \tilde{R}_{t-n} + \sum_{k=0}^{n-1} \rho^k \epsilon_{t-k}.
\end{aligned}$$

For an AR(1) process, the moments that we need are:

$$\mathbb{E}[\tilde{R}_t] = \frac{\mu}{1-\rho}, \tag{A-15}$$

$$\sigma_R^2 = \frac{\sigma_\epsilon^2}{1-\rho^2}, \tag{A-16}$$

$$\text{cov}(\tilde{R}_t, \tilde{R_{t-k}}) = \rho^k \frac{\sigma_\epsilon^2}{1-\rho^2}. \tag{A-17}$$

We define the signal as the average of the past n months of returns:

$$\tilde{S}_{t,n} = \frac{1}{n}\sum_{k=1}^{n} \tilde{R_{t-k}} = \frac{1}{n}\big(\tilde{R}_{t-1} + \tilde{R}_{t-2} + \cdots + \tilde{R}_{t-n}\big). \tag{A-18}$$

Our objective is to modify the factor weighting program to accommodate the signal of this average-return form. We compute the expected value of the signal, its variance, and its covariance with expected returns:

$$\mathbb{E}[\tilde{S}_{t,n}] = \frac{1}{n}\big(n\frac{\mu}{1-\rho}\big) = \frac{\mu}{1-\rho}, \tag{A-19}$$

$$\begin{aligned}
\text{var}(\tilde{S}_{t,n}) &= \frac{1}{n^2}\left[n\frac{\sigma_\epsilon^2}{1-\rho^2} + \sum_{\substack{i=1 \\ i\neq j}}^{n}\sum_{j=1}^{n}\text{cov}(\tilde{R}_{t-i}, \tilde{R}_{t-j})\right] \\
&= \frac{1}{n^2}\left[n\frac{\sigma_\epsilon^2}{1-\rho^2} + 2\sum_{k=1}^{n}(n-k)\rho^k\frac{\sigma_\epsilon^2}{1-\rho^2}\right] \\
&= \frac{\sigma_\epsilon^2}{n(1-\rho^2)}\left[1 + \frac{2}{n}\sum_{k=1}^{n}(n-k)\rho^k\right],
\end{aligned} \tag{A-20}$$

$$\begin{aligned}
\text{cov}(\tilde{R}_t, \tilde{S}_{t,n}) &= \text{cov}\bigg(\tilde{R}_t, \frac{1}{n}\big(\tilde{R}_{t-1} + \tilde{R}_{t-2} + \cdots + \tilde{R}_{t-n}\big)\bigg) = \frac{1}{n}\sum_{k=1}^{n}\rho^k\frac{\sigma_\epsilon^2}{1-\rho^2} \\
&= \frac{\sigma_\epsilon^2}{n(1-\rho^2)}\sum_{k=1}^{n}\rho^k.
\end{aligned} \tag{A-21}$$

The slope coefficient from a regression of R_t on the signal, the average of past n months of returns, is:

$$\beta_n = \frac{\mathrm{cov}(\tilde{R}_t, \tilde{S}_{t,n})}{\mathrm{var}(\tilde{S}_{t,n})} = \frac{\frac{\sigma_\epsilon^2}{n(1-\rho^2)} \sum_{k=1}^{n} \rho^k}{\frac{\sigma_\epsilon^2}{n(1-\rho^2)} \left[1 + \frac{2}{n} \sum_{k=1}^{n} (n-k)\rho^k\right]} = \frac{\sum_{k=1}^{n} \rho^k}{1 + \frac{2}{n} \sum_{k=1}^{n} (n-k)\rho^k}. \tag{A-22}$$

When the value of ρ is not too large, this expression is approximately equal to ρ. This result means that, regardless of the choice of lag, we can write:

$$\mathbb{E}[R_t|S_{t,n}] = \mathbb{E}[R_t|R_{t-n,t-1}] = \beta_0 + \beta_n \tilde{R}_{t-n,t-1} \sim \mu + \rho R_{t-n,t-1}. \tag{A-23}$$

