Understanding the Relation between the Statistical and Economic Significance of Predictability

BARRAS, Laurent Richard

Abstract

We examine the economic gains produced by equity (and bond) conditional strategies when the out-of-sample predictive power, R2, is close to zero. To address this issue, we design a model which expresses the performance of the conditional strategy as a function of any given level of R2. The relation between R2 and performance is studied in different settings, including one and multiple assets, as well as short-selling constraints. In all cases, a level of R2 as small as 0.5% is sufficient to greatly outperform the unconditional (passive) strategy. For a given R2, the profits are much higher if the source of predictability comes from asset-selectivity rather than factor-timing. In addition, we find that under asset-selectivity, a single-asset strategy requires a very high R2 in order to perform as well as a multi-asset strategy.

Reference


Available at:
http://archive-ouverte.unige.ch/unige:5733

Disclaimer: layout of this document may differ from the published version.
Understanding the Relation between the Statistical and Economic Significance of Predictability

Laurent BARRAS
Understanding the Relation between the Statistical and Economic Significance of Predictability

Laurent Barras*

JEL Classification: G11
Keywords: Predictability, Conditional Strategy, Factor-Timing, Asset-Selectivity
First version: May 2006
This version: April 2007

*Swiss Finance Institute at HEC-University of Geneva, Boulevard du Pont d’Arve 40, 1211 Geneva 4, Switzerland. Tel: +41223798141. Fax: +41223798104. E-mail: barras@hec.unige.ch. I am grateful to Frédéric Sonney, Olivier Scailllet, and Georges Gatopoulous for helpful comments.
Understanding the Relation between the Statistical and Economic Significance of Predictability

Abstract

We examine the economic gains produced by equity (and bond) conditional strategies when the out-of-sample predictive power, $R^2$, is close to zero. To address this issue, we design a model which expresses the performance of the conditional strategy as a function of any given level of $R^2$. The relation between $R^2$ and performance is studied in different settings, including one and multiple assets, as well as short-selling constraints. In all cases, a level of $R^2$ as small as 0.5% is sufficient to greatly outperform the unconditional (passive) strategy. For a given $R^2$, the profits are much higher if the source of predictability comes from asset-selectivity rather than factor-timing. In addition, we find that under asset-selectivity, a single-asset strategy requires a very high $R^2$ in order to perform as well as a multi-asset strategy.
Introduction

Active investors and academics are interested in assessing the economic significance of equity and bond return predictability. The latter is measured with the performance of conditional strategies based on this predictability (e.g., Breen, Glosten, and Jagannathan, 1989; Pesaran and Timmermann, 1995; Solnik, 1993). The key input of these strategies is the out-of-sample $R^2$ defined as the explanatory power of the predicted values estimated only with past available data\(^1\). Empirically, producing a positive $R^2$ is extremely difficult, as confirmed by the negligible statistical significance of predictability documented by Bossaerts and Hillion (1999) and Goyal and Welch (2003, 2006). One reason for this result is the estimation errors contained in the predicted values. As shown by Cochrane (2006), estimation risk drives the out-of-sample $R^2$ to zero, even if the predictive model based on the true (but unknown) parameters can forecast asset returns. A second reason for this poor $R^2$ is the specification uncertainty faced by the investor. Since there is no theoretical arguments to guide him, the investor cannot detect the predictive models yielding accurate forecasts in the future (Barras, 2007; Cooper and Gulen, 2006). Therefore, a central question for active investors is to know whether an extremely small out-of-sample $R^2$ is still sufficient to generate positive economic gains. In a single-asset framework, Kandel and Stambaugh (1996) and Xu (2004) show that levels of $R^2$ between 2.5% and 5.5% have a substantial impact on the investor portfolio decision. While these studies provide important insights into the relation between the statistical and economic significance of predictability, values between 2.5% and 5.5% do not reflect the poor level of out-of-sample $R^2$. In addition, these papers do not examine the sensitivity of the results to the inclusion of multiple risky assets in the investment universe.

In this paper, we determine the economic significance of predictability produced by levels of $R^2$ close to zero. To address this issue, we design a model which expresses the performance of the mean-variance conditional strategy as a function of any given level of $R^2$. A main virtue of our model is its flexibility. This allows us to examine the relation between $R^2$ and the performance of single- as well as multi-asset strategies in a unified framework. In addition, our model successfully incorporates the two possible sources of predictability generating a given level of $R^2$, namely factor-timing and asset-selectivity. Since these two sources yield different correlations among the predicted values, they are likely to affect the relation between $R^2$ and performance. Similar to Kandel and Stambaugh (1996) and Xu (2004), we calibrate the parameters of our model with real return observations on developed market equity and bond indices\(^2\).

\(1\) For ease of exposition, we refer hereafter to the term "$R^2$" as the out-of-sample $R^2$.

\(2\) There is a growing list of papers using calibration on real data to address different issues in predictability, asset-pricing, or portfolio decision. An non-exhaustive list includes Barberis (2000), Cochrane (2006), Ferson, Sarkissian, and Simin (2003), and Ferson, Siegel, and Xu (2006).
This empirical calibration offers the advantage of reproducing closely the data characteristics of the investment universe used in previous empirical studies, while easily controlling for the variation in the level of $R^2$.

Our results show that levels of predictability considered as negligible from a statistical standpoint are sufficient to produce substantial economic gains. To illustrate this superior performance, consider a conditional strategy implemented with equity indices. When the out-of-sample $R^2$ is as small as 0.5%, the single-asset strategy produces an additional expected return over the unconditional (passive) strategy equal to 0.5% per year (for the same standard deviation). If the investor maintains the same level of $R^2$ for five equity indices, this gain amounts to 1% per year if predictability is entirely due to factor-timing, and 4.7% per year in case of asset-selectivity. This example also shows that the relation between $R^2$ and performance depends on the source of predictability. Under factor-timing, the predicted values are high (or low) at the same time, because all assets depend on the common factors. In order to avoid taking excessive risk, the active investor cannot move the portfolio weights far away from the unconditional (passive) weights. On the contrary, asset-selectivity produces uncorrelated forecasts and allows the investor to take large bets on the few assets with high predicted values. Further, we find that a small $R^2$ generates a positive skewness similar to the one observed for some hedge fund directional strategies such as global macro or managed futures (Lhabitant and Learned, 2002). This result strengthens the economic significance of predictability, as risk-averse investors have a preference for positively-skewed distributions.

