Journal of Financial Econometrics, 2023, Vol. 21, No. 4, 1228–1257
https://doi.org/10.1093/jjfinec/nbac002
Advance Access Publication Date: 4 February 2022
Article

OXFORD

The Determinants of Volatility Timing Performance

Nick Taylor

School of Accounting and Finance, University of Bristol, Bristol, BS8 1PQ, UK

Address correspondence to Nick Taylor, School of Accounting and Finance, University of Bristol, Bristol, BS8 1PQ, UK, or e-mail: nick.taylor@bristol.ac.uk.

Received March 11, 2021; revised January 16, 2022; editorial decision January 18, 2022

Abstract

The exact conditions under which volatility timing strategies yield value are documented. These conditions include: the ability to correctly forecast next period stochastic variance, and violation of a strict version of Merton's intertemporal capital asset pricing model (ICAPM). While the empirical evidence supports the first of these conditions, the latter remains open to debate. Our empirical results confirm the former, but demonstrate a significant violation of the (strict) ICAPM. It follows that volatility timing strategies appear to have value. However, using reasonable parameter values plugged into the derived formulae, the results also show that extreme leverage is often required for success. A method of tempering leverage is proposed, which is somewhat able to loosen the requirement of high leverage while still maintaining a good performance level. Given the likely variation in (strict) ICAPM violations across time and assets, it follows that volatility timing success (or failure) is very much sample dependent.

Key words: financial forecasting, forecasting skill, performance fees, risk preference, volatility timing

JEL classification: C22, C58, G11, G17

The recent volatility observed in financial markets serves as a reminder of the enormous risks associated with trading financial assets.[1] This uncertainty has, unsurprisingly, led to renewed interest in trading strategies that seek to take positions in order to mitigate the adverse effects of such events (see, e.g., Smith, 2020, for cautionary advice). To demonstrate the association between volatility and interest in such market timing strategies consider the anecdotal evidence in Figure 1. This contains a plot of the number of worldwide Google

1 During the first wave of the COVID-19 pandemic, the spike in volatility was greater than that observed in the 2008–2009 global financial crises. The VIX index stood at 85.47 on March 18, 2020. This compares to a high of 81.48 observed on November 20, 2008 (data downloaded from http://www.cboe.com/products/vix-index-volatility/vix-options-and-futures/vix-index/vix-historical-data).

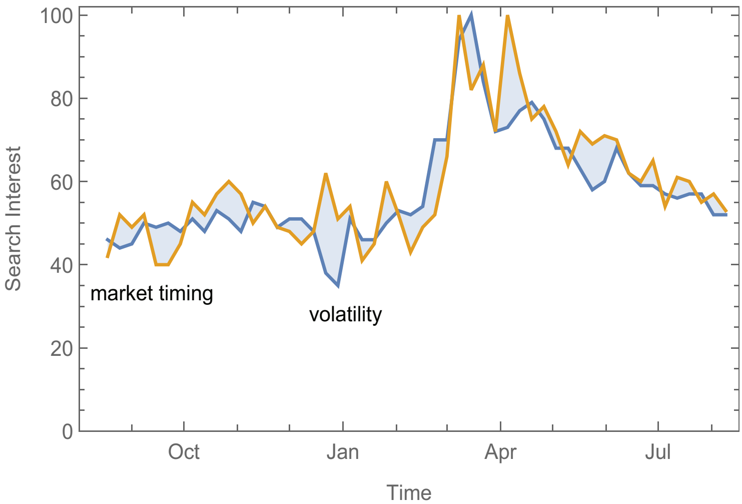


Figure 1 Volatility and market timing Google searches. This figure contains a plot of the number of worldwide Google searches of the terms "volatility" (blue line) and "market timing" (orange line) over the period August 18, 2019 to August 9, 2020. The data frequency is weekly, and each series is standardized to have a maximum value of 100.

searches of the terms "volatility" and "market timing" over the period August 18, 2019 to August 9, 2020 (the data frequency is weekly, each series is standardized to have a maximum value of 100, and are obtained from Google Trends). There is peak interest around early March 2020, and the two series are clearly positively related. With this observation in mind, we analyze the performance of volatility timing strategies within a theoretical setting. In doing this, we are able to distil the factors that determine this performance.

Market timing strategies have been extensively studied by the academic community. Merton (1981) provides a seminal attempt to theoretically determine the value of forecasting skill in the context of asset returns (see, e.g., Sharpe, 1975, for an early empirical examination). Using a simple binary market timing strategy, Merton is able to demonstrate that the necessary and sufficient conditions for success are the conditional probabilities of correctly forecasting when the risky asset return will exceed the safe asset return (conditional on the market return). This result led directly to testing criteria for market timing (Henriksson and Merton, 1981; Breen et al., 1986;), which together with the Treynor and Mazuy (1966) measure have been extensively applied within a variety of contexts; for recent applications, see Brandt et al. (2019), Ding et al. (2020), Merkoulova (2020), and Mascio et al. (2021), for hedge funds, mutual funds, commodity futures, and equity indices, respectively.

Investigations of market timing have provided scant evidence of success. For instance, within the context of U.S. equity funds, Chang and Lewellen (1984), Henriksson (1984), Grinblatt and Titman (1988), Becker et al. (1999), and Jiang (2003) all find that market timing actually has negative value.[2] The problem inherent in market timing is evinced by

2 See Chance and Hemler (2001), Bollen and Busse (2005), and Mamaysky et al. (2008) for contrary evidence based on daily data.

the difficulty in predicting market returns via a set of publicly available predictors; see Welch and Goyal (2008). It is therefore not surprising that alternative forms of market timing have emerged. Most notably a large literature considers the virtues of trading strategies that seek to change their exposure to the risky asset(s) in response to changes in volatility forecasts. For studies that seek to maximize utility by altering portfolio weights according to volatility forecasts (henceforth *utility maximizing* strategies) see, for example, Fleming et al. (2001, 2003), Marquering and Verbeek (2004), Chiriac and Voev (2011), and Taylor (2014). Moreover, a related literature exists which alters the exposure to the risky asset(s) such that the volatility is spread evenly over time by applying weights that are inversely proportional to the time-varying risk associated with the underlying asset (so-called *volatility-managed* (VM), *dynamic risk parity*, or *volatility targeting* strategies); see, for example, Maillard et al. (2010), Asness et al. (2012), Dopfel and Ramkumar (2013), Hocquard et al. (2013), Barroso and Santa-Clara (2015), Harvey et al. (2018), Daniel and Moskowitz (2016), Hoyle and Shephard (2019), Liu et al. (2019), Moreira and Muir (2017, 2019), Bongaerts et al. (2020), Cederburg et al. (2020), and Barroso and Detzel (2021).[3]

We complement these literatures by considering the performance of a volatility timing strategy with respect to the forecasting skill and risk preferences (RPs) of the user within a pure theoretical setting. Two utility-maximizing volatility timing strategies are considered. One is based on maximizing the *conditional* expectation of benefits, while the benchmark strategy is based on maximizing the *unconditional* expectation of benefits, both within a mean–variance (MV) setting. Consequently, the advantage to the latter over the former represents a measure of the value of conditioning information. By using a theoretical setting, we are able to isolate the determinants of this value, and thus enable volatility timers to assess the virtues of their strategy given *market conditions* (as indicated by the prevailing risk–return relationship), and *forecasting skill* (as indicated by their ability to predict next period stochastic variance). In doing this, we are able to provide the specific conditions under which volatility forecasts have economic significance.

The relative success of volatility timing strategies (cf. market timing strategies based on return prediction) relies on two features. First, as volatility is highly persistent, it is relatively easy for users to take positions in the current period that are complimentary to prevailing conditions in the next period. Second, the nature of the risk–return relationship must be such that returns are not exclusively determined by *ex ante* stochastic variance. Instead returns must be determined (at least in part) by their covariance with other state variables. Equivalently, a *strict version* of Merton's ICAPM (in which a time-invariant investment opportunity set is assumed) must fail to hold. It is the specifics of these features that we tease out in this article. Under a small and simple set of assumptions (that find support empirically), we are able to document the exact conditions for volatility timing strategy success not only in terms of forecasting skill, but also in terms of the risk–return relationship and RPs of users.

The theoretical results of this article indicate that the key driver of volatility timing success is that the market-determined risk–return tradeoff must be such that a 'risk' premium exists in the absence of market risk. Under this condition, if the forecaster has skill (i.e., they are able to predict volatility), then the conditional MV efficient strategy earns a higher

3 A number of indices are available that seek to maintain a given level of risk; see Taylor (2017) for empirical analysis of these *risk control* indices.

risk-adjusted return than the unconditional MV efficient strategy. The magnitudes of these benefits are calculated by calibrating to daily U.S. data over the period 1927–2019. Under realistic forecasting skill levels, the maximum expected benefit can be as high as 2–3% per annum (per unit of leverage); however, this does require an extreme market-determined risk–return tradeoff and/or high leverage levels.

The results also reconcile the performance of the VM strategy with the conditional MV efficient strategy. In particular, we demonstrate that the former is a restricted version of the latter, and relies on a particular risk–return relationship. If the latter assumption fails to hold, then the conditional MV efficient strategy offers substantial improvements over the VM strategy. This superiority is a function of the degree of risk–return misspecification, and forecasting skill. Consequently, it is not clear why the VM strategy should be pursued instead of the conditional MV efficient strategy.

The rest of the article is organized as follows. Section 1 contains the main theoretical findings of the article, with the corresponding empirical findings provided in Section 2. The final section concludes.

1 Theory

The development of the main result consists of two stages. First, the trading environment is outlined in which a trading rule is considered in the context of the dynamics of returns and a function of these returns. The second part contains a series of results based on the assumed trading environment and return/benefit dynamics.

1.1 The Environment

Various features of a trading rule are considered under specific assumptions regarding the dynamics of returns and the benefits of trading. These dynamics are specified below.

Assumption 1 (Return dynamics). Excess returns are determined by an *ex post* version of Merton's (1973) ICAPM such that they vary positively with their conditional variance; specifically,

$$R_t = \lambda_0 + \lambda_1 \mathrm{E}[\sigma_t^2 | \mathcal{F}_{t-1}] + \epsilon_t, \quad t = 1, 2, \ldots, T, \tag{1}$$

where R_t is the difference between the return to a diversified risky asset portfolio and a safe asset, λ_0 captures the compensation for covariance of the market return with other state variables (other than *ex ante* variance), $\lambda_1 (> 0)$ is the coefficient of relative risk aversion of the representative agent, ϵ_t is a zero-mean error term with variance during t given by σ_t^2, and $\mathrm{E}[\sigma_t^2 | \mathcal{F}_{t-1}]$ is the consensus conditional expectation of σ_t^2 given all information available at $t-1$. Under the strict set of the ICAPM assumptions (that includes a time-invariant investment opportunity set), λ_0 should equal zero. Here σ_t^2 is the latent stochastic variance process that contains a predictable component such that

$$\ln[\sigma_t^2] = \mathrm{E}[\ln[\sigma_t^2] | \mathcal{F}_{t-1}] + \nu_t, \tag{2}$$

where $\mathrm{E}[\ln[\sigma_t^2] | \mathcal{F}_{t-1}] \sim \mathrm{N}(A, C^2)$ and $\nu_t \sim \mathrm{IN}(0, D^2)$. It follows that $\ln[\sigma_t^2] \sim \mathrm{N}(A, B^2)$, where $B^2 = C^2 + D^2$. The correlation between the volatility forecast and its realization, ρ, measures the quality of the conditional expectation such that $\rho = C/B$, and can be interpreted as forecasting skill. It follows that $C^2 = \rho^2 B^2$ and $D^2 = (1 - \rho^2)B^2$.