The correlation coefficient between the return R_t and the signal $S_{t,n}$ equals,

$$\rho_n = \frac{\mathrm{cov}(\tilde{R}_t, \tilde{S}_{t,n})}{\sqrt{\mathrm{var}(\tilde{S}_{t,n})}\sqrt{\sigma_R^2}} = \frac{\frac{\sigma_\epsilon^2}{n(1-\rho^2)} \sum_{k=1}^{n} \rho^k}{\sqrt{\frac{\sigma_\epsilon^2}{n(1-\rho^2)} \left[1 + \frac{2}{n} \sum_{k=1}^{n} (n-k)\rho^k\right]}\sqrt{\frac{\sigma_\epsilon^2}{1-\rho^2}}} = \frac{\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \rho^k}{\sqrt{\left[1 + \frac{2}{n} \sum_{k=1}^{n} (n-k)\rho^k\right]}}. \tag{A-24}$$

The value of ρ_n decreases rapidly in the number of lags. This decrease, in turn, implies that we should expect to be able to explain a smaller share of the variation in expected returns as we increase the lag. That is, under the assumption of the AR(1) process, our ability to improve factors' Sharpe ratios should weaken. In the data, however, factor returns do not appear to follow AR(1) processes. Panel C of Table 1 shows that the prior one-year return correlates more with month t returns than what Panel B together with an AR(1) assumption would suggest. In the data factor premiums therefore seem "stickier" than what they should be under the AR(1) assumption.

We can compute the expected factor return conditional on the signal (the average past return), and the correlation between the factor's return and signal. This correlation, as before, is also the R^2 from a regression of the expected return on the signal. The efficient factor's time-t weight to the minimum-variance efficient factor, conditional on average past n-month return, becomes:

$$x(S_t) = \frac{\mu_p}{\zeta}\Big(\frac{\mu(S_t)}{\mu(S_t)^2 + \sigma_\epsilon^2}\Big), \tag{A-25}$$

$$\zeta = \frac{SR^2 + \rho_n^2}{SR^2 + 1}, \tag{A-26}$$

$$\mu(S_t) = \mu_p\,(1-\rho) + \beta_n\,R_{t-n,t-1}. \tag{A-27}$$

Compared to equation (8), ζ is smaller because ρ_n decreases rapidly in n. A smaller ζ is undesirable from the efficiency viewpoint because the expected reduction in variance is its function; thus, our ability to improve Sharpe ratios falls. Put differently, if factor returns indeed follow AR(1) processes, the trader should use only the most recent returns to generate time-series efficient factors; adding more lags will reduce the explanatory power.

The average factor in the sample has unconditional sample moments of about $\rho = 0.10$. We derive in equation (A-24) and equation (A-22) the correlation (between return at time t and average past T month returns) as a function of the lag length, assuming returns are AR(1). However, studies such as Arnott et al. (2019) and Ehsani and Linnainmaa (2019) find that factor returns exhibit autocorrelation at every lag between 1 and 12, followed by reversal from lag 13 and after. according to equation (A-24), AR(1) monthly returns with an autocorrelation coefficient of 0.10 should result in a correlation of less than 0.03 between month t and average return between $t-12$ and $t-1$. But in data, this correlation is 0.09, almost as high as the AR(1) estimate, consistent with positive serial-correlation at lags beyond 1, and the profitability of momentum strategies that use returns over the $\{t-12, t-1\}$ period as their signal. For this estimate to be consistent with a simple AR(1) process, the AR(1) coefficient should be as high as 0.30. We use the following weight program, which assume some correlation beyond lag 1 for the program:[10]

[10] Slightly more aggressive weighting schedules than the one used here give stronger results for the mean-variance investor that uses last year return as her trading signal.

$$x(S_t) = 8.01 \left(\frac{0.28 + 0.173 R_{t-12,t-1}}{[0.28 + 0.173 R_{t-12,t-1}]^2 + 3.99} \right). \qquad \text{(A-28)}$$

A.4 Time-series efficient factors when a dominant factor is informative about individual factor returns

We extend the model to accommodate information in a dominant factor's prior return, that is, in the first "level" principal component. This extension is motivated by Kozak et al. (2018), who find that only the first few principal component obtained by long short anomaly portfolios are informative about the cross section of returns out-of-sample. In this section we create time-series efficient factors under the assumption that individual factor returns are linear in their own past returns and that of the dominant factor.