Examination of the portfolio weights of the conditional strategy reveals that the total long and short positions can be very large. Since regulatory constraints may prevent the investor from taking these positions, we also examine the performance of the conditional strategy subject to short-selling constraints. Imposing these constraints reduces the positive relation between $R^2$ and performance in all cases, and especially under asset-selectivity. However, a negligible level of $R^2$ is still sufficient to produce positive economic gains. For instance, when $R^2 = 0.5\%$, the additional gain of the conditional strategy based on five equity indices amounts to 1.2% per year under timing, and 2% per year under selectivity. All of these results remain qualitatively unchanged if the investment universe is formed with bond indices.

It can be argued that following an increasing number of assets requires more time, and may reduce the overall forecast accuracy. For this reason, the investor is likely to face a trade-off between the number of assets and the achievable level of $R^2$. Under asset-selectivity, we find that it is easier to improve performance by increasing the number of assets. To illustrate it, consider the performance of the unconstrained conditional strategy based on five equity indices.
when \( R^2 = 1.5\% \) for the five assets. In order to obtain the same performance with a single asset, the investor must produce an extremely high level of \( R^2 \) equal to 14.6% (6.8% with short-selling constraints). By contrast, we find that the reward for increasing the number of assets is less pronounced under factor-timing. We find that achieving a level of \( R^2 = 1.5\% \) for five equity indices is equivalent to producing a level of \( R^2 \) equal to 4.1% for a single asset (3.4% under short-selling constraints).

A further application of our paper examines the importance of the Dybvig and Ross critique (1985) on the reliability of the Sharpe ratio. In a single-asset framework, they show that the relation between \( R^2 \) and the Sharpe ratio is negative under the (sufficient) condition that \( \mu_j > \sigma_j \), where \( \mu_j \) and \( \sigma_j \) are the excess mean and standard deviation of the risky asset \( j \). Using a wide range of values for \( R^2 \) (including high values) in both single- and multi-asset strategies, we observe that the relation between \( R^2 \) and the Sharpe ratio is always positive. It implies that the Sharpe ratio is a robust performance measure, and that the condition of Dybvig and Ross (1985) is too stringent to have empirical implications.

The remainder of the paper is as follows. Section I describes the statistical and economic significance of predictability. It defines the level and the source of predictability, and specifies the design of the conditional strategy. Section II describes the parameterization of the model. Section III contains the results of our analysis.

I. The Statistical and Economic Significance of Predictability

A. General Framework

We consider an investment universe of \( N \) risky assets and a riskless asset. The \( N \times 1 \) vector of excess returns over the riskfree rate between time \( t \) and \( t + 1 \) is denoted by \( r_{t+1} \). At each time \( t \), the investor uses the predictive variables included in his information set \( I_t \) to forecast \( r_{t+1} \):

\[
r_{t+1} = E(r_{t+1} | I_t) + \varepsilon_{t+1} = \Pi_t + \varepsilon_{t+1},
\]

where \( \Pi_t \) denotes the \( N \times 1 \) conditional excess mean vector (i.e., the predicted values), and \( \varepsilon_{t+1} \) is the \( N \times 1 \) vector of unpredictable terms. We assume that the \( N \times N \) conditional covariance matrix is constant: \( \text{Var}(r_{t+1} | I_t) = \Sigma \). In addition, we assume that \( \Pi_t \) and \( \varepsilon_{t+1} \) are normally distributed:

\[
\Pi_t \sim N(\mu, \Omega), \quad \varepsilon_{t+1} \sim N(0, \Sigma),
\]
where $\mu$ denotes the $N \times 1$ unconditional excess mean vector, and $\Omega$ the $N \times N$ covariance matrix of $\Pi_t$. Time-varying volatility as well as fat-tail distributions are more likely over shorter periods, such as daily or weekly frequencies. On the contrary, the conditional strategy examined in this paper should be implemented over a monthly or yearly time-horizon to benefit from the increase in return predictability over time (e.g., Cochrane, 2004, ch.1). This provides an empirical justification to the two assumptions mentioned above. The robustness of our results to non-normal distributions are examined in Section III.

B. The Statistical Significance of Predictability

We determine the statistical significance of predictability with two measures: the level of out-of-sample predictability, $R^2$, and its source, $\eta$. To define these two measures, let us denote by $r_{it+1}$ the excess return of the risky asset $i$ ($i = 1, ..., N$), and $\pi_{it}$ its predicted value estimated with past available data only. For each asset $i$, the explanatory power $R^2_i$ of the predictive model is defined as $R^2_i = \sigma^2_{\pi_{it}}/\sigma^2_{r_{it}}$, where $\sigma^2_{\pi_{it}}$ and $\sigma^2_{r_{it}}$ denote the variance of $\pi_{it}$ and $r_{it+1}$, respectively. Since the investment universe may contain more than a single asset, we measure the level of predictability, $R^2$, as the average predictive power across all assets:

$$R^2 = \frac{1}{N} \sum_{i=1}^{N} R^2_i,$$  \hspace{1cm} (3)  

A given level of $R^2$ can be driven by two sources of predictability. The first one, called factor-timing, forecasts the evolution of the common factors. The second one, called asset-selectivity, predicts the price variation of individual assets. Separating these two cases is important, as they imply different correlation structures among the predicted values. Under factor-timing, correlations are likely to be higher because of the dependence of all assets on the common factors. To address this issue, we postulate a factor structure for $r_{it+1}$ ($i = 1, ..., N$):

$$r_{it+1} = b_i f_{t+1} + \psi_{it+1},$$ \hspace{1cm} (4)  

where $f_{t+1}$ denotes the $K \times 1$ excess return vector of $K$ factor-mimicking portfolios, $b_i$ is the $K \times 1$ vector of factor exposures, and $\psi_{it+1}$ stands for the zero-mean idiosyncratic term. We

\footnote{Using daily returns, Fleming, Kirby, and Ostdiek (2001) find that predicting volatility brings positive profits. However, this is not the case with monthly returns (e.g., Cavaglia et al., 1997; Marquering and Verbeek, 2004; Han, 2005).}
assume that the factors are orthogonal, and that \( \text{cov}(\psi_{it+1}, \psi_{jt+1}) = 0 \) for each pair of assets \( i \) and \( j \). Using Equation (4), \( \pi_{it} \) can be decomposed as:

\[
\pi_{it} = b_i' \pi_{ft} + \pi_{\psi it},
\]

where \( \pi_{ft} = E(f_{t+1} | I_t) \) denotes the \( K \times 1 \) conditional excess mean vector of \( f_{t+1} \), and \( \pi_{\psi it} = E(\psi_{it+1} | I_t) \) is the conditional mean of \( \psi_{it+1} \). Following Grinblatt and Titman (1989), there is factor-timing if \( \pi_{ft} = E(f_{t+1}) \) for at least one realization of \( I_t \). Similarly, there is asset-selectivity if \( \pi_{\psi it} = E(\psi_{it+1}) \) for at least one realization of \( I_t \), and one asset \( i \). Holding the predictive power \( R^2_i \) of each asset \( i \) fixed \((i = 1, ..., N)\), we denote by \( \pi_{T it} \) the asset \( i \) predicted value under pure factor-timing \( \text{i.e., } \pi_{\psi it} = E(\psi_{it+1}) = 0, \forall i \). Similarly, \( \pi_{S it} \) stands for the asset \( i \) predicted value under pure asset-selectivity \( \text{i.e., } \pi_{ft} = E(f_{t+1}) \). In order to define the source of predictability, we use the parameter \( \eta \in [0, 1] \). For all assets, a value \( \eta = 0 \) indicates that predictability is only due to asset-timing. At the other extreme when \( \eta = 1 \), predictability is entirely caused by asset-selectivity:

\[
\pi_{it} = (1 - \eta) \cdot \pi_{T it} + \eta \cdot \pi_{S it}.
\]