The parameter A measures that average level of (log) variance (realized and in expectation). Moreover, B can be considered a measure of volatility of volatility (*vol-of-vol* henceforth), with C and D representing the expected and unexpected vol-of-vol. Adding (or removing) white noise to the stochastic variance process (i.e., changing B) will not affect the expected vol-of-vol (i.e., C), but will simply increase (or decrease) the unexpected vol-of-vol (i.e., D).

The log transformation is used in Equation (2) because stochastic variance (σ_t^2) is likely to have a skewed (and hence non-Gaussian) distribution. This transformation has been previously used in the extant literature to good effect. For instance, Andersen et al. (2003) demonstrate that the use of log transformations in the context of realized volatility enables use of standard linear Gaussian approaches for modeling and forecasting.

Assumption 2 (Realized benefits). The *realized benefit* derived from consumption of excess returns is measured by the function $f[R_t, w_t, \theta]$, where w_t is the leverage measure given by the proportion of wealth invested (which, in turn, can be dependent on other parameters), and θ measures the user's RP. Specifically, this benefit is captured by the first two moments of returns such that it is given by

$$Y_{a,t} \equiv f[R_t, w_{a,t}, \theta] = w_{a,t}R_t - w_{a,t}^2\theta(R_t - \mathrm{E}[R_t|\mathcal{F}_{t-1}])^2, \quad \theta > 0, \tag{3}$$

where $w_{a,t}$ is the leverage measure associated with an active trading strategy, and risk itself is captured by the squared prediction error.

It follows that the conditional expectation of the benefits is given by

$$Y_{a,t}^e \equiv \mathrm{E}[f[R_t, w_{a,t}, \theta]|\mathcal{F}_{t-1}] = w_{a,t}\mathrm{E}[R_t|\mathcal{F}_{t-1}] - w_{a,t}^2\theta\mathrm{Var}[R_t|\mathcal{F}_{t-1}]. \tag{4}$$

The expression for the conditional expectation of returns follows directly from Equation (1). The conditional variance expression can be derived as follows. First, note that Equation (2) can be written as

$$\sigma_t^2 = \exp[\nu_t]X_t, \tag{5a}$$

where $X_t \equiv \exp[\mathrm{E}[\ln[\sigma_t^2]|\mathcal{F}_{t-1}]] \sim \mathrm{LN}(A, C^2)$. Taking conditional expectations we obtain the following expression for the conditional variance of returns:

$$\mathrm{Var}[R_t|\mathcal{F}_{t-1}] = \mathrm{E}[\sigma_t^2|\mathcal{F}_{t-1}] = \mathrm{E}[\exp[\nu_t]]X_t = \exp[D^2/2]X_t. \tag{5b}$$

Here as ν_t is assumed to be independently distributed it follows that $\mathrm{E}[\exp[\nu_t]|\mathcal{F}_{t-1}] = \mathrm{E}[\exp[\nu_t]]$.

Assumption 3 (The trading rule). An active trading rule is considered in which a position is taken each period such that the unconditional expectation of realized benefits is maximized. This rule is based on the above return dynamics. As the dynamics are driven by predictability in the stochastic volatility process, the rule is henceforth referred to as the *volatility timing strategy*. Moreover, as the realized benefits are based on the first two moments of returns, we henceforth refer to volatility strategies based on these moments as *MV volatility strategies* (or *MV strategies* for short).

The amount of leverage, $w_{a,t}$, required in the active position such that it maximizes the *conditional* expectation of next period benefit (and in so doing maximizes the unconditional expectation of the realized benefits) is given by

$$w^*_{a,t} = \underset{w_{a,t}}{\operatorname{argmax}} \, \mathrm{E}[f[R_t, w_{a,t}, \theta]|\mathcal{F}_{t-1}] = \frac{\mathrm{E}[R_t|\mathcal{F}_{t-1}]}{2\theta \mathrm{Var}[R_t|\mathcal{F}_{t-1}]} = \frac{\lambda_0 + \lambda_1 \exp[D^2/2]X_t}{2\theta \exp[D^2/2]X_t}. \tag{6}$$

In contrast, we also consider a benchmark buy-and-hold strategy that consists of fixed leverage (denoted w_b, with realized benefits given by $Y_{b,t}$) designed to maximize the *unconditional* expectation of benefits—it follows that $w_b = \mu/2\theta\sigma^2$, where μ $(= \lambda_0 + \lambda_1\sigma^2)$ is the unconditional mean return and σ^2 $(= \exp[A + B^2/2])$ is the unconditional variance of errors in Equation (1).[4] We henceforth refer to strategies based on these leverage levels as the *conditional MV strategy* and the *unconditional MV strategy*, respectively.

Finally, we note that if the strict version of the ICAPM holds (i.e., $\lambda_0 = 0$), it follows from Equation (6) that $w^*_{a,t} = \lambda_1/2\theta$. Noting that $\mu = \lambda_0 + \lambda_1\sigma^2$ gives $w^*_{a,t} = \mu/2\theta\sigma^2$. Thus, in this situation, the conditional MV leverage will coincide with the unconditional MV leverage, and the conditional and unconditional MV strategies are equivalent.

1.2 Strategy Performance

Within the context of the above trading strategies, two performance measures are considered: the *advantage* and the *performance fee set*. The first of these is formally defined as follows.

Definition 1 (The advantage). The advantage (gain) achieved by adopting the active strategy is given by

$$\mathcal{G} \equiv \mathrm{E}[Y_{a,t} - Y_{b,t}], \tag{7}$$

which represents the difference between the expected benefits to the active and benchmark strategies.

Under Assumptions 1–3, and the advantage definition, we have the following proposition.

Proposition 1. The advantage to the conditional MV strategy over the unconditional MV strategy (equivalently the value of conditioning information) is given by

$$\mathcal{G} = \frac{(\exp[C^2] - 1)\lambda_0^2}{4\theta\sigma^2}, \tag{8}$$

which is positive if $\lambda_0 > 0$ and $\exp[C^2] > 1$.

Proof. See Appendix A. □

4 Expressions for μ and σ^2 are obtained by taking unconditional expectations of Equation (1) and using Equation (2), respectively. Moreover, the unconditional variance of returns is given by $\sigma_R^2 = (\exp[C^2] - 1)(\mu - \lambda_0)^2 + \sigma^2$.

From this expression, we note that the advantage is given by the product of three separate components: the (inverse) RP component ($1/\theta$); the market conditions component ($\lambda_0^2/4\sigma^2$); and the forecaster skill component ($\exp[C^2]-1$). The RP parameter is inversely related to the advantage, while the other components have a positive impact on advantage. Other more subtle features are also evident. First, the advantage is invariant to changing the amount of noise in the stochastic variance measure. This can be seen by noting that the advantage is a function of C, which is itself invariant to such noise. Any increase in the stochastic variance noise would increase the vol-of-vol level (B) and reduce the apparent skill level (ρ) such that $C=\rho B$ remains constant. Second, conditional on the user having forecasting skill (i.e., $\exp[C^2]>1$), a statistical test of the null hypothesis of no advantage can be conducted via a statistical test of the null hypothesis that $\lambda_0=0$. This null coincides with the strict version of the ICAPM being true.

The above result has been derived under the assumption that the log of stochastic variance, and the associated conditional expectation and error, each have a Gaussian distribution (see Equation (2) in Assumption 1). This assumption is necessary in the context of the conditional expectation. However, it is not necessary to assume that the error (and hence the log of stochastic variance) has this distribution. Indeed, it is possible to show that the error (ν_t) can have any distribution as long as it is independently distributed and the unconditional expectation $\exp[\nu_t]$ exists. Further details are provided in Appendix B.

In assessing the practicality of the conditional MV strategy, it is useful to gage the magnitudes of the leverage, $w_{a,t}$, in this strategy. To this end, consider the following proposition.

Proposition 2. Conditional MV leverage has a probability density function given by

$$g[w_a^*]=\frac{1}{\sqrt{2\pi}}\frac{\kappa_2}{(w_a^*-\kappa_1)^2}\frac{\exp\left[-\left(\frac{-A+\log[\kappa_2/(w_a^*-\kappa_1)]}{\sqrt{2}C}\right)^2\right]}{C\kappa_2/(w_a^*-\kappa_1)},\qquad \frac{\kappa_2}{w_a^*-\kappa_1}>0, \tag{9}$$

and zero otherwise. Here $\kappa_1=\lambda_1/2\theta$ and $\kappa_2=\lambda_0/2\theta\exp[D^2/2]$.

Proof. See Appendix A. □

The above density can be used to calculate the leverage moments. For instance, the MV are given by

$$\mathrm{E}[w_a^*]=\frac{\mu+\lambda_0(\exp[C^2]-1)}{2\theta\sigma^2}, \tag{10a}$$

$$\mathrm{Var}[w_a^*]=\frac{\lambda_0^2\exp[2C^2](\exp[C^2]-1)}{4\theta\sigma^4}. \tag{10b}$$

The mean leverage is necessarily greater than the benchmark leverage ($w_b^*=\mu/2\theta\sigma^2$) providing the user has forecasting skill and the strict version of the ICAPM does not hold (i.e., $C>0$ and $\lambda_0>0$). The higher the forecasting skill level the greater the expected leverage. If there is no forecasting skill, and/or the strict version of the ICAPM holds, then the conditional MV leverage will coincide with the (constant) unconditional MV leverage.

An alternative measure of performance is also considered, in which an economic value to the above advantage is provided. This is given in the following definition.

Definition 2 (The performance fee set). The performance fee set (per unit of leverage) contains the fees the user of the active strategy is willing to pay in order to enjoy the advantage over the benchmark strategy. Specifically, this set is given by $\{\delta \in \mathbb{R} | \Phi[\delta]\}$ with

$$\Phi[\delta] = \mathrm{E}[(f[R_t - \delta, w_{a,t}, \theta] - Y_{b,t})] \geq 0. \tag{11}$$

Here the performance fee set is standardized by leverage. To obtain the total fee, we would multiply by leverage ($w_{a,t}$).

Proposition 3. The maximum performance fee (per unit of leverage) a user is prepared to pay to use the conditional MV strategy over the unconditional MV strategy is given by

$$\delta^* = \frac{\lambda_0^2}{2\left(\lambda_0 + \frac{\mu}{\exp[C^2]-1}\right)}, \tag{12}$$

which is positive if $\lambda_0 > 0$ *and* $\exp[C^2] > 1$.

Proof. See Appendix A. □

The performance fee is standardized by leverage and is thus invariant to RPs (evinced by the absence of the θ parameter). A positive performance fee does, however, require that the strict version of the ICAPM fails to hold (i.e., $\lambda_0 > 0$) and that forecasting skill exists ($\exp[C^2] > 1$)—the same requirements as for the advantage in Proposition 1. Moreover, increases in these parameters lead to increases in the performance fee. We can go further and decompose the advantage such that it is possible to see how it is related to the above performance fees and leverage levels. This decomposition is given in the following proposition.

Proposition 4. The conditional MV advantage can be decomposed into two components: the leverage component and the performance fee component. Specifically,

$$\mathcal{G} = \mathrm{E}[w_a^*]\delta^*, \tag{13}$$

where expressions for $\mathrm{E}[w_a^*]$ and δ^* are provided in Propositions 2 and 3.