A factor return is described by the following model:[11]

$$\tilde{R_{f,t}} = \mu + \rho \tilde{R}_{f,t-1} + \rho' \tilde{R}'_{t-1} + \epsilon_t, \qquad \text{(A-29)}$$

where $\tilde{R}'_{t-1}$ is the dominant factor's past period return. We assume that $\tilde{R}_{f,t}$ and $\tilde{R}'_t$ are not correlated for analytical convenience. The investor's signal in this case is a function of the factor's own past return and that of the dominant factor, $\mu(S_t) = \mu + \rho \tilde{R}_{f,t-1} + \rho' \tilde{R}'_{t-1}$.

To compute Ferson-Siegel weights, we need factor means, variances, and the estimate of zeta. Assuming that all factors' unconditional returns are equal, unconditional expected return to factors equals $\mathbb{E}[\tilde{R}_t] = \frac{\mu}{1-\rho-\rho'}$. If factors are uncorrelated with similar volatilities ($\sigma^2_{R_f} = \sigma^2_R$), $\sigma^2_{R'} = \frac{1}{K} \sum_{f=1}^{N} \sigma^2_{R_f} = \frac{\sigma^2_R}{N}$, and factor variances become $\sigma^2_R = \frac{\sigma^2_\epsilon}{1-\rho^2-\frac{\rho'^2}{N}}$.

We compute the ratio of the variance in expected returns explained by the signal. Previously,

[11] This model can also be interpreted as a Spatial Autoregressive Model (SAR) in which a "neighbor" signal impacts future levels of the factor.

under AR(1) returns, this ratio was the squared autocorrelation coefficient. If we estimate the following regression for factor f,

$$\tilde{R_{f,t}} = a + b_1 \tilde{R}_{f,t-1} + b_2 \tilde{R}'_{t-1} + \epsilon_t, \tag{A-30}$$

then the ratio of $\frac{\text{cov}(\hat{b_1}\tilde{R}_{f,t-1}, \tilde{R_{f,t}})}{\text{var}(\tilde{R_{f,t}})}$ is the variation in factor returns explained by its past returns. Moreover, the ratio of $\frac{\text{cov}(\hat{b_2}\tilde{R}'_{t-1}, \tilde{R_{f,t}})}{\text{var}(\tilde{R_{f,t}})}$ is the variation in factor returns explained by the dominant factor,

$$R^2 = \frac{\text{cov}(\hat{b_1}\tilde{R}_{f,t-1}, \tilde{R_{f,t}}) + \text{cov}(\hat{b_2}\tilde{R}'_{t-1}, \tilde{R_{f,t}})}{\text{var}(\tilde{R_{f,t}})}. \tag{A-31}$$

When the signal uses last month's returns, $\hat{b_1} = \rho$ and $\hat{b_2} = \rho'$, and the R^2 of the regression becomes:

$$R^2 = \frac{\rho \text{cov}(\tilde{R}_{f,t-1}, \tilde{R_{f,t}}) + \rho' \text{cov}(\tilde{R}'_{t-1}, \tilde{R_{f,t}})}{\sigma_R^2} = \frac{\rho^2 \sigma_R^2 + {\rho'}^2 \sigma_R \sigma_{R'}}{\sigma_R^2}, \tag{A-32}$$

where $\sigma_{R'}$ is the standard deviation of the dominant factor represented by the average return to all uncorrelated factors with similar standard deviations (σ_R). Since we do not directly observe the dominant factor, in our empirical analysis, we compute the dominant factor as the average return of all factors *excluding* the factor that is being modeled. If the number of factors used to calculate the dominant factor is $N-1$, we can further simplify R^2,

$$R^2 = \frac{\rho^2 \sigma_R^2 + {\rho'}^2 \frac{\sigma_R^2}{\sqrt{N-1}}}{\sigma_R^2} = \rho^2 + \frac{{\rho'}^2}{\sqrt{N-1}}. \tag{A-33}$$