C. The Economic Significance of Predictability

In order to assess the economic significance of predictability, we measure the performance of the conditional strategy. The latter uses the predictive information \( \text{contained in } I_t \) in order to determine the \( N \times 1 \) vector of risky asset weights, \( w_{Ct} \). We use the standard mean-variance conditional framework proposed among others by Solnik (1993), Harvey (1994), and Handa and Tiwari (2006). Replacing the riskless asset weight, \( w_{Ht} \), with \( 1 - w_{Ct} \) \( \text{the budget constraint} \), the investor maximizes the following mean-variance conditional function at each time \( t \):

\[
\max_{w_{Ct}} w_{Ct}' \Pi_t - \frac{1}{2} A w_{Ct}' \Sigma w_{Ct},
\]

where \( A \) denotes the investor’s risk aversion coefficient. The optimal solution to Equation (7) is given by \( w_{Ct} = (1/A) \Sigma^{-1} \Pi_t \). Using this result, the excess return of the conditional asset allocation, \( r_{Ct+1} \), is equal to:

\[
r_{Ct+1}(\mu, \Omega, \Sigma) = w_{Ct}' r_{t+1} = \frac{1}{A} \Pi_t' \Sigma^{-1} r_{t+1} = \frac{1}{A} \left[ \Pi_t' \Sigma^{-1} \Pi_t + \Pi_t' \Sigma^{-1} \varepsilon_{t+1} \right],
\]

where the terms \( \Pi_t' \Sigma^{-1} \Pi_t \) and \( \Pi_t' \Sigma^{-1} \varepsilon_{t+1} \) correspond to two sums of products of normal variables. These two terms are not normally distributed, as they do not follow a known parametric distribution (Wade and Lade, 2003). It implies that \( r_{Ct+1} \) is not normally distributed. In order
to measure the performance of the conditional strategy, we compare it to the unconditional mean-variance strategy which does not account for predictability. Replacing the conditional moments $\Pi_t$ and $\Sigma$ with the unconditional ones $\mu$ and $V$ in Equation (7), the $N \times 1$ unconditional weight vector is equal to $w_U = (1/A)V^{-1}\mu$. Since $w_U$ is constant, the excess return of the unconditional strategy, $r_{Ut+1}$, is a linear combination of $r_{t+1}$, and is therefore normally distributed.

We use the Sharpe ratio to compare the performance of these two strategies. The main advantage of this measure is that it does not rely on an asset pricing model. As a result, it can be used regardless of the asset-class (equities or bonds) chosen by the investor. Using Equation (8), the Sharpe ratio of the conditional strategy is equal to (see the appendix):

$$S_C(\mu, \Omega, \Sigma) = \frac{Tr(\Sigma^{-1}\Omega) + \mu'\Sigma^{-1}\mu}{[2Tr((\Sigma^{-1}\Omega)^2) + 4\mu'\Sigma^{-1}\Omega\Sigma^{-1}\mu + Tr(\Sigma^{-1}\Omega) + \mu'\Sigma^{-1}\mu]^2}$$

(9)

where $Tr$ denotes the trace operator. Standard analysis shows that the Sharpe ratio of the unconditional strategy takes the form: $S_U = [\mu/V^{-1}\mu]^2$. While the Sharpe ratio provides important information about the risk-return trade-off of the conditional strategy, the higher moments of $r_{Ct+1}$ can also affect the economic significance of predictability. It is well-known that risk-averse investors prefer positive skewness (Kraus and Litzenberger, 1976; Harvey and Siddique, 2000), and generally dislike excess kurtosis (Scott and Horvath, 1980). To address this issue, we compute the skewness and the kurtosis of the conditional strategy. Because of the difficulty in obtaining tractable expressions for these higher moments, they are computed with a Monte-Carlo simulation-based method. After specifying values for $\mu$, $\Omega$, and $\Sigma$, we draw $Q$ realizations of $\Pi_t$ and $\varepsilon_t$ from the normal distributions shown in Equation (2). Then, we use Equation (8) to form a $Q \times 1$ vector of realization of $r_{Ct+1}$, from which we compute the skewness and kurtosis. We set $Q$ equal to 10,000.

As discussed by Best and Grauer (1991), mean-variance optimization is very sensitive to variations in expected returns, and can produce large (long and short) positions. As regulatory constraints may prevent the investor from taking these positions, we consider a second case with short-selling constraints:

$$\max_{w_{Ct}} w_{Ct}'\Pi_t - \frac{1}{2}Aw_{Ct}'\Sigma w_{Ct} \quad \text{s.t.} \quad 0 \leq w_{Ct} \leq 1, \; w_{Ct}'1 \leq 1,$$

(10)

Since Equation (9) is no longer valid, we use the simulation-based approach to compute the Sharpe ratio of the conditional strategy. After drawing one realization of $\Pi_t$, we determine $w_{Ct}$.
by maximizing the mean-variance function in Equation (10). Then, we draw one realization of 
\( \varepsilon_{t+1} \) and compute 
\[ r_{Ct+1} = w'_C (\Pi_t + \varepsilon_{t+1}) . \]
By repeating this procedure \( Q \) times, we obtain a \( Q \times 1 \) vector of realizations of 
\( r_{Ct+1} \) from which we compute the Sharpe ratio, the skewness, and the kurtosis. For the constrained unconditional strategy, the Sharpe ratio is given by 
\[ S_U = w'_U \mu / [w'_U V w_U]^{1/2} , \]
where \( w_U \) is the optimal unconditional weight vector under short-selling constraints.