Proof. See Appendix A. □

Thus the advantage is the product of the (maximum) performance fee and the expected leverage. RPs determine the advantage via their impact on leverage only; from Equation (10a) we see that the reciprocal of the RP parameter determines expected leverage (ceteris paribus). More risk-averse investors enjoy lower advantages because they are less willing to take higher leverage levels. This also means that less risk-averse investors may be willing to take extremely large positions. To temper the use of excessively high leverage rates (and/or use of highly volatility leverage rates), it is possible to shrink leverage toward the benchmark leverage. Formally,

Definition 3 (Leverage shrinkage). Conditional MV leverage is shrunk toward unconditional MV leverage such that

$$\tilde{w}_{a,t} = \beta w^*_{a,t} + (1-\beta) w^*_b \tag{14}$$

where β determines the degree of shrinkage.

This definition enables a derivation of the leverage shrinkage distribution.

Proposition 5. Conditional MV leverage shrinkage has a probability density function given by

$$g[\tilde{w}^*_a] = \frac{1}{\sqrt{2\pi}} \frac{\kappa_4}{(\tilde{w}^*_a - \kappa_3)^2} \frac{\exp\left[-\left(\frac{-A+\log[\kappa_4/(\tilde{w}^*_a-\kappa_3)]}{\sqrt{2}C}\right)^2\right]}{C\kappa_4/(\tilde{w}^*_a - \kappa_3)}, \qquad \frac{\kappa_4}{\tilde{w}^*_a - \kappa_3} > 0, \tag{15}$$

and zero otherwise. Here $\kappa_3 = \beta\kappa_1 + (1-\beta)w_b$ and $\kappa_4 = \beta\kappa_2$.

Proof. See Appendix A. □

The use of shrinkage leverage leads to lower levels of leverage. To see note that the MV leverage levels are now given by

$$\mathrm{E}[\tilde{w}_a] = \frac{\mu + \beta\lambda_0(\exp[C^2]-1)}{2\theta\sigma^2} \leq \mathrm{E}[w_a], \tag{16a}$$

$$\mathrm{Var}[\tilde{w}_a] = \frac{\beta^2\lambda_0^2\exp[2C^2](\exp[C^2]-1)}{4\theta\sigma^4} \leq \mathrm{Var}[w_a], \tag{16b}$$

where increases in the degree of shrinkage (i.e., as β falls to zero) lead to lower levels of the MV of leverage. These lower leverage levels necessarily imply lower transaction costs. However, they limit the size of the maximum performance fee. This can be seen in the following proposition.

Proposition 6. The maximum performance fee (per unit of leverage) for a user of leverage shrinkage is given by

$$\delta^* = \frac{(2-\beta)\beta\lambda_0^2}{2\left(\beta\lambda_0 + \frac{\mu}{\exp[C^2]-1}\right)}, \tag{17}$$

which is non-negative under the restriction that $\lambda_0 \geq 0$ and $0 \leq \beta \leq 2$. The product of this fee and the expected leverage associated with the conditional MV strategy with shrinkage gives the advantage.

Proof. See Appendix A. □

Interestingly, there is a nonlinear relationship between the maximum performance fee and the degree of shrinkage. Indeed, there is an optimal level of shrinkage such that this fee is maximized.

Proposition 7. The optimal degree of leverage shrinkage is given by

$$\beta^* = \frac{2\mu}{\mu + \sqrt{\mu^2 + 2\mu\lambda_0(\exp[C^2] - 1)}}, \tag{18}$$

where previous notation is maintained.

Proof. See Appendix A. □

This finding requires some discussion. We can see from Proposition 4 that the performance fee is given by the advantage divided by the expected leverage. If we interpret the advantage as an output and the expected leverage as an input, then this ratio can be considered as a kind of productivity rate. As such, the optimal degree of leverage shrinkage delivers the maximum productivity rate (alternatively, the maximum advantage per unit of leverage). It does not, however, deliver the maximum advantage as this is achieved by not shrinking leverage. The lower advantage obtained by shrinking leverage represents the price investors are willing to pay to avoid using high leverage. By maximizing the productivity rate, we are assuming a value function consisting of a benefit (the advantage) divided by a cost (the expected leverage). The latter implies an aversion to leverage (in turn, due to the presence of transaction costs) such that the investor is better off using shrinkage leverage. While other more complex value functions may be more appropriate, the productivity rate is virtuous in its simplicity.

The above findings pertain to the conditional MV strategy. However, we can extend the analysis by considering an alternative (related) volatility timing strategy. This is contained in the following section.

1.3 Volatility-Managed Portfolios

The VM strategy (henceforth the *conditional VM strategy*) is widely used in practice, and has been the subject of much academic inquiry. This strategy consists of a leverage level that is, proportional to the inverse of the conditional variance of returns and is scaled such that the returns to the strategy have an unconditional variance that equals the unconditional variance of (unscaled) returns. To marry this approach with the conditional MV strategy, it is useful to show what restrictions are required on the RP measure θ such that the VM leverage is achieved.

Definition 4 (The VM portfolio return). The VM portfolio (excess) return is constructed as follows:

$$R_{vm,t} = \frac{L}{\mathrm{Var}[R_t|\mathcal{F}_{t-1}]} R_t, \tag{19}$$

where L is selected such that $\mathrm{Var}[R_{vm,t}] = \mathrm{Var}[R_t]$. The VM literature assumes that $\mathrm{Var}[R_t|\mathcal{F}_{t-1}]$ is given by the realized variance in the previous period (in most cases, this is the previous trading day).[5]

Here we see that the implied VM leverage measure is $L/\mathrm{Var}[R_t|\mathcal{F}_{t-1}]$, which henceforth is denoted $w_{vm,t}$ (though noting that it is known at time $t-1$). The absence of a conditional mean in Equation (19) implies that the assumed VM data generating process underlying R_t is not dependent on the conditional variance. That is, λ_1 is assumed to equal zero such that $R_t = \lambda_0 + \epsilon_t$, with λ_0 assumed to equal μ.

Under the above definition, we have the following proposition.

Proposition 8. The RP implied by the VM approach is given by

$$\theta_{vm} = \frac{\mu \exp[C^2/2]\sqrt{\exp[C^2](\exp[C^2]-1)\lambda_0^2 + \sigma^2}}{2\sigma^2\sqrt{(\exp[C^2]-1)(\mu-\lambda_0)^2 + \sigma^2}}. \tag{20}$$

where previous notation is maintained.

Proof. See Appendix A. □

The conditional VM strategy is misspecified when λ_0 does not equal μ. To see the impact of this, it is possible to derive an expression for the maximum performance fee (per unit of leverage) an investor is willing to pay to use the conditional MV strategy over the conditional VM strategy, under this misspecification. This is contained in the following proposition.

Proposition 9. The maximum performance fee (per unit of leverage) a user is prepared to pay to use the conditional MV strategy over the conditional VM strategy is given by

$$\delta^* = \frac{1}{2} \times \frac{(\exp[C^2]-1)(\mu-\lambda_0)^2}{(\exp[C^2]-1)\lambda_0 + \mu}. \tag{21}$$

The product of this fee and the expected leverage associated with the conditional MV strategy (with θ_{vm} imposed) gives the advantage.

Proof. See Appendix A. □

We would expect the VM-implied RP level to increase as the forecasting skill level increases. This is because over this space the leverage levels become more volatile, which leads to VM returns having a higher unconditional variance. To temper this increase, the RP level must also increase over this space so that the unconditional variance of VM returns remains equal to the unconditional variance of unscaled returns. Rather than impose these restrictions, the conditional MV strategy is preferable on three counts: (i) all RPs are

5 Other, related, definitions have recently emerged in the literature. For instance, Bongaerts et al. (2020) consider a *conditional volatility targeting* strategy in which leverage is nonlinearity related to volatility, with variation driven by (predicted) extreme volatility episodes.

permitted; (ii) a more general DGP is specified based on Merton's ICAPM; and (iii) a range of forecasting skill levels are allowed beyond those implicitly assumed in the VM approach (i.e., those based on the previous realized variance level).

Having derived a set of theoretical results associated with volatility timing performance, we proceed with conduct a calibration exercise based on a long-span U.S. dataset.

2 EMPIRICAL ANALYSIS

To illustrate the magnitudes of the gains to volatility timing, and to investigate which particular users are likely to gain the most from such activity, we estimate the parameters in the above formulae.

2.1 Data

Excess returns to the risky asset portfolio are constructed using the U.S. value-weighted index (inclusive of dividends) provided in the CRSP database. These data were obtained via the WRDS interface. The risk-free rate is based on short-dated Treasury bills obtained via the Kenneth French data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). These data span the period from January 1, 1927 to December 31, 2019. In addition, we consider realized variance measures based on four major U.S. equity indices: the S&P 500, DJIA, NASDAQ 100, and Russell 2000 indices; and two international indices: the FTSE 100 and Nikkei 225 indices. A total of seven different measures of realized variance associated with each of these indices were collected from the Oxford-Man Institute of Quantitative Finance Realized Library (http://realized.oxford-man.ox.ac.uk/data). These measures are: realized variance (5-min sub-sampled), realized variance (10-min sub-sampled), realized kernel variance (non-flat Parzen), realized kernel variance (Tukey-Hanning), realized kernel variance (two-scale/Bartlett), realized semi-variance (5-min sub-sampled), and bi-power variation (5-min sub-sampled). See Shephard and Sheppard (2010) for further details of these measures. The data span the period from January 1, 2000 to December 31, 2019. All data are daily in frequency.

2.2 Parameter Estimation

The parameters identified in the previous section are estimated using daily data over the period 1927–2019. This sample period represents the largest span available such that parameter accuracy is maximized. While some of the parameters are trivial to calculate such as the unconditional MV of returns, ICAPM-based parameters need to be estimated with caution. With this in mind, we estimate the ICAPM (see Equation (1)) using daily frequency returns over the longest span of data available, that is, 1927–2019. Realized variance measures based on intraday data are not available over this period. Consequently, to be able to use such a long span of data, we use a crude log realized variance measure based on the log of squared daily returns when estimating the variance equation. A two-step estimation procedure is adopted in which the log realized variance equation is estimated, then the fitted value is transformed (and adjusted) into a variance forecast for use in the ICAPM (mean) equation. The former consists of various ARMA-type models each estimated by maximum likelihood, while the latter is estimated using ordinary least squares. See Meddahi (2003) for details of the relationship between the ARMA representation of realized variances, and the commonly employed GARCH representation of returns. The estimated parameters are provided in Table 1.[6]

Results associated with four ARMA-type models are presented: the AR(1), MA(1), ARMA(1,1), and ARMA(2,1) models. The fit of these models (as evinced by the forecasting skill level, ρ, which is the correlation between the forecast of log variance and its realization) varies across the models. The best model appears to be the ARMA(1,1) model and has an associated forecasting skill level of 34.3% (note that out-of-sample skills levels would be necessarily lower). Such forecasting skill levels are consistent with those found in other financial markets; see, for example, Andersen et al. (2003) in the context of foreign exchange markets. The ICAPM parameters associated with these models also exhibit variation over this space. For the ARMA(1,1) model, the ICAPM parameters λ_0 and λ_1 are 5.186 (annualized) and 0.225, both of which are significant at the 10% level. As the value of λ_0 plays a key role in determining the overall success of the conditional MV strategy (see Proposition 1), this result suggests that market conditions appears suitable in this regard. A positive and significant compensation for "other risk" (given by λ_0) is also found in previous studies; see, for example, Ghysels et al. (2005), who use the MIDAS technique.

To investigate the likely benefits of the conditional MV strategy, its advantages are calculated based on the above estimated parameter values. For simplicity the RP parameter is assumed to equal unity and an ARMA(1,1) model is used to model log-transformed stochastic volatility. In addition to use of the full sample parameter values, we also consider parameter values based on subsets of the sample. In particular, given the previously observed variation in volatility forecasting performance and the risk–return relationship over low and high volatility regimes, conditional MV strategy performance and leverage levels are documented in each of these regimes; see Taylor (2014) for evidence of the former and Lettau and Ludvigson (2004) for evidence of the latter.[7] The advantage is decomposed into three components: the (inverse) RP parameter ($1/\theta$); the skill component (given by $\exp[C^2] - 1$); and the market conditions component (given by $\lambda_0^2/4\sigma^2$), with the product of these components giving the advantage. Results are provided in Table 2.