In the five-factor model, for example, $N-1$ is always four. It is reasonable to assume that using more factors in computing the dominant factor creates a cleaner proxy, increasing the correlation between the dominant factor and future individual factor returns—increasing N increases ρ'. Also given that we assumed factors are uncorrelated, the dominant factor's volatility is decreasing in N, which is reflected in the apparently negative relationship between R^2 and number of factors used in calculating the dominant factor. These two effects—the increase in ρ'^2 and the decrease in the dominant factor's volatility as N increases—may offset each other so that the regression's R^2 is insensitive to the choice of the number of factors used to estimate the dominant factor. We thus assume that this term is constant regardless of the number of the factors used; we conjecture that there exists a ρ'' that satisfies

$$\rho'(N)^2 = \sqrt{N-1}\rho''^2. \tag{A-34}$$

This assumption means that R^2 of the regression of an individual factor return on the lagged dominant factor is always constant regardless of the choice of the factors used in creating the dominant factor. We have

$$R^2 = \rho^2 + \rho''^2. \tag{A-35}$$

According to equation (A-35), if returns follow the process in equation (A-30), adding the dominant factor to the regression increases the R^2 by an amount equal to the squared cross-serial correlation between the dominant factor and the individual factor. With the new signal that contains two components, $\mu(S_t) = \mu + \rho\tilde{R}_{f,t-1} + \rho''\tilde{R}'_{t-1}$, we have

$$\sigma_\epsilon^2 = \left(\frac{1}{\rho^2 + \rho''^2} - 1\right)\sigma^2_{\mu(\tilde{S})}. \tag{A-36}$$

We can now compute ζ:

$$\zeta = \mathbb{E}\left[\frac{\mu^2(\tilde{S})}{\mu^2(\tilde{S}) + \sigma_\epsilon^2}\right] = \frac{SR^2 + \rho^2 + \rho''^2}{SR^2 + 1}. \tag{A-37}$$

where derivation is the same as before. The expected ratio of the efficient factor's Sharpe ratio to the original factor's Sharpe ratio becomes

$$\frac{SR^*}{SR} = \sqrt{\frac{1 + \left(\frac{\rho + \rho''}{SR}\right)^2}{1 - \rho^2 - \rho''^2}}. \tag{A-38}$$

What is a plausible value for ρ''^2? In the data, when the signals use last month's returns, the estimates for this parameter range from 0.1% to 0.5% depending on the choice of factors; ρ'', therefore, typically lies between 0.03 and 0.07. An exception is the market factor, which always negatively relates to past and future returns of all other factors.

Using equation (A-37), the investor's optimal weight on the risky asset in this new setting is

$$x(S_t) = \frac{\mu_p}{\zeta}\left(\frac{\mu(S_t)}{\mu(S_t)^2 + \sigma_\epsilon^2}\right), \tag{A-39}$$

where the constant ζ and the conditional expected return are

$$\zeta = \frac{SR^2 + \rho^2 + \rho''^2}{SR^2 + 1}, \tag{A-40}$$

$$\mu(S_t) = \mu_p (1 - \rho) + \rho \tilde{R}_{f,t-1} + \rho'' \tilde{R}'_{t-1}. \tag{A-41}$$

The only new required parameter for estimating the weight under this return process is ρ''. Similar to the results in Section 3.3, an investor does not need to know the exact parameter value to use the information embedded in the dominant factor's past returns; an investor who considers a positive ρ'' that is not too large can improve the typical factor's Sharpe ratio. In our implementation, for simplicity, we assume that $\rho'' = \frac{\rho}{2}$.