II. Parameterization of the Model

A. Specification of the Moments \( \mu, V, \Omega, \) and \( \Sigma \)

The previous analysis shows that the performance measures depend on the four moments \( \mu, V, \Omega, \) and \( \Sigma. \) This section explains how to compute these moments, and how to express them as a function of our measures of statistical predictability, \( R^2 \) and \( \eta. \) First, we specify the unconditional moments \( \mu \) and \( V \) using the factor structure in Equation (4). For each factor \( k \) \((k = 1, ..., K)\) and each asset \( i \) \((i = 1, ..., N)\), we draw beta coefficients from the normal distribution: 
\[ b_{ik} \sim N(\mu_{b_k}, \sigma_{b_k}) . \]
The \( N \times 1 \) unconditional excess mean vector \( \mu \) is given by
\[ \mu = B \mu_f , \] (11)
where \( B \) denotes the \( N \times K \) matrix of asset betas, and \( \mu_f \) is the \( K \times 1 \) factor excess mean vector. We draw the idiosyncratic variance \( \sigma_{\Psi_i}^2 \) for each asset \( i \) from the inverted gamma distribution (Pastor and Stambaugh, 2002): 
\[ \sigma_{\Psi_i}^2 \sim s_0^2 \cdot v_0 / \chi^2(v_0) , \]
where \( s_0^2 \) is a scalar, and \( \chi^2(v_0) \) denotes the chi-square distribution with \( v_0 \) degrees of freedom. The \( N \times N \) unconditional covariance matrix \( V \) is given by
\[ V = BV_f B' + V_\Psi , \] (12)
where \( V_\Psi \) is the \( N \times N \) diagonal matrix containing the idiosyncratic variances \( \sigma_{\Psi_i}^2 \), and \( V_f \) is the \( K \times K \) diagonal factor covariance matrix. The method used to determine the parameters \( \mu_{b_k}, \sigma_{b_k} \) \((k = 1, ..., K)\), \( \mu_f \), \( V_f \), \( s_0^2 \), and \( v_0 \) are explained below.

The next step consists in determining the covariance matrix of the predicted values, \( \Omega \), which depends on both \( R^2 \) and \( \eta. \) To separate these two effects, we model the relation between \( \eta \) and \( \Omega \), holding \( R^2 \) fixed. The variance of each predicted value, \( \sigma_{\pi_i}^2 \), must be independent of \( \eta \) to keep \( R^2 \) constant. This condition, along with Equation (6), implies the following restriction for each asset \( i: \sigma_{\pi_i}^2 = \sigma_{\pi_i^T}^2 = \sigma_{\pi_i^S}^2. \) To determine the covariance for each pair of assets \( i \) and \( j \), \( \text{cov}(\pi_{it}, \pi_{jt}) \), we follow Admati et al. (1986) and assume that the asset \( i \) selectivity
signal is uncorrelated with factor-timing signals, and with asset \( j \) selectivity signal \((\forall j \neq i)\): 
\[
cov \left( \pi_{it}^S, \pi_{jt}^T \right) = \text{cov} \left( \pi_{it}^S, \pi_{jt}^S \right) = 0.
\]
From Equation (6), 
\[
cov \left( \pi_{it}^T, \pi_{jt} \right) = (1 - \eta) \text{cov} \left( \pi_{it}^T, \pi_{jt}^T \right),
\]
indicating that a decrease in \( \eta \) increases the importance of factor-timing, which in turn affects the covariance term\(^5\). Denoting by \( R_f^2 \) the level of factor predictability (assumed to be identical across the \( K \) factors), we show in the appendix that 
\[
cov \left( \pi_{it}^T, \pi_{jt}^T \right) = R_f^2 \beta_i V_f \beta_j.
\]
Consistent with Equation (3), 
\[
R_f^2 (R^2) \text{ is determined such that } R^2 = 1/N \sum_{i=1}^N \left( \frac{\sigma_{\pi_i^T}^2}{\sigma_{\pi_i}^2} \right). 
\]
After collecting all variance and covariance terms, the \( N \times N \) covariance matrix of the predicted values is given by
\[
\Omega \left( R^2, \eta \right) = (1 - \eta) R_f^2 (R^2) B V_f B' \eta V_f,
\]
where \( V_f \) is the \( N \times N \) diagonal matrix containing the elements \( \sigma_{\pi_i}^2 \). The final step is to determine the \( N \times N \) conditional covariance matrix \( \Sigma \). Since \( V = \Omega + \Sigma \), we can use the values for \( V \) and \( \Omega \) given in Equations (12) and (13) to find:
\[
\Sigma \left( R^2, \eta \right) = V - \Omega \left( R^2, \eta \right).
\]

### B. Investment Universe and Parameter Values

We determine the parameters \( \mu_{b_k}, \sigma_{b_k} \) \((k = 1, \ldots, K)\), \( \mu_f \), \( V_f \), \( s_0^2 \), and \( \nu_0 \) by calibrating the factor structure in Equation (4) on the assets forming the investment universe. Since the conditional strategy produces high turnover, it must be implemented using futures markets to reduce transaction costs (e.g., Solnik, 1993). We consider a universe of 12 developed market equity indices on which stock index futures are actively traded (Sutcliffe, 2006): Australia, Belgium, Canada, Denmark, France, Germany, Italy, Japan, Netherlands, Switzerland, the United Kingdom, and the United States. Since most previous studies on the statistical and economic significance of predictability are based on monthly returns, we use the same time-horizon for comparability. Using stock index futures removes the currency risk of foreign investments (Solnik and McLeavey, 2004). As a result, the excess return of each country, \( r_{it+1} \), is computed in local currency, and is independent of the investor’s reference currency (Barras, 2007). \( r_{it+1} \) is proxied by the monthly return of the Datastream country index (in local currency) minus its respective one-month euro-market interest rate.

As shown by Illmanen (1996) and Solnik (1993), the conditional strategy can also be imple-

---

\(^5\)The modification of \( \text{cov} \left( \pi_{it}, \pi_{jt} \right) \) after changes in \( \eta \) depends on the sign of \( \text{cov} \left( \pi_{it}^T, \pi_{jt}^T \right) \) or, alternatively, on the signs of the betas for the different factors. For instance, a one-factor model with positive betas for all assets implies that a decrease in \( \eta \) rises the correlation across all predicted values.
mented with bond indices, as government bond futures are actively traded in many countries. To address this issue, we also calibrate the factor structure on long-term government bond indices from the 12 developed countries listed above. We find that the relation between \( R^2 \) and performance remain qualitatively similar when bonds instead of equities are used. For sake of brevity, the results obtained with bond indices are not shown, but are available upon request.