Using the full sample parameter values it is evident that there is an advantage to using the conditional MV strategy over the unconditional MV strategy; the (annualized) advantage equals 1.626%. In order to achieve this level of performance, high leverage levels are required with the mean leverage level equal to 1.801. The results also show interesting variation in performance over the low and high volatility regimes. In the low volatility regime, a healthy advantage of 2.455% (per annum) is possible (with an associated mean leverage of 4.507). However, this advantage disappears in the high volatility regime, with only 0.080% (per annum) available. The cause of this collapse is variation in market conditions. In particular, over both volatility regimes the skill component remains strong (indeed it is

6 Other volatility models are available. For instance, one could augment a model to include the VIX as a predictor. In this article, the sample period used is such that VIX data are not available for the majority of the sample. Consequently, we refrain from considering the VIX as a predictor, but note that it could deliver higher forecasting skill levels than those presented in Table 1. This possibility motivates presentation of a wide range of forecasting skill levels in the subsequent analysis.

7 The volatility regimes are based on a log-transformed GARCH(1,1) model applied to returns with the regimes defined by periods in which this conditional variance is below and above the median level.

Table 1 Estimated model parameters

Parameter		Specification			
		AR(1)	MA(1)	ARMA(1,1)	ARMA(2,1)
Panel A: Mean equation parameters					
Other risk cost/compensation (λ_0)		−11.628**	−13.801*	5.186**	5.236**
		(5.860)	(7.142)	(2.268)	(2.267)
Rep. RRA coefficient (λ_1)		1.923***	2.157***	0.225*	0.220*
		(0.556)	(0.695)	(0.124)	(0.124)
Panel B: Volatility equation parameters					
AR(1) coefficient		0.146***		0.995***	0.991***
		(0.006)		(0.001)	(0.007)
AR(2) coefficient					0.004
					(0.007)
MA(1) coefficient			−0.118***	0.959***	0.958***
			(0.006)	(0.002)	(0.002)
Panel C: Hypothesis Test (p-values)					
$H_0 : \lambda_0 = 0$	$H_1 : \lambda_0 \neq 0$	0.047	0.053	0.022	0.021
Panel D: Other parameters					
Mean of returns (μ, annualized)		7.624			
Volatility of returns (σ, annualized)		18.017			
Mean of (log) volatility (A)		−11.039			
Volatility of (log) volatility (B)		2.536			
Volatility of expected (log) volatility (C)		0.371	0.296	0.870	0.870
Volatility of unexpected (log) volatility (D)		2.509	2.514	2.382	2.382
Forecasting skill ($\rho = C/B$)		0.146	0.117	0.343	0.343

Notes: This table contains the estimated ICAPM parameter values. The mean equation parameters (and standard errors) are given in Panel A, and the ARMA-based log-variance equation parameters (and standard errors) are given in Panel B. The p-values associated with likelihood ratio tests of the null hypotheses associated with the "other risk" compensation parameter (λ_0) are given in Panel C. Estimates of the other parameters used in the subsequent analysis are given in Panel D. The data used cover the period from January 1, 1927 to December 31, 2019.

Significance is indicated as follows: *** indicates significance at the 1% level; ** indicates significance at the 5% level; and * indicates significance at the 10% level.

actually higher in the high volatility regime). However, the market conditions component falls dramatically in the high volatility regime indicating a change in the risk–return relationship such that λ_0 is close to zero. These results underline the importance of market conditions on the performance of the conditional MV strategy.

2.3 Volatility Measures

The above analysis is based on a noisy measure of realized variance, that is, the squared return. The primary motivation for this choice is that these data can be observed over a very

Table 2 Performance and volatility regimes

Measure	Formula	Regime		
		Low Vol.	High Vol.	All
Panel A: Performance				
RP	$1/\theta$	1	1	1
Forecasting skill	$\exp[C^2]-1$	0.432	0.574	0.785
Market conditions	$\lambda_0^2/4\sigma^2$	5.678	0.139	2.071
Advantage (= RP × skill × conditions)	$(\exp[C^2]-1)\lambda_0^2/4\theta\sigma^2$	2.455	0.080	1.626
Panel B: Leverage				
Mean	$(\mu+\lambda_0(\exp[C^2]-1))/2\theta\sigma^2$	4.507	0.790	1.801
Variance	$(\lambda_0^2\exp[2C^2](\exp[C^2]-1))/4\theta\sigma^4$	3.082	0.036	2.279

Notes: This table contains the estimated performance and leverage measures for low and high volatility periods, and for all periods. The volatility regimes are based on a log-transformed GARCH(1,1) model applied to returns, with conditional volatility levels above (below) the median level defined as the high (low) volatility regime. In each regime, the ICAPM MV equations are estimated with the latter based on an log-transformed ARMA(1,1) variance equation. Performance measures (viz. RP, forecasting skill, market conditions, and advantage) are provided in Panel A and leverage measures (MV) are provided in Panel B. The data used cover the period from January 1, 1927 to December 31, 2019.

long period (93 years of daily observations yielding 24,493 observations). However, other more accurate realized variance measures (based on intraday data) are available that are less noisy in nature. Use of these measures would lead to a considerably lower vol-of-vol parameter estimate (i.e., B). However, two issues require further discussion. First, such measures are only available over much shorter time periods (typically from 2000 onwards only). This could lead to less accurate parameter estimates and less power associated with tests of the null hypotheses involving λ_0. Second, we know from Proposition 1 that the quality of the measure does not affect the performance of the strategy. This is because as vol-of-vol (B) decreases, there is an offsetting increase in forecasting skill ρ such that $C=\rho B$ remains broadly constant. To examine this issue in further detail, consider the results in Figure 2.

Here we have estimated a heterogeneous autoregressive model of (log) realized variance using seven daily frequency realized volatility measures based on intraday data, and the crude (log) realized variance measure based on squared returns.[8] All are estimated over the period 2000–2019, with six national and international indices used. The results show a clear negative relationship between forecasting skill and vol-of-vol. For the crude volatility measure, forecasting skills levels around 35% and a vol-of-vol value of 2.5 are achieved. By contrast, the sophisticated volatility measures yield respective values of approximately 85% and 1.[9] Even within this set of realized volatility measures there appears to exist a negative forecasting skill/vol-of-vol relationship. Thus the choice of the ρ and B parameter

8 See Corsi (2009) for more details of the heterogeneous autoregressive model.

9 These forecasting skills levels are broadly in line with those reported by Corsi (2009).

Figure 2 Forecasting skill and realized variance measure quality. This figure contains the forecasting skill level (ρ) and vol-of-vol parameter (B) values associated with the log-heterogeneous autoregressive model applied to various stock indices. The following realized variance measures are used: squared return denoted RV0; realized variance (5-min sub-sampled) denoted RV1; realized variance (10-min sub-sampled) denoted RV2; realized kernel variance (nonflat Parzen) denoted RV3; realized kernel variance (Tukey–Hanning) denoted RV4; realized kernel variance (two-scale/Bartlett) denoted RV5; realized semi-variance (5-min sub-sampled) denoted RV6; and bi-power variation (5-min sub-sampled) denoted RV7. The data span the period from January 1, 2000 to December 31, 2019, and are daily in frequency. (**a**) FTSE 100. (**b**) Nikkei 225. (**c**) S&P 500. (**d**) DJIA

values appears to be dependent on the quality of the realized volatility measure, while the product ρB (i.e., C) does not. As it is this product that appears in our derived formulae then it follows that the proposed approach is robust to variation in the quality of the realized variance measure.

2.4 The Distribution of Leverage

A prevailing issue for active trading strategies is that they require potentially excessive leverage. Thus, before examining the performance levels associated with the conditional and unconditional MV strategies, we examine their leverage requirements. To this end consider the results in Table 3. Here we present the probability that the conditional MV leverage level exceeds 1, 2, 4, 8, 16, 32, and 64. These probabilities are presented for a wide range of forecasting skill levels and different market conditions; specifically, we consider forecasting skill levels from zero to one (in increments of 0.1), and (strict) ICAPM violations of

Table 3 The distribution of leverage (conditional MV strategy)

Leverage	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
		Forecasting skill (ρ)									
Panel A: $\lambda_0 = \mu$											
1	0.500	0.550	0.599	0.646	0.691	0.734	0.773	0.809	0.841	0.870	0.894
2	0.000	0.004	0.128	0.291	0.423	0.528	0.613	0.684	0.743	0.793	0.835
4	0.000	0.000	0.006	0.070	0.188	0.314	0.431	0.533	0.621	0.695	0.757
8	0.000	0.000	0.000	0.008	0.057	0.150	0.262	0.377	0.484	0.580	0.662
16	0.000	0.000	0.000	0.000	0.012	0.056	0.136	0.239	0.350	0.457	0.556
32	0.000	0.000	0.000	0.000	0.002	0.016	0.059	0.134	0.232	0.339	0.446
64	0.000	0.000	0.000	0.000	0.000	0.003	0.022	0.067	0.140	0.235	0.340
Panel B: $\lambda_0 = 3/4 \times \mu$											
1	0.500	0.550	0.599	0.646	0.691	0.734	0.773	0.809	0.841	0.870	0.894
2	0.000	0.001	0.074	0.225	0.364	0.479	0.573	0.652	0.718	0.773	0.819
4	0.000	0.000	0.001	0.038	0.134	0.254	0.373	0.482	0.577	0.659	0.728
8	0.000	0.000	0.000	0.003	0.033	0.107	0.210	0.323	0.433	0.535	0.624
16	0.000	0.000	0.000	0.000	0.005	0.035	0.100	0.194	0.301	0.410	0.513
32	0.000	0.000	0.000	0.000	0.001	0.009	0.040	0.103	0.191	0.295	0.402
64	0.000	0.000	0.000	0.000	0.000	0.002	0.013	0.048	0.111	0.198	0.299
Panel C: $\lambda_0 = 1/2 \times \mu$											
1	0.500	0.550	0.599	0.646	0.691	0.734	0.773	0.809	0.841	0.870	0.894
2	0.000	0.000	0.026	0.138	0.275	0.400	0.507	0.598	0.674	0.738	0.791
4	0.000	0.000	0.000	0.013	0.074	0.176	0.292	0.406	0.511	0.603	0.681
8	0.000	0.000	0.000	0.001	0.014	0.062	0.146	0.251	0.362	0.469	0.566
16	0.000	0.000	0.000	0.000	0.002	0.017	0.062	0.138	0.237	0.344	0.451
32	0.000	0.000	0.000	0.000	0.000	0.004	0.022	0.068	0.142	0.237	0.342
64	0.000	0.000	0.000	0.000	0.000	0.001	0.007	0.029	0.077	0.152	0.246
Panel D: $\lambda_0 = 1/4 \times \mu$											
1	0.500	0.550	0.599	0.646	0.691	0.734	0.773	0.809	0.841	0.870	0.894
2	0.000	0.000	0.001	0.038	0.134	0.254	0.373	0.482	0.577	0.659	0.728
4	0.000	0.000	0.000	0.001	0.019	0.077	0.169	0.277	0.389	0.494	0.589
8	0.000	0.000	0.000	0.000	0.002	0.019	0.067	0.147	0.247	0.355	0.461
16	0.000	0.000	0.000	0.000	0.000	0.004	0.023	0.070	0.146	0.241	0.347
32	0.000	0.000	0.000	0.000	0.000	0.001	0.007	0.030	0.079	0.154	0.248
64	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.011	0.039	0.091	0.168

Notes: This table contains the probabilities that conditional MV leverage will exceed a particular level (1, 2, 4, 8, 16, 32, and 64). These are based on forecasting skill levels from zero to one (in increments of 0.1), and (strict) ICAPM violations of $\lambda_0 = \mu$, $\lambda_0 = 3\mu/4$, $\lambda_0 = \mu/2$, and $\lambda_0 = \mu/4$. The RP level is selected such that the unconditional MV leverage equals one (implying $\theta = \mu/2\sigma^2 = 1.17$).

$\lambda_0 = \mu$, $\lambda_0 = 3\mu/4$, $\lambda_0 = \mu/2$, and $\lambda_0 = \mu/4$. The RP level is selected such that unconditional MV leverage equals one (implying that $\theta = \mu/2\sigma^2 = 1.17$).[10]

The greater the skill level, the greater the range of leverage levels required. This range falls as the (strict) ICAPM violations fall. For instance, for a forecasting skill level of 0.4 and a (strict) ICAPM violation of $\lambda_0 = \mu$, leverage exceeds one 69.1% of the time, exceeds two 42.3% of the time, exceeds four 18.8% of the time, and exceeds eight 5.7% of the time. These are strong requirements, and would be beyond the borrowing capabilities of most investors. While the leverage requirements lessen as λ_0 falls, even at lower levels excessive leverage pervades. This observation motivates the shrinkage approach proposed in this article.

An appreciation of the leverage requirements with respect to RP is possible via Figure 3. The plots in this figure contain the expected leverage (panels (a) and (b)), and the probability of leverage being greater than one (panels (a) and (b))—henceforth *excessive leverage*. These results are based on a forecasting skill level, ρ, of 0.4; the "other risk" compensation parameter, λ_0, equal to μ and $\mu/2$; and RPs, θ, from zero to five.

The results indicate that high leverage is a undeniable feature of volatility timing strategies. Moreover, the conditional MV leverage levels are expected to be universally higher than the unconditional MV leverage levels. However, excessive leverage is not a universal feature of the conditional MV strategy; indeed, leverage requirements fall to reasonable levels as the RP parameters increases, and the extent of the (strict) ICAPM failure decreases (i.e., when λ_0 changes from μ to $\mu/2$).

A further illustration of the leverage requirements is given in Figure 4. Here the probability density functions are provided for a forecasting skill level, ρ, of 0.4; the "other risk" compensation parameter, λ_0, equal to μ, $3\mu/4$, $\mu/2$, and $\mu/2$; and an RP parameter, θ, such that the unconditional MV leverage level equals unity. In addition, we also plot the density functions when optimal shrinkage leverage is adopted. The formulae used to calculate these densities are provided in Propositions 2 and 5.

The densities are widely dispersed. However, this dispersion decreases as the degree of (strict) ICAPM violation decreases (i.e., when λ_0 becomes smaller). Moreover, when the proposed shrinkage approach is adopted, there is a noticeable reduction in leverage dispersion with a more peaked distribution apparent.

2.5 Leverage Constraints

The demands on leverage documented above are likely to lead to binding constraints in the presence of leverage restrictions. To investigate this issue we consider the relative efficiency of the advantage (as defined previously) to the conditional MV strategy (over the unconditional MV strategy) under leverage constraints with respect to the unrestricted leverage advantage. The formula is derived under the assumptions laid out in Section 1.1, but with leverage truncated at an upper level (denoted $\overline{w}$) when the optimal conditional MV leverage

10 Given the λ_0, λ_1, and ρ parameter variation that is likely to be observed empirically, we consider a range of λ_1 and ρ values, with the λ_0 variation achieved by the link $\lambda_0 = \mu - \lambda_1\sigma^2$. Moreover, we assume (annualized) MV of returns values (μ and σ_R, respectively) of 7.6% and 18%, respectively. Finally, the vol-of-vol parameter B will take a value of 2.5, with the mean of (log) volatility restricted to $A = \ln[\sigma^2] - B^2/2$. These parameter values are maintained in the subsequent analysis.

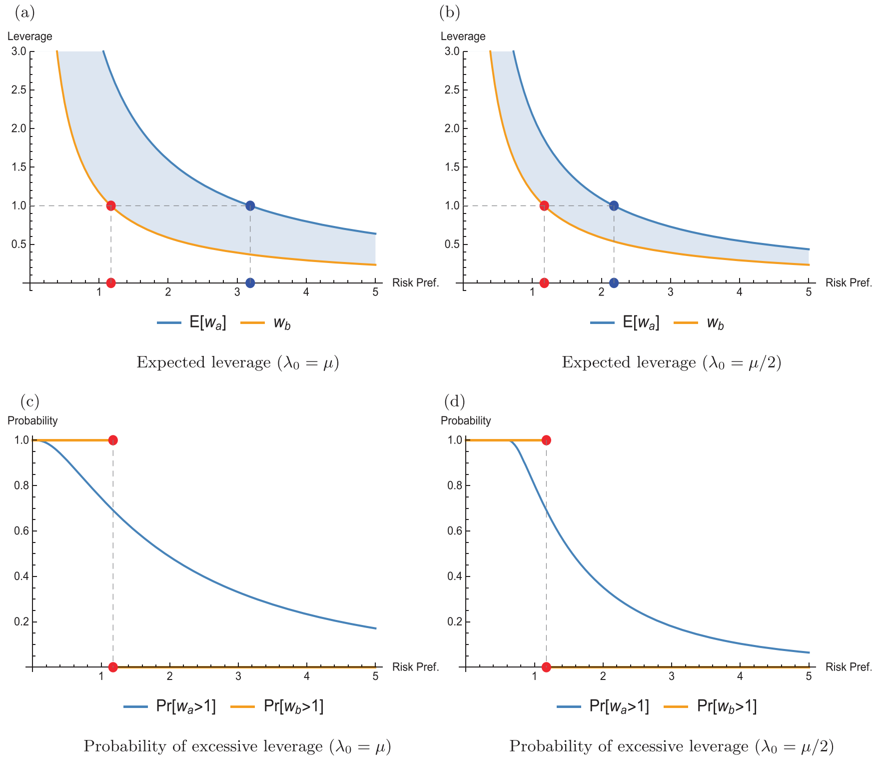


Figure 3 Leverage and RP. This figure contains the (**a**) Expected leverage ($\lambda_0 = \mu$). (**b**) Expected leverage ($\lambda_0 = \mu/2$). (**c**) Probability of excessive leverage ($\lambda_0 = \mu$). (**d**) Probability of excessive leverage ($\lambda_0 = \mu/2$). These results are based on a forecasting skill level, ρ, of 0.4; the 'other risk' compensation parameter, λ_0, equal to μ and $\mu/2$; and RPs, θ, from zero to five.

exceeds this level (a similar approach is adopted for the constrained unconditional MV leverage).[11] Figure 5 provides plots of these (percentage) relative efficiencies as a function of the (buy-and-hold) unconditional MV leverage, and are calculated using the same parameter values as used above. Unconstrained leverage ($\overline{w} = \infty$), a high margin constraint ($\overline{w} = 1$), the standard margin constraint ($\overline{w} = 1.5$), and a low margin constraint ($\overline{w} = 2$) are considered.

The (theoretical) profiles in this figure are similar in shape to the empirical profiles provided previously; see, for example, Figure 4 in Moreira and Muir (2017). Efficiency decreases as the unconditional MV leverage increases because the demands on leverage increase over this space. Moreover, as conditions become more favorable to volatility timing (i.e., as λ_0 increases), efficiencies decrease more quickly (again because of the increased leverage requirements). However, there is some robustness to leverage constraints, particularly when these constraints are moderate in nature. For instance, when $\lambda_0 = \mu/4$ and the unconditional MV leverage equals unity, an efficiency level above 60% is attained in the

11 The *Mathematica* notebook used to derive this formula is available on request.

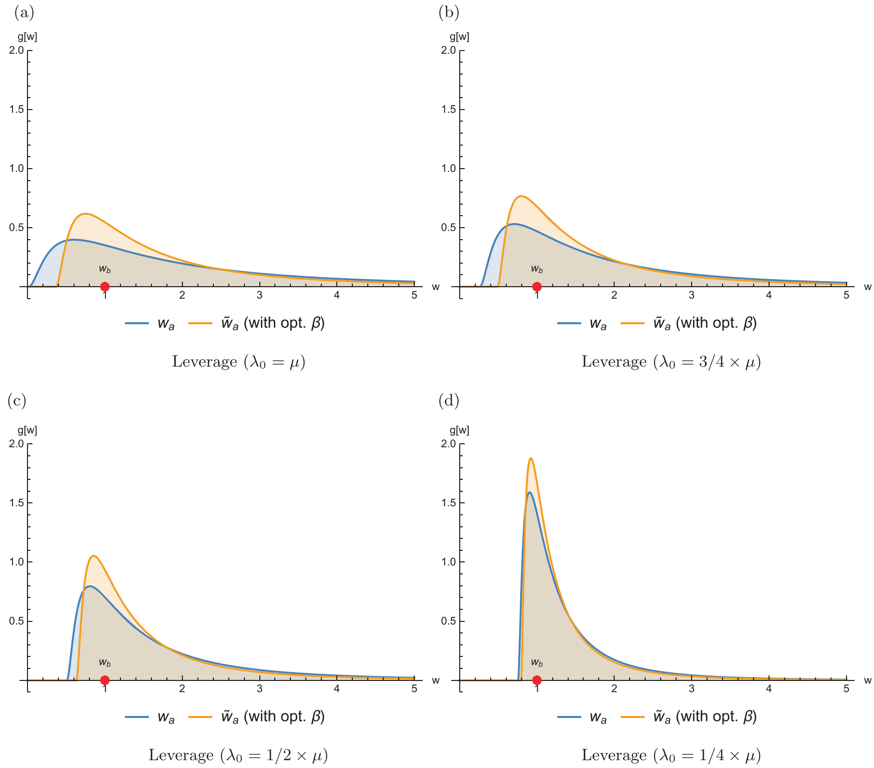


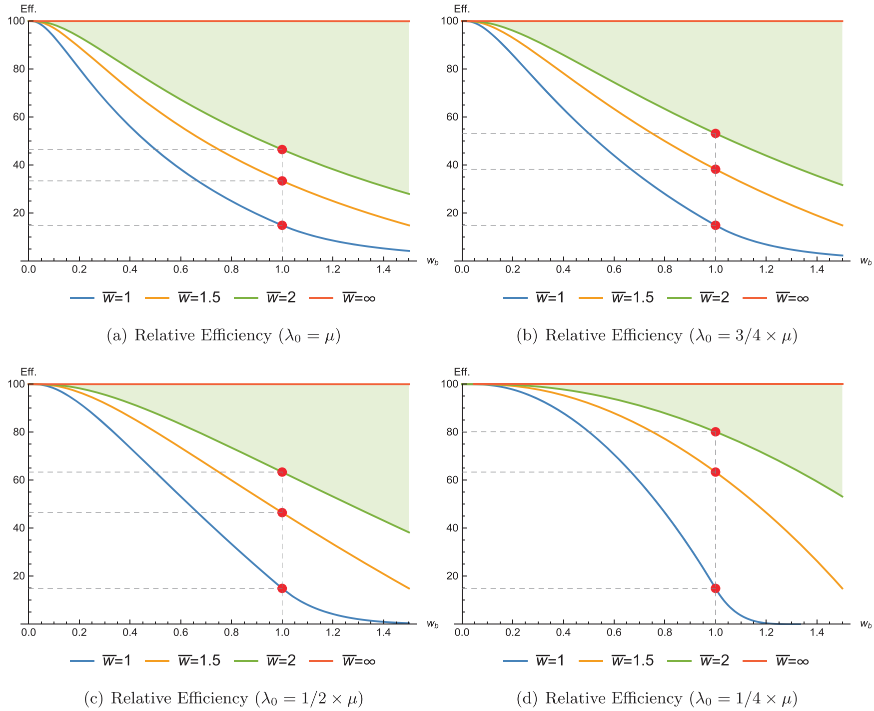
Figure 4 Leverage densities. This figure contains the probability densities for MV leverage under unrestricted and optimal shrinkage leverage. These results are based on a forecasting skill level, ρ, of 0.4; the "other risk" compensation parameter, λ_0, equal to μ (panel (a)), $3/4 \times \mu$ (panel (b)), $1/2 \times \mu$ (panel (c)), and $1/4 \times \mu$ (panel (d)); and RP level, θ, such that the unconditional MV leverage equals unity. (a) Leverage ($\lambda_0 = \mu$). (b) Leverage ($\lambda_0 = 3/4 \times \mu$). (c) Leverage ($\lambda_0 = 1/2 \times \mu$). (d) Leverage ($\lambda_0 = 1/4 \times \mu$).

presence of the standard margin constraint. A similar observation is made by Moreira and Muir (2017).