The parameters are obtained with a Principal Component Analysis (PCA) applied to the empirical covariance matrix of the 12 country excess returns. The monthly excess mean and standard deviation of the factor \( k \) are computed as \( \mu_{f_k} = c_k^T \hat{\mu} / c_k^T 1 \) and \( \sigma_{f_k} = \zeta_k^2 / c_k^T 1 \), where \( \hat{\mu} \) is the 12 × 1 vector of the country average monthly excess returns, \( 1 \) is a 12 × 1 vector of ones, and \( c_k \) is the 12 × 1 eigenvector associated with the \( k \)th eigenvalue, \( \zeta_k \). Each beta is calculated as \( b_{ik} = c_{ik} / c_k^T 1 \), where \( c_{ik} \) is the \( i \)th element of \( c_k \). We set \( \mu_{b_k} \) and \( \sigma_{b_k} \) equal to the cross-sectional mean and standard deviation of the \( b_{ik} \) across the 12 countries. Table 1 displays the estimated parameters for the three highest eigenvalues over the full period (January 1980-December 2005), and over two subperiods of equal length (January 1980-December 1992, and January 1993-December 2005). Over the full period (312 observations), the three-factor model explains on average 76.6% of the country return variance. The bulk of the explanatory power comes from the first factor (63.7%). This factor can be interpreted as the world market portfolio, since its annualized excess mean (7.4%), standard deviation (13.8%), and average beta (0.967) are similar to those obtained with the world market index (Ferson and Harvey, 1994). We observe that the parameters related to the first factor (\( \mu_{f_1}, \sigma_{f_1}, \mu_{b_1}, \) and \( \sigma_{b_1} \)) are very stable over the two subperiods. On the contrary, the empirical moments of the second and third factors experience large swings. This variation is not an issue for the calibration of our model, because the simulations are not sensitive to the choice of \( \mu_{f_2}, \mu_{f_3}, \sigma_{f_2}, \) and \( \sigma_{f_3} \). The reason is that the exposures to these factors, measured by \( \mu_{b_2} \) and \( \mu_{b_3} \), are extremely low. If we omit these two factors from the analysis, the results remain unchanged. This confirms the dominance of the first factor in explaining equity index returns.

To calibrate the three-factor model, we use the parameter values computed over the full period\(^6\). \( s_2^0 \) and \( v_0 \) are determined according to the following equations (Pastor and Stambaugh, 2002):

\[
\begin{align*}
v_0 &= 4 + 2 E (\sigma_{\Psi_i}^2) / \text{var} (\sigma_{\Psi_i}^2) \quad \text{and} \quad s_2^0 = E (\sigma_{\Psi_i}^2) \cdot (v_0 - 2) / v_0,
\end{align*}
\]

where \( E (\sigma_{\Psi_i}^2) \) and \( \text{var} (\sigma_{\Psi_i}^2) \) denote the cross-sectional mean and variance of \( \sigma_{\Psi_i}^2 \). From the estimated idiosyncratic variances of the 12 countries, we set \( s_2^0 = 0.0005 \) and \( v_0 = 8 \). Table 2 compares the characteristics of the actual and simulated country excess returns. In Panel A, we show descriptive statistics for each

\(^6\)Expressed on a monthly basis, the elements of the 3 × 1 vector \( \mu_\lambda \) are equal to 0.0063, 0.0032, and −0.0033. The diagonal elements of the 3 × 3 covariance matrix \( V_\lambda \) are set to 0.0016, 0.0156, and 0.0334.
country between January 1980 and December 2005. These statistics are the annualized excess mean and standard deviation, the cross-country correlation, and the explanatory power of the three-factor model obtained with the PCA. In Panel B, we compare these four statistic averages across the 12 countries (Actual) with those obtained in five simulated sets (Simulated). For each simulation, we draw $b_{ik}$ and $\sigma^2_{\psi_i}$ from their respective distributions for 12 hypothetical countries, and compute the four statistic averages. The results indicate that the values obtained with the simulations closely reproduce the salient features of the data.

[please insert Table 2 here]

III. Analysis of the Results

A. The Statistical and Economic Significance of Predictability

We fix different values for $R^2$ and $\eta$ and compute the Sharpe ratio, the skewness, and the kurtosis implied by these values. The in-sample predictive power documented in early studies (e.g., Keim and Stambaugh, 1986; Fama and French, 1988, 1989) is likely to overestimate the out-of-sample $R^2$ achievable by the investor for several reasons. First, the predicted values are affected by estimation risk. In addition to their high standard deviation, the estimated coefficients of the predictive regression are biased in small samples (e.g., Goetzmann and Jorion, 1993), and are affected by structural breaks (Paye and Timmermann, 2006). Another reason is that the investor faces specification uncertainty. Since there is no theoretical ground to justify the choice of the predictive model, it is difficult to detect the ones which will yield a positive $R^2$ in the future (Barras, 2007; Cooper and Gulen, 2006). These problems are clearly illustrated by Bossaerts and Hillion (1999) for developed market equity index monthly returns. During the 90's, they find that the average in-sample predictive power equal to 5% drops to an out-of-sample value of 1.7%. In addition, the out-of-sample $R^2$ is inferior to 1% in more than 40% of the cases. These poor forecasts are also documented by Goyal and Welch (2003, 2006) over a longer period. In light of these results, we select levels of out-of-sample $R^2$ close to zero: 0.5%, 1%, and 1.5%. We consider three sources of predictability: pure asset-timing, $Tim$ ($\eta = 0$), mixed asset-timing and factor-selectivity, $Mix$ ($\eta = 0.5$), and pure factor-selectivity, $Sel$ ($\eta = 1$). For a given simulation, the performance measures (i.e., the Sharpe ratio, the skewness, and the kurtosis) are conditioned on the values taken by $\mu, V, \Omega,$ and $\Sigma$. To compute their unconditional values, we run 100 different simulations, where $R^2$ and $\eta$ are left unchanged, but the four moments $\mu, V, \Omega,$ and $\Sigma$ are recomputed at each trial. Then, we simply take the average of each measure across these 100 repetitions (in total, we simulate 1,000,000 return
observations (10,000-100)). We set the investor’s risk aversion coefficient $A$ equal to 5, but the results obtained with other values are similar\(^7\).

### A.1. Unconstrained Case

We first consider an investment universe composed of one risky asset and the risk-free asset. In Panel A of Table 3, we display the (annualized) Sharpe ratio (along with the excess mean and standard deviation) of the conditional and unconditional strategies, as well as their skewness and kurtosis. With only one risky asset, $\eta$ has no impact on performance as $\Omega$ reduces to a single element. Since the columns $Mix$ and $Sel$ are identical to $Tim$, we leave them empty. We observe that a negligible $R^2$ of 0.5% leads to a 15%-increase in the annualized Sharpe ratio compared with the unconditional strategy (from 0.41 to 0.47). Further, a level of $R^2$ equal to 1.5% boosts the Sharpe ratio by 41% (from 0.41 to 0.58). In both cases, this increase comes along with a rise in both the excess mean and the standard deviation. We also see that small levels of $R^2$ substantially affect the skewness of the conditional strategy\(^8\). For instance, a level of $R^2$ equal to 1.5% increases the skewness up to 0.75. This level is similar to the one observed for some hedge fund directional strategies, such as global macro or managed futures (Lhabitant and Learned, 2002). This result strengthens the economic significance of predictability, since risk-averse investors have a preference for positive skewness. On the negative side, we also see that the kurtosis rises with $R^2$, which tends to increase the investor’s exposure to extreme losses.

[please insert Table 3 here]

How does the performance of the conditional strategy change when the investor has a small predictive power across several assets, instead of a single one? To address this issue, we form multi-asset strategies based on five risky assets and the risk-free asset. The results in Panel B of Table 3 show that the response of the Sharpe ratio to small levels of $R^2$ is greatly affected by the source of predictability. Under $Tim$, the relative increase compared with the unconditional strategy is equal to 25% (from 0.52 to 0.65) when $R^2 = 0.5\%$, and 63% (from 0.52 to 0.85) when $R^2 = 1.5\%$. By contrast, under $Sel$, the Sharpe ratio experiences a massive increase of 102% (from 0.52 to 1.05) when $R^2 = 0.5\%$, and 213% (from 0.52 to 1.63) when $R^2 = 1.5\%$. The relation between $R^2$ and the skewness of the conditional strategy is always positive, regardless of the source of predictability. While it is approximately equal to 0.50 when $R^2 = 1.5\%$ under $Tim$ and $Mix$, it is slightly higher under $Sel$ (0.77). We also observe that the inclusion of

\(^7\)A decrease in $A$ increases the excess mean and the standard deviation, but leaves the Sharpe ratio unchanged. Since the skewness and the kurtosis are scaled by $\sigma^2$ and $\sigma^4$, they are also not sensitive to changes in $A$.

\(^8\)The reason for this positive skewness is that the portfolio is tilted at each rebalancing time toward assets with high conditional expected returns, thus increasing the probability of observing high positive returns.
multiple assets greatly reduces the kurtosis of the conditional strategy. While the kurtosis amounts to 7.90 in the single-asset case when $R^2 = 1.5\%$, its maximum value with five risky assets is equal to 5.48 (under Tim).

In Figure 1, we plot the (annualized) additional expected return provided by the conditional strategy for the same level of standard deviation as the unconditional strategy. If the predictive model yields a level of $R^2$ equal to 0.5\% for a single asset, the additional gain amounts to 0.5\% per year. If the investor has the same predictive ability for five risky assets (instead of only one), the gains become substantial. The annual return difference is equal to 1.3\% under Tim, 3.7\% under Mix, and 5.6\% under Sel. In addition, the slopes of the curves are steep, especially under Mix and Sel. Therefore, a small rise in $R^2$ from 0.5\% to 1.5\% increases this gain by 4.6\% per year under Mix, and 6.1\% per year under Sel. These results clearly show that levels of predictability considered as negligible from a statistical standpoint are sufficient to generate large economic gains.

Figure 2 compares the distributions of the conditional and unconditional strategies when $R^2 = 0.5\%$. The distributions with a single asset are shown in Panel A, while those with five assets under Tim, Mix, and Sel are displayed in Panels B, C, and D, respectively. In all cases, the distribution of the conditional strategy is clearly asymmetric. These graphs also confirm the dramatic impact of predictability. While the probability of observing extreme returns higher than 10\% per month is null for the unconditional strategy, a negligible $R^2$ of 0.5\% rises this probability to 5.3\% under Mix, and 8.9\% under Sel.

A.2. Explaining the Impact of the Source of Predictability on Performance

In order to explain the impact of $\eta$, we use a simple example with two risky assets. Since $R^2$ is very low, we further assume that $\Sigma$ is equal to $V$, implying that:

$$
\Sigma^{-1} \approx V^{-1} = \frac{1}{D} \begin{bmatrix}
\sigma_2^2 & -\sigma_{12} \\
-\sigma_{12} & \sigma_1^2
\end{bmatrix} = \begin{bmatrix}
\varphi_{11} & \varphi_{12} \\
\varphi_{12} & \varphi_{22}
\end{bmatrix},
$$

(15)

where $D$ denotes the determinant of $V$. As shown in Table 2, all international equity indices have a strong positive exposure to one dominant factor. This characteristic, commonly shared by assets belonging to the same asset-class, implies that the covariance $\sigma_{12}$ is positive, and
that $\varphi_{12}$ is negative. Since $w_{Ct} = (1/A) V^{-1} \Pi_t$, the $2 \times 1$ vector of weight sensitivity to $\Pi_t$, $\Delta w_{Ct}$, can be written as:

$$\Delta w_{Ct} = w_{Ct} - w_U = \frac{1}{A} V^{-1} (\Pi_t - \mu) = \frac{1}{A} V^{-1} K_t,$$

(16)

where $w_U = \frac{1}{A} V^{-1} \mu$, and $K_t = \Pi_t - \mu$. Under Sel, the predicted values are uncorrelated. For instance, the investor may think that $\pi_{1t} - \mu_1 = k_t > 0$, while his views for the second asset are neutral: $\pi_{2t} - \mu_2 = 0$. In this case, $K_t = [k_t, 0]'$, and the two elements of $\Delta w_{Ct}$ take the following form:

$$\Delta w_{Ct}^1 (Sel) = \varphi_{11} \cdot k_t + \varphi_{12} \cdot 0 = \varphi_{11} \cdot k_t > 0,$$

$$\Delta w_{Ct}^2 (Sel) = \varphi_{12} \cdot k_t + \varphi_{22} \cdot 0 = \varphi_{12} \cdot k_t < 0.$$

(17)

Under Tim, the predictions for both assets are positively correlated. For instance, suppose that $\pi_{1t} - \mu_1 = \pi_{2t} - \mu_2 = k_t$. In this case, $K_t = [k_t, k_t]'$, and $\Delta w_{Ct}^1$, $\Delta w_{Ct}^2$ are equal to:

$$\Delta w_{Ct}^1 (Tim) = (\varphi_{11} + \varphi_{12}) \cdot k_t < \Delta w_{Ct}^1 (Sel),$$

$$\Delta w_{Ct}^2 (Tim) = (\varphi_{11} + \varphi_{12}) \cdot k_t > \Delta w_{Ct}^2 (Sel).$$

(18)

Under Sel, the investor takes a large position in the first asset to benefit from the high predicted value $\pi_{1t}$. At the same time, $\Delta w_{Ct}^2$ is negative in order to hedge the risk produced by the first asset. The hedge is set up because its cost measured by $\pi_{2t} - \mu_2$ is null. Under Tim, the hedging strategy becomes costly, because $\pi_{2t} - \mu_2 = k_t > 0$. For this reason, the investor decides jointly to reduce his hedging position (in the second asset) as well as his risk exposure (in the first asset), making $w_{Ct}$ closer to the unconditional weights, $w_U$. Therefore, the weak response of $w_{Ct}$ to changes in $\Pi_t$ explains why the performance of the conditional strategy is much lower under Tim.

Determining the value of $\eta$ is a difficult question, since the true values of $\Pi_t$ cannot be observed. In developed markets, it may be argued that predictability is mostly driven by factor-timing, since these markets are well integrated. However, Ferson and Harvey (1993) find that local variables, which proxy for asset-selectivity information, still have some predictive power in integrated markets. Therefore, a value of $\eta$ between Tim and Mix seems consistent with the empirical evidence. By contrast, if the conditional strategy is based on emerging markets (Cavaglia et al., 1997; Harvey, 1994), the selectivity component of predictability is likely to be higher, as these markets are less integrated to the world capital market.
A.3. Analysis of the Portfolio Weights

In order to examine the portfolio weights implied by the conditional strategy, we determine the mean and the standard deviation of the total weight invested in risky assets. They are computed from the \( N \times Q \) matrix of portfolio weights obtained from the simulation, where \( N \) denotes the number of risky assets and \( Q \) the number of draws. For each draw \( q (q = 1, \ldots, Q) \), we sum up the \( q^{th} \) column-elements of the matrix. Then, we simply compute the mean and standard deviation of this sum across all \( q \). Further, we decompose the mean weight into the average long and short positions invested in risky assets. The results for the single-asset strategy are shown in Panel A of Table 4. The unconditional strategy splits wealth almost equally into the risky and the riskfree asset. As \( R^2 \) rises, the active investor keeps on average the same weight invested in the risky asset (46%). However, he can substantially modify his portfolio according to the forecast \( \pi_{it} \). When \( R^2 = 1.5\% \), the investor invests on average 51% of his wealth in the risky asset if \( \pi_{it} > \mu_i \), and takes an average short position of -5% if \( \pi_{it} \) is lower than zero. A further evidence of these active bets is given by the weight standard deviation, as it rises up to 49% when \( R^2 = 1.5\% \).

In the multi-asset framework examined in Panel B, we observe that the unconditional strategy invests 69% of the wealth in the five risky assets (the long position of 71% being partly financed by a short position of 2%), and 31% in the riskfree asset. While the mean weight remains constant regardless of \( R^2 \), the conditional strategy produces important variations in the long and short positions according to the source of predictability. Under \( Sel \), a small \( R^2 \) of 0.5% generates on average a long position of 230%. The small weight standard deviation (41%) in comparison with this long position indicates that the latter is not financed by borrowing at the riskfree rate, but rather by short positions taken in other risky assets (-161%). On the contrary under \( Tim \), the investor does not take large opposite bets across the risky assets. This behaviour is consistent with Equations (17) and (18).

Table 4 also shows the great impact of small changes in \( R^2 \) on the spread between the average long and short positions, especially under \( Mix \) and \( Sel \). A common explanation to the observation of these extreme portfolio weights is the estimation errors contained in the expected return vector (Jobson and Korkie, 1981). This is not the case here since the returns of the conditional strategy are simulated using levels of out-of-sample \( R^2 \) which are truly equal to 0.5%,
1%, or 1.5%\(^9\). Weight variability is therefore an inherent feature of the conditional strategy, and highlights the need of using futures markets to reduce transaction costs. Since regulatory reasons may prevent the investor from taking these large long and short positions, we examine below the relation between \(R^2\) and performance when short-selling is prohibited.

A.4. No Short-Selling Case

Panel A of Table 5 shows the relation between \(R^2\) and the performance of the single-asset strategy under short-selling constraints. Compared with the unconstrained case (shown in Table 3), there is little difference in the Sharpe ratio. As \(R^2\) rises, the excess mean and standard deviation both decrease, leaving the Sharpe ratio nearly unchanged. Further, the relation between \(R^2\) and the skewness is also similar to the unconstrained case. On the contrary, we observe a strong decrease in the kurtosis, as its level declines from 7.91 (in the unconstrained case) to 6.27 when \(R^2 = 1.5\%\).

The results obtained with five risky assets are displayed in Panel B. The impact of the short-selling constraints on performance logically depends on the total short position taken in the unconstrained case. Under Tim, the relation between \(R^2\) and the Sharpe ratio is slightly affected, because these short positions are moderate. On the contrary, the reduction in performance is important under Sel. The Sharpe ratio rises by 44% (from 0.52 to 0.75) when \(R^2 = 0.5\%\), and by 84% (from 0.52 to 0.96) when \(R^2 = 1.5\%\), as opposed to 102% and 213% in the unconstrained case. Imposing short-selling constraints also reduces the skewness and the kurtosis, especially under Mix and Sel. For instance, when \(R^2 = 0.5\%\), the skewness under Sel declines from 0.36 to 0.08, while the kurtosis goes down from 4.35 to 3.34.

In Figure 3, we plot the (annualized) additional expected return provided by the constrained conditional strategy for the same level of standard deviation as the unconditional strategy. Compared with the unconstrained case, we see that the impact of the source of predictability on performance is greatly reduced. But even though the curves become flatter, the economic significance of negligible levels of predictability remains important. For instance, when \(R^2 = 1.5\%\), the annual return difference is equal to 2.4% under Tim, 3.8% under Mix, and 4.6% under Sel.

\(^9\)Stated differently, we account for estimation risk in the sense that we purposely choose levels of out-of-sample \(R^2\) close to zero. But once a given level of \(R^2\) is chosen, the predicted values produced by our model truly achieve this level of explanatory power.
Figure 4 compares the distributions of the constrained conditional and unconditional strategies when $R^2 = 0.5\%$. The distribution of the single-asset conditional strategy shown in Panel A is close to its unconstrained counterpart. It is peaked around zero, and presents a slight asymmetry. The distributions of the conditional strategy with multiple assets under $Tim$, $Mix$, and $Sel$ are displayed in Panels B, C, and D, respectively. Compared with the unconstrained case, the asymmetry, as well as the thickness of the tails are less pronounced, making these distributions closer to normality.

[please insert Figure 4 here]

B. The Trade-off between Breadth and Skill

It may be argued that it takes more time to generate a given level of $R^2$ for five assets rather than for a single one. The investor may be tempted to follow a fewer number of assets in order to increase the forecast accuracy on the remaining ones. Stated differently, there is a trade-off between breadth (the number of assets) and skill (the level of out-of-sample $R^2$). In order to examine this trade-off, we determine the level of $R^2$ for one single asset necessary to yield the same performance as the five-asset conditional strategy. We measure performance with the Sharpe ratio differential, $SD$, between the conditional and unconditional strategies for a given level of $R^2$ for the five assets. This relation shown in Figure 6 is represented for different values of $R^2$ for the five assets ranging from 0\% to 3\% under the three sources of predictability ($Tim$, $Mix$, and $Sel$).

[please insert Figure 5 here]

In the unconstrained case (Panel A), we see that the required $R^2$ for one asset has to be very high to achieve the same $SD$ as the five-asset unconstrained strategy under $Mix$ and $Sel$. For instance, when $R^2 = 1.5\%$ for the five assets, the required $R^2$ for the single asset is equal to 7.8\% under $Mix$ and 14.6\% under $Sel$. These levels are much higher than the cumulated $R^2$ across the five assets equal to 7\% (1.5\% · 5), and are extremely difficult to reach. It implies that when predictability is partly due to asset-selectivity, breadth is highly rewarded in comparison with skill. By contrast, under $Tim$, the required $R^2$ for the single asset is much lower, as it amounts to 4.1\% when the level of $R^2$ for the five assets equals 1.5\%.

In the no-short selling case (Panel B), the results under $Tim$ remain similar to the unconstrained case. However, the trade-off between breadth and skill greatly changes under $Mix$.

\[10\] The terms "breadth" and "skill" are proposed by Grinold and Khan (1999, ch. 5). They show that the portfolio information ratio is a function of breadth (the number of independent bets) and skill (the correlation between the forecasts and the returns).
and Sel. For instance, when $R^2 = 1.5\%$ for the five assets, the required $R^2$ for one asset decreases from 7.8\% to 5.4\%, and from 14.6\% to 6.8\%. Although these figures indicate a decline in the impact of breadth on performance, the investor focusing on a single asset still has to reach a sizeable level of $R^2$ in order to produce the same $SD$ as the five-asset strategy.

C. The Dybvig and Ross Critique against the Sharpe Ratio

It is commonly asserted that the Sharpe ratio differential, $SD$, between the conditional and unconditional strategies is not a reliable performance measure, as it can be negative for conditional strategies based on valuable predictive information. The theoretical justification proposed by Dybvig and Ross (1985) can be summarized as follows. If $\varepsilon_{t+1}$ is normally distributed (as it is assumed in this paper), the optimization in Equation (7) is equivalent to the maximization of the exponential utility function conditionally on the information set $I_t$:

$$\max_{\Pi_t} \Pi_t \Pi_t - \frac{1}{2} A \cdot w_{Ct} \Sigma w_{Ct} \Leftrightarrow \max_{w_{Ct}} E \left( -e^{-bW_{t+1}} \left| I_t \right. \right),$$

(19)

where $W_{t+1}$ denotes the investor’s wealth at time $t + 1$, and $A = b \cdot W_t$. This portfolio is located on the conditional mean-variance frontier, defined as the set of returns $R_{t+1}$ that maximize $E(R_{t+1} \mid I_t)$ given $\text{var}(R_{t+1} \mid I_t)$. However, this portfolio is not on the unconditional mean-variance frontier. This frontier contains the set of returns that maximize $E(R_{t+1})$ given $\text{var}(R_{t+1})$, allowing for dynamic strategies based on $I_t$ (Cochrane, 2004, ch. 8). The reason for this result is that the investor with an exponential utility function is ready to trade a lower Sharpe ratio for a higher skewness. As a result, maximizing Equation (7) is not equivalent to maximizing the Sharpe ratio, and may even generate a negative $SD^{11}$.

Considering a single risky asset $j$, Dybvig and Ross (1985) show that a sufficient condition to obtain a negative relation between $R^2$ and $SD$ is that $\mu_j > \sigma_j$. From an empirical point of view, this condition is extremely stringent, as none of the countries examined in Table 2 satisfy this condition. However, we cannot base our conclusions exclusively on the empirical validation of this condition, because the latter is only sufficient (and not necessary), and does not apply to multiple risky assets. Further, Bernardo and Ledoit (2000) show that positively skewed distribution can yield low Sharpe ratios$^{12}$. It may be the case that, for a sufficiently high $R^2$, the positive skewness produced by the conditional asset allocation goes together with a low $SD$.

---

$^{11}$Ferson and Siegel (2001) show that maximizing the Sharpe ratio is equivalent to maximizing the quadratic utility function conditionally on $I_t$ (instead of the exponential one):

$$\max_{w_{Ct}} E(W_{t+1} - \frac{A}{2} \cdot W_{t+1}^2 \mid I_t).$$

$^{12}$To illustrate this problem, they consider a lottery ticket costing one cent and giving 50 billions with a 10\% probability and zero otherwise. This ticket has a skewness of 2.6, but a low Sharpe ratio of 0.33.
To address these issues, we compute the Sharpe ratio along with the skewness of the conditional strategy for levels of $R^2$ equal to 3%, 5%, and 7%. These values represent extreme cases probably well above the true achievable out-of-sample predictability, but provide a robust test of the relation between $R^2$ and the Sharpe ratio. We consider in turn an investment universe of one, five, and ten risky assets under the three sources of predictability Tim, Mix, and Sel. The results in the unconstrained and no-short selling cases are shown in Panels A and B of Table 6. We see that the relation between $R^2$ and the Sharpe ratio is always positive. Further, the conditional strategy is not subject to the skewness problem raised by Bernardo and Ledoit (2000). Although a higher $R^2$ comes together with a higher skewness, it does not lead to a reduction in the Sharpe ratio\textsuperscript{13}. Therefore, the Sharpe ratio is a robust indicator of the performance of the conditional asset allocation.

\[\text{please insert Table 6 here}\]

D. Sensitivity Analysis to Fat-Tail Distributions

Levy and Duchin (2004) find that the empirical distribution of US stock and bond monthly returns is symmetric, but presents fat tails\textsuperscript{14}. The presence of fat tails leaves the portfolio weights and the Sharpe ratio of the mean-variance conditional strategy unchanged, as their values only depend on the asset return first and second moments. However, it can substantially affect its kurtosis and expose the investor to the risk of extreme losses. To address this issue, we model asset return distributions with fat tails. We draw the kurtosis of each asset unpredictable term, $\varepsilon_{it+1}$, from the following uniform distribution: $\text{Kurt}_{\varepsilon_i} \sim U [5, 6]$\textsuperscript{