The benefit of using a theoretical approach in the current context is that the determinants of performance can be isolated and examined (in contrast to previous empirical studies). To demonstrate, consider a special case in which the unconditional MV leverage is equal to the margin constraint (i.e., when $w_b = \overline{w}$). The plots demonstrate that the efficiency is approximately equal to 15%. Interestingly, this efficiency is invariant to conditions (as measured by λ_0)—evinced by noting that this efficiency is the same across all four plots. The expression in this special case is relatively compact (unlike in the more general case when $w_b \neq \overline{w}$). Specifically (percentage) relative efficiency (denoted $\nabla\mathcal{G}$) is given by

$$\nabla\mathcal{G} = 100 \times \left(\frac{1}{2} - \frac{\exp[C^2]\mathrm{erf}\left[\frac{3C}{2\sqrt{2}}\right] - 3\mathrm{erf}\left[\frac{C}{2\sqrt{2}}\right]}{2(\exp[C^2] - 1)} \right), \tag{22}$$

where erf[.] is the Gauss error function. The absence of the conditions variable (λ_0) means that efficiency is entirely determined by skill (as measured by C). Higher skill levels lead to



(a) Relative Efficiency ($\lambda_0 = \mu$)

(b) Relative Efficiency ($\lambda_0 = 3/4 \times \mu$)

(c) Relative Efficiency ($\lambda_0 = 1/2 \times \mu$)

(d) Relative Efficiency ($\lambda_0 = 1/4 \times \mu$)

Figure 5 Performance with leverage constraints. This figure contains the relative efficiency (in percentage terms) of the advantage to the conditional MV strategy (over the unconditional MV strategy) under leverage constraints with respect to the unrestricted leverage advantage. Efficiency is plotted against (buy-and-hold) unconditional MV leverage (w_b). Unconstrained leverage ($\overline{w} = \infty$), a high margin constraint ($\overline{w} = 1$), the standard margin constraint ($\overline{w} = 1.5$), and a low margin constraint ($\overline{w} = 2$) are considered. These results are based on a forecasting skill level, ρ, of 0.4; the "other risk" compensation parameter, λ_0, equal to μ (panel (a)), $3/4 \times \mu$ (panel (b)), $1/2 \times \mu$ (panel (c)), and $1/4 \times \mu$ (panel (d)); and RP levels, θ, such that unconditional MV leverage varies from 0 to 1.5. (a) Relative efficiency ($\lambda_0 = \mu$). (b) Relative efficiency ($\lambda_0 = 3/4 \times \mu$). (c) Relative efficiency ($\lambda_0 = 1/2 \times \mu$). (d) Relative efficiency ($\lambda_0 = 1/4 \times \mu$).

lower efficiencies because such skill levels require higher leverage levels (hence the constraint becomes more restrictive). The expression also provides the upper bound of efficiency. Specifically, the limit of efficiency is 50% as skill levels approach zero. While these appear meager, they are the result of a highly restrictive leverage constraint. In the more general case when constraints are less restrictive (i.e., when $w_b < \overline{w}$) then high-efficiency levels are possible (as shown in Figure 5), with conditions contributing to this efficiency.

2.6 Economic Significance

To assess the economic significance of volatility timing strategies, we decompose the advantage of the conditional MV strategy (in comparison to the unconditional MV strategy) in the leverage effect and the performance fee effect (as per Proposition 4). Results are provided in panel A of Table 4 (with the same parameter values as used above). The performance fees (per unit of leverage) appear to be economically significant. For instance, for a

Table 4 Decomposed performance (conditional MV versus unconditional MV strategies)

Measure	Conditions (λ_0)	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
		Forecasting skill (ρ)										
Panel A: Performance with unrestricted leverage												
$\theta\mathrm{E}[w_a^*]$	μ	1.173	1.248	1.506	2.058	3.188	5.595	11.130	25.080	64.030	185.300	607.500
	$3/4 \times \mu$	1.173	1.230	1.423	1.837	2.684	4.490	8.642	19.120	48.430	140.200	466.600
	$1/2 \times \mu$	1.173	1.211	1.339	1.616	2.181	3.386	6.159	13.170	32.920	95.890	335.000
	$1/4 \times \mu$	1.173	1.192	1.256	1.395	1.678	2.282	3.674	7.207	17.260	50.350	192.300
δ^*	μ	0.000	0.230	0.841	1.635	2.402	3.003	3.399	3.622	3.730	3.776	3.793
	$3/4 \times \mu$	0.000	0.131	0.500	1.030	1.605	2.105	2.463	2.675	2.781	2.826	2.843
	$1/2 \times \mu$	0.000	0.059	0.236	0.521	0.878	1.242	1.538	1.730	1.832	1.876	1.893
	$1/4 \times \mu$	0.000	0.015	0.063	0.151	0.285	0.461	0.646	0.794	0.884	0.926	0.943
Panel B: Performance with 50% shrinkage leverage												
$\theta\mathrm{E}[w_a^*]$	μ	1.173	1.211	1.339	1.616	2.180	3.384	6.150	13.120	32.600	93.230	304.400
	$3/4 \times \mu$	1.173	1.201	1.298	1.505	1.929	2.832	4.908	10.150	24.800	70.710	233.900
	$1/2 \times \mu$	1.173	1.192	1.256	1.394	1.677	2.280	3.667	7.175	17.050	48.550	168.100
	$1/4 \times \mu$	1.173	1.182	1.215	1.284	1.426	1.728	2.425	4.195	9.227	25.800	96.910
δ^*	μ	0.000	0.178	0.709	1.562	2.634	3.725	4.613	5.191	5.495	5.628	5.678
	$3/4 \times \mu$	0.000	0.101	0.411	0.943	1.675	2.504	3.253	3.780	4.072	4.204	4.253
	$1/2 \times \mu$	0.000	0.045	0.189	0.453	0.856	1.383	1.937	2.382	2.652	2.779	2.828
	$1/4 \times \mu$	0.000	0.011	0.049	0.123	0.252	0.456	0.734	1.023	1.240	1.356	1.403
Panel C: Performance with optimal-β shrinkage leverage												
$\theta\mathrm{E}[w_a^*]$	μ	1.173	1.246	1.469	1.858	2.470	3.428	4.973	7.579	12.200	20.810	37.730
	$3/4 \times \mu$	1.173	1.228	1.401	1.713	2.218	3.026	4.348	6.596	10.610	18.160	33.450
	$1/2 \times \mu$	1.173	1.210	1.329	1.554	1.934	2.563	3.618	5.443	8.749	15.160	29.380
	$1/4 \times \mu$	1.173	1.192	1.253	1.377	1.600	1.995	2.695	3.955	6.319	11.150	23.790
δ^*	μ	0.000	0.230	0.851	1.718	2.707	3.725	4.699	5.563	6.267	6.789	7.142
	$3/4 \times \mu$	0.000	0.132	0.504	1.066	1.757	2.516	3.277	3.978	4.563	5.004	5.305
	$1/2 \times \mu$	0.000	0.059	0.237	0.531	0.931	1.413	1.937	2.449	2.894	3.240	3.480
	$1/4 \times \mu$	0.000	0.015	0.063	0.152	0.292	0.492	0.745	1.025	1.294	1.517	1.678

Notes: This table contains measures of decomposed performance associated with the conditional MV strategy in comparison to the unconditional MV strategy. The decomposition consists of two measures: the (annualize percentage) performance fee and the expected leverage (multiplied by θ). These are based on forecasting skill levels from zero to one (in increments of 0.1), and (strict) ICAPM violations of $\lambda_0 = \mu$, $\lambda_0 = 3\mu/4$, $\lambda_0 = \mu/2$, and $\lambda_0 = \mu/4$. The measures are provided under three scenarios: unrestricted leverage (panel A), 50% shrinkage leverage (panel B), and optimal-β shrinkage leverage (panel C).

forecasting skill, ρ, of 0.4, the (per annum) performance fee ranges from 2.402% ($\lambda_0 = \mu$) to 0.285% ($\lambda_0 = \mu/4$). However, to achieve these rewards, extreme leverage levels are required. Using the same forecasting skill level, the expected leverage multiplied by the RP parameter (i.e., $\theta\mathrm{E}[w_a^*]$) ranges from 3.188 ($\lambda_0 = \mu$) to 1.678 ($\lambda_0 = \mu/4$). It follows that an investor with a RP of say $\theta = 3$, would have expected leverage levels of 1.063 (= 3.188/3) and 0.559 (= 1.678/3) over this space. What is also noticeable is that as forecasting skill increases, the proportional increase in the performance fee is always lower than the proportional increase in the expected leverage. Thus the benefits associated with higher forecasting skill place a disproportionate loading on expected leverage.

The need for excessive leverage to achieve the above gains can be tempered by shrinking the leverage. Panels B and C provide results for 50% shrinkage leverage, and optimal-β shrinkage leverage, respectively. In both cases, there is a noticeable lowering in the required leverage rates, while maintaining or increasing the performance fee. For instance, under optimal-β shrinkage with a forecasting skill, ρ, of 0.4, and an RP, θ, of 3, performance fees of 2.707% ($\lambda_0 = \mu$) and 0.292% ($\lambda_0 = \mu/4$) are obtained. Moreover, the expected leverage levels are 0.823 (= 2.470/3) and 0.533 (= 1.600/3) over this space. These are reasonable in terms of the reward to the strategy and do not seem to rely on excessively high leverage levels. Such a strategy would appear to suit those investors who are sensitive to transaction costs, or who face constraints on the amount of borrowing that they can undertake. See Moreira and Muir (2017, 2019) and Barroso and Detzel (2021) for empirical evidence of the impact of transaction costs and borrowing restrictions on conditional VM performance.

2.7 Conditional MV versus Conditional VM Performance

It is interesting to consider the advantage of using the conditional MV strategy over the conditional VM strategy—this is particularly true because the latter is a more popular choice of strategy both from a practitioner and academic perspective. As demonstrated in Section 1.3, the VM strategy represents a restrictive version of the conditional MV strategy. To quantify the advantage over the VM strategy, we calculate the VM-implied RP, and the associated performance fee and expected leverage levels. The same parameter set is used as above. Moreover, we have assumed that the conditional MV and VM forecasting skills are equal (even though the latter strategy assumes a more basic forecasting model based on the previously observed realized variance). Results are provided in Table 5.

The VM-implied θ values demonstrate that the VM strategy is only applicable to a tiny subset of investors. Even within this tight RP space, there are considerable benefits to the conditional MV strategy over the VM strategy, though the advantage is highly nonlinear over the λ_0 space. When $\lambda_0 = \mu$, the performance fee equals zero (as expected as both strategies share this assumption). As λ_0 falls the advantage of the conditional MV strategy starts to reveal itself. For instance, when forecasting skill, ρ, equals 0.4 and $\lambda_0 = \mu/2$, the performance fee equals 0.878% and the expected leverage is 1.126. As λ_0 decreases further the performance fee increases, but the expected leverage falls, until λ_0 reaches zero where there are no benefits to using the conditional MV strategy over the conditional VM strategy (or indeed over the unconditional MV strategy). Despite this nonlinear profile, the users of the conditional MV strategy are never worse off in comparison to the conditional VM strategy users.

2.8 Consistency with Previous Studies

Recent empirical studies show that the performance of volatility timing strategies is somewhat mixed; see, for example, Moreira and Muir (2019) and Cederburg et al. (2020). Such findings are consistent with the empirical predictions of this article. First, forecasting skill is an important driver of performance. However, from an empirical perspective, this skill is unlikely to vary considerably. This is because of the strong persistence in volatility such that it is unlikely that out-of-sample forecasting skill would fall below 0.3 (see the in-

Table 5 Decomposed performance (conditional MV versus VM strategies)

Measure	Conditions (λ_0)	0.0	Forecasting skill (ρ) 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
VM-implied θ		1.173	1.210	1.329	1.554	1.937	2.578	3.714	6.203	15.180	63.480	368.400
$E[w_a^*]$	μ	1.000	1.032	1.133	1.324	1.646	2.170	2.996	4.043	4.217	2.919	1.649
	$3/4 \times \mu$	1.000	1.016	1.070	1.182	1.386	1.742	2.326	3.079	3.180	2.186	1.224
	$1/2 \times \mu$	1.000	1.000	1.008	1.039	1.126	1.313	1.656	2.115	2.140	1.449	0.788
	$1/4 \times \mu$	1.000	0.985	0.945	0.897	0.866	0.884	0.986	1.151	1.104	0.721	0.370
δ^*	μ	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$3/4 \times \mu$	0.000	0.015	0.056	0.114	0.178	0.234	0.274	0.297	0.309	0.314	0.316
	$1/2 \times \mu$	0.000	0.059	0.236	0.521	0.878	1.242	1.538	1.730	1.832	1.876	1.893
	$1/4 \times \mu$	0.000	0.136	0.567	1.358	2.569	4.149	5.811	7.147	7.956	8.338	8.484

Notes: This table contains the VM-implied RPs and measures of decomposed performance associated with the conditional MV strategy in comparison to the conditional VM strategy. The decomposition consists of two measures: the (annualized percentage) performance fee and the expected leverage. These are based on forecasting skill levels from zero to one (in increments of 0.1), and (strict) ICAPM violations of $\lambda_0 = \mu$, $\lambda_0 = 3\mu/4$, $\lambda_0 = \mu/2$, and $\lambda_0 = \mu/4$.

sample forecasting skill of 0.343 in Table 1). Second, the most likely cause of the previously observed variation in performance is the variation in the degree to which the (strict) ICAPM fails to hold over time. If previous studies are using data that cover periods in which the (strict) ICAPM is only marginally failing, then their performance levels will be close to zero. This finding is consistent with the time variation observed in the risk–return relationship (see, e.g., Lettau and Ludvigson 2004) and the results presented in Table 2. Third, for studies that incorporate transaction costs, they are likely to find only marginal performance levels. This is because volatility timing strategies generally have incredibly high leverage requirements (the exception being those users with high-risk aversion levels and those that adopt leverage shrinkage techniques). Finally, many previous studies make use of the conditional VM strategy. Given the evidence in Table 5, it follows that the findings of such studies may underestimate the true performance advantages of volatility timing strategies such as the conditional MV strategy.

3 Conclusion

The results demonstrate the potential virtues of volatility timing strategies. Their success depends on two factors. Perhaps, most obviously the user requires forecasting skill, that is, the ability to predict volatility in the next period. This is however a necessary, but not sufficient condition. One other condition is required. Specifically, the strict version of the ICAPM must fail to hold in that the intercept must be different from zero. This amounts to saying that the investment opportunity set varies over time and returns are determined by "other risks" apart from market risk (Ghysels et al. 2005). A nonzero intercept in the ICAPM equation yields time-varying leverage. This delivers potential benefits to the conditional MV volatility timing strategy over the unconditional MV strategy, and the conditional VM strategy—though the advantage over the latter is more nuanced. In most cases high levels of leverage are required; however, this can be tempered by the use of a proposed leverage shrinkage method.

The requirement of (strict) ICAPM failure over all periods may be a tough condition to be met. However, it is quite possible that the ICAPM intercept may vary over time such that the benefits to volatility timing may be confined to certain periods (see Lustig and Verdelhan 2012, for empirical evidence of business cycle variation in the risk–return relationship). If this is indeed the case then the use of volatility timing strategies should be an occasional pursuit. Similar variation may exist in cross-section such that portfolios (or factors) may exhibit differences in the risk–return relationship. Consequently, differences in the performance of volatility timing strategies applied to these portfolios (or factors) are likely; see, for example, Moreira and Muir (2017) for empirical variation across a range of pricing factor portfolios.

APPENDIX

A Proofs

A.1 Proof of Proposition 1

Proof. Using the law of total expectation we have $\mathrm{E}[Y_{a,t} - Y_{b,t}] = \mathrm{E}[Y^e_{a,t} - Y^e_{b,t}]$. The expressions for $Y^e_{a,t}$ and $Y^e_{b,t}$ are obtained using Equations (4) and (6), such that

$$\mathcal{G} = \mathrm{E}[Y^e_{a,t} - Y^e_{b,t}] = \frac{\lambda_0^2}{4\theta\sigma^4} \times \mathrm{E}\left[\frac{(\mathrm{Var}[R_t|\mathcal{F}_{t-1}] - \sigma^2)^2}{\mathrm{Var}[R_t|\mathcal{F}_{t-1}]}\right]. \tag{A1}$$

Under the assumption that innovations to the log of stochastic variance (ν_t) are i.i.d. with a Gaussian distribution with zero MV D^2 we have

$$\mathrm{Var}[R_t|\mathcal{F}_{t-1}] = \exp[D^2/2]X_t, \tag{A2}$$

where $X_t \equiv \exp\left[\mathrm{E}[\ln[\sigma_t^2]|\mathcal{F}_{t-1}]\right] \sim \mathrm{LN}(A, C^2)$. Using this result and taking the expectation in Equation (A1) delivers the result in this proposition. □

A.2 Proof of Proposition 2

Proof. The expression in Equation (6) indicates that leverage is a constant plus the reciprocal of a log-normal distributed variable, that is $w_a^* = \kappa_1 + \kappa_2/X$. Rearranging this expression such that $X = \kappa_2/(w_a^* - \kappa_1)$, recalling that X has a log-normal distribution, and making use of the change-of-variable technique delivers the desired density function. □

A.3 Proof of Proposition 3

Proof. Under the performance fee set definition, we set the advantage to zero such that

$$\mathrm{E}[w_{a,t}\mathrm{E}[R_t - \delta|\mathcal{F}_{t-1}] - w_{a,t}^2\theta\mathrm{Var}[R_t - \delta|\mathcal{F}_{t-1}] - Y^e_{b,t}] = 0. \tag{A3}$$

Solving for δ leads to the expression in the proposition. □

A.4 Proof of Proposition 4

Proof. Rearranging the definition of the maximum performance fee, we have

$$\underbrace{\mathrm{E}[w_{a,t}\mathrm{E}[R_t|\mathcal{F}_{t-1}] - w_{a,t}^2\theta\mathrm{Var}[R_t|\mathcal{F}_{t-1}] - Y_{b,t}^e]}_{\mathcal{G}} - \mathrm{E}[w_{a,t}\delta] = 0. \tag{A4}$$

Further rearranging gives us the expression in the proposition. □

A.5 Proof of Proposition 5

Proof. This follows from the definition of the maximum performance fee and the expression for the shrinkage leverage. □

A.6 Proof of Proposition 6

Proof. Leverage shrinkage is given by the weighted average of the conditional MV leverage, $w_a^* = \kappa_1 + \kappa_2/X$, and the unconditional MV leverage, w_b^*. Combining we have $\tilde{w}_a^* = \beta(\kappa_1 + \kappa_2/X) + (1-\beta)w_b^*$. Rearranging this expression such that $X = \beta\kappa_2/(\tilde{w}_a^* + (\beta-1)w_b^* - \beta\kappa_1)$, recalling that X has a log-normal distribution, and making use of the change-of-variable technique delivers the desired density function. □

A.7 Proof of Proposition 7

Proof. Differentiating the expression in Equation (15), setting to zero, and solving for β leads to the result in this proposition. □

A.8 Proof of Proposition 8

Proof. Using the law of total variance and invoking the log-normal distribution assumed in this article, the unconditional variance of (unscaled) returns is therefore given by

$$\mathrm{Var}[R_t] = \mathrm{E}[\mathrm{Var}[R_t|\mathcal{F}_{t-1}]] + \mathrm{Var}[\mathrm{E}[R_t|\mathcal{F}_{t-1}]]. \tag{A5}$$

The first term on the RHS is given by

$$\mathrm{E}[\mathrm{Var}[\epsilon_t|\mathcal{F}_{t-1}]] = \sigma^2, \tag{A6}$$

while the second term is given by

$$\mathrm{Var}[\mathrm{E}[R_t|\mathcal{F}_{t-1}]] = (\exp[C^2] - 1)(\mu - \lambda_0)^2. \tag{A7}$$

Thus, using Equations (A6) and (A7), we have

$$\mathrm{Var}[R_t] = (\exp[C^2] - 1)(\mu - \lambda_0)^2 + \sigma^2. \tag{A8}$$

By contrast, the VM unconditional variance is given by

$$\mathrm{Var}[R_{vm,t}] = \mathrm{Var}[w_{vm,t}R_t] = \mathrm{E}[\mathrm{Var}[w_{vm,t}R_t|\mathcal{F}_{t-1}]] + \mathrm{Var}[\mathrm{E}[w_{vm,t}R_t|\mathcal{F}_{t-1}]]. \tag{A9}$$

The first term on the RHS can be expressed as

$$\mathrm{E}[\mathrm{Var}[w_{vm,t}R_t|\mathcal{F}_{t-1}]] = \frac{L^2\exp[C^2]}{\sigma^2}, \tag{A10}$$

while the second term on the RHS is given by

$$\mathrm{Var}[\mathrm{E}[w_{vm,t}R_t|\mathcal{F}_{t-1}]] = \frac{L^2(\exp[C^2]-1)\exp[2C^2]\lambda_0^2}{\sigma^4}. \tag{A11}$$

Combining Equations (A10) and (A11) leads to

$$\text{Var}[R_{vm,t}] = \frac{L^2 \exp[C^2](\exp[C^2](\exp[C^2]-1)\lambda_0^2) + \sigma^2)}{\sigma^4}. \tag{A12}$$

The value of L is obtained by setting the unconditional variances in Equations (A8) and (A12) equal to each other, and solving for L. Doing this we have

$$L = \frac{\exp[-C^2/2]\sigma^2\sqrt{(\exp[C^2]-1)(\mu-\lambda_0)^2+\sigma^2}}{\sqrt{\exp[C^2](\exp[C^2]-1)\lambda_0^2+\sigma^2}}. \tag{A13}$$

Moreover, note that in this article, the conditional MV leverage is given by $\text{E}[R_t|\mathcal{F}_{t-1}]/(2\theta\text{Var}[R_t|\mathcal{F}_{t-1}])$. Comparing this with the VM leverage $L/\text{Var}[R_t|\mathcal{F}_{t-1}])$, we can see that the approaches are equivalent when $\theta = \text{E}[R_t|\mathcal{F}_{t-1}]/2L$. As such, only a time-varying RP ensures equivalence. To force a fixed RP that equates both approaches we impose the restriction that $\lambda_1 = 0$ (and therefore, $\text{E}[R_t|\mathcal{F}_{t-1}] = \mu$) such that $\theta = \mu/2L$. Substituting the expression in Equation (A13) into this expression for θ leads to the RP implied by the VM approach. Thus the VM approach represents a restricted version of the approach adopted in this article, with the restriction given in Equation (20). □

A.9 Proof of Proposition 9

Proof. Under the performance fee set definition, we set the advantage to zero such that

$$\begin{aligned}&\text{E}[w_{a,t}\text{E}[R_t-\delta|\mathcal{F}_{t-1}] - w_{a,t}^2\theta_{vm}\text{Var}[R_t-\delta|\mathcal{F}_{t-1}]\\&\quad-(\text{E}[w_{vm,t}\text{E}[R_t|\mathcal{F}_{t-1}] - w_{vm,t}^2\theta_{vm}\text{Var}[R_t|\mathcal{F}_{t-1}])]\\&\quad=0.\end{aligned} \tag{A14}$$

Solving for δ leads to the result in this proposition. □

B Alternative distribution assumptions

Recall the general expression for the advantage:

$$\mathcal{G} = \text{E}[Y_{a,t}^e - Y_{b,t}^e] = \frac{\lambda_0^2}{4\theta\sigma^4} \times \text{E}\left[\frac{(\text{Var}[R_t|\mathcal{F}_{t-1}]-\sigma^2)^2}{\text{Var}[R_t|\mathcal{F}_{t-1}]}\right]. \tag{A15}$$

Under the assumption that innovations to the log of stochastic variance (ν_t) are independently distributed (though not necessarily Gaussian), we have

$$\text{Var}[R_t|\mathcal{F}_{t-1}] = \text{E}[\sigma_t^2|\mathcal{F}_{t-1}] = \text{E}[\exp[\nu_t]|\mathcal{F}_{t-1}]X_t = \text{E}[\exp[\nu_t]]X_t, \tag{A16}$$

where $X_t \equiv \exp[\text{E}[\ln[\sigma_t^2]|\mathcal{F}_{t-1}]]$. For the expectations component of this expression, we have the general result that if $X_t \sim \text{LN}(A, C^2)$ then

$$\text{E}\left[\frac{(X_t\text{E}[\exp[\nu_t]] - \text{E}[X_t]\text{E}[\exp[\nu_t]])^2}{X_t\text{E}[\exp[\nu_t]]}\right] = \exp[A + C^2/2]\text{E}[\exp[\nu_t]](\exp[C^2]-1). \tag{A17}$$

Furthermore, note that

$$\sigma^2 = \mathrm{E}[\mathrm{Var}[R_t|\mathcal{F}_{t-1}]] = \mathrm{E}[\exp[\nu_t]]\mathrm{E}[X_t] = \mathrm{E}[\exp[\nu_t]]\exp[A + C^2/2]. \tag{A18}$$

Using this expression in Equation (A17) and substituting into Equation (A15), we obtain

$$\mathcal{G} = \frac{(\exp[C^2] - 1)\lambda_0^2}{4\theta\sigma^2}, \tag{A19}$$

which is the same formula as in Proposition 1. Thus the result is robust to alternative distribution assumptions for ν_t providing that the corresponding expectation exists.

References

Andersen, T., T. Bollerslev, F. Diebold, and P. Labys. 2003. Modeling and Forecasting Realised Volatility. *Econometrica* 71: 579–625.

Asness, C., A. Frazzini, and L. Pedersen. 2012. Leverage Aversion and Risk Parity. *Financial Analysts Journal* 68: 47–59.

Barroso, P., and P. Santa-Clara. 2015. Momentum Has Its Moments. *Journal of Financial Economics* 116: 111–120.

Barroso, P., and A. Detzel. 2021. Do Limits to Arbitrage Explain the Benefits of Volatility-Managed Portfolios? *Journal of Financial Economics* 140: 744–767.

Becker, C., W. Ferson, D. Myers, and M. Schill. 1999. Conditional Market Timing with Benchmark Investors. *Journal of Financial Economics* 52: 119–148.

Bollen, N., and J. Busse. 2001. On the Timing Ability of Mutual Fund Managers. *The Journal of Finance* 56: 1075–1094.

Bongaerts, D., X. Kang, and M. van Dijk. 2020. Conditional Volatility Targeting. *Financial Analysts Journal* 76: 54–71.

Brandt, M., F. Nucera, and G. Valente. 2019. Can Hedge Funds Time the Market? *International Review of Finance* 19: 459–469.

Breen, W., R. Jagannathan, and A. Ofer. 1986. Correcting for Heteroscedasticity in Tests for Market Timing Ability. *The Journal of Business* 59: 585–598.

Cederburg, S., M. O'Doherty, F. Wang, and X. Yan. 2020. On the Performance of Volatility-Managed Portfolios. *Journal of Financial Economics* 138: 95–117.

Chance, D., and M. Hemler. 2001. The Performance of Professional Market Timers: Daily Evidence from Executed Strategies. *Journal of Financial Economics* 62: 377–411.

Chang, E., and W. Lewellen. 1984. Market Timing and Mutual Fund Investment Performance. *The Journal of Business* 57: 57–72.

Chiriac, R., and V. Voev. 2011. Modelling and Forecasting Multivariate Realised Volatility. *Journal of Applied Econometrics* 26: 922–947.

Corsi, F. 2009. A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics* 7: 174–196.

Daniel, K., and T. Moskowitz. 2016. Momentum Crashes. *Journal of Financial Economics* 122: 221–247.

Ding, J., L. Jiang, X. Liu, and L. Peng. 2020. "Nonparametric Tests for Market Timing Ability Using Daily Mutual Fund Returns." Working Paper. SSRN. Available at SSRN: https://ssrn.com/abstract=3542845 or http://dx.doi.org/10.2139/ssrn.3542845 (Accessed on 27 January 2022).

Dopfel, F., and S. Ramkumar. 2013. Managed Volatility Strategies: Applications to Investment Policy. *The Journal of Portfolio Management* 40: 27–39.

Fleming, J., C. Kirby, and B. Ostdiek. 2001. The Economic Value of Volatility Timing. *The Journal of Finance* 56: 329–352.

Fleming, J., C. Kirby, and B. Ostdiek. 2003. The Economic Value of Volatility Timing Using "Realized" Volatility. *Journal of Financial Economics* 67: 473–509.

Ghysels, E., P. Santa-Clara, and R. Valkanov.2005. There is a Risk-Return Trade-off after All. *Journal of Financial Economics* 76: 509–548.

Grinblatt, M., and S. Titman. 1988. "The Evaluation of Mutual Fund Performance: An Analysis of Monthly Returns." Working Paper. University of California, Los Angeles (03 1988).

Harvey, C., E. Hoyle, R. Korgaonkar, S. Rattray, M. Sargaison, and O. van Hemert. 2018. The Impact of Volatility Targeting. *The Journal of Portfolio Management* 45: 14–33.

Henriksson, R. 1984. Market Timing and Mutual Fund Performance: An Empirical Investigation. *The Journal of Business* 57: 73–96.

Henriksson, R., and R. Merton. 1981. On Market Timing and Investment Performance. II. Statistical Procedures for Evaluating Forecasting Skills. *The Journal of Business* 54: 513–533.

Hocquard, A., S. Ng, and N. Papageorgiou. 2013. A Constant-Volatility Framework for Managing Tail Risk. *The Journal of Portfolio Management* 39: 28–40.

Hoyle, E., and N. Shephard. 2019. "When Does Volatility Scaling Improve the Unconditional Sharpe Ratio Compared to a Buy-and-hold Position." Working Paper. Man Institute.

Jiang, W. 2003. A Nonparametric Test of Market Timing. *Journal of Empirical Finance* 10: 399–425.

Lettau, M., and S. Ludvigson. 2004. "Measuring and Modeling Variation in the Risk-Return Tradeoff." In Y. Ait-Sahalia and L. P. Hansen (eds.), *Handbook of Financial Economics*. North-Holland, Amsterdam: Elsevier Science.

Liu, F., X. Tang, and G. Zhou. 2019. Volatility-Managed Portfolio: Does It Really Work? *The Journal of Portfolio Management* 46: 38–51.

Lustig, H., and A. Verdelhan. 2012. Business Cycle Variation in the Risk-Return Trade-off. *Journal of Monetary Economics* 59: S35–S49.

Maillard, S., T. Roncalli, and J. Teïletche. 2010. The Properties of Equally Weighted Risk Contribution Portfolios. *The Journal of Portfolio Management* 36: 60–70.

Mamaysky, H., M. Spiegel, and H. Zhang. 2008. Estimating the Dynamics of Mutual Fund Alphas and Betas. *Review of Financial Studies* 21: 233–264.

Marquering, W., and M. Verbeek. 2004. The Economic Value of Predicting Stock Index Returns and Volatility. *Journal of Financial and Quantitative Analysis* 39: 407–429.

Mascio, D., F. Fabozzi, and K. Zumwalt. 2021. Market Timing Using Combined Forecasts and Machine Learning. *Journal of Forecasting* 40: 1–16.

Merkoulova, Y. 2020. Predictive Abilities of Speculators in Energy Markets. *Journal of Futures Markets* 40: 804–815.

Merton, R. 1981. On Market Timing and Investment Performance I. An Equilibrium Theory of Value for Market Forecasts. *The Journal of Business* 54: 363–406.

Moreira, A., and T. Muir. 2017. Volatility-Managed Portfolios. *The Journal of Finance* 72: 1611–1643.

Moreira, A., and T. Muir. 2019. Should Long-Term Investors Time Volatility? *Journal of Financial Economics* 131: 507–527.

Meddahi, N. 2003. ARMA Representation of Integrated and Realized Variances. *The Econometrics Journal* 6: 335–356.

O'Hagan, A., and T. O. M. Leonard. 1976. Bayes Estimation Subject to Uncertainty about Parameter Constraints. *Biometrika* 63: 201–203.

Sharpe, W. 1975. Likely Gains for Market Timing. *Financial Analysts Journal* 31: 60–69.

Shephard, N., and K. Sheppard. 2010. Realising the Future: Forecasting with High Frequency Based Volatility (Heavy) Models. *Journal of Applied Econometrics* 25: 197–231.

Smith, T. 2020. “There are only two types of investor.” *Financial Times*, 2 July. Available at https://www.ft.com/content/d4e6dc11-816a-44d6-86d1-b195624b1c64 (Accessed 10 August 2020).

Taylor, N. 2014. The Economic Value of Volatility Forecasts: A Conditional Approach. *Journal of Financial Econometrics* 12: 433–478.

Taylor, N. 2017. Risk Control: Who Cares? *European Financial Management* 23: 153–179.

Treynor, J., and K. Mazuy. 1966. Can Mutual Funds Outguess the Market? *Harvard Business Review* 44: 131–136.

Welch, I., and A. Goyal. 2008. A Comprehensive Look at the Empirical Performance of Equity Premium Prediction. *Review of Financial Studies* 21: 1455–1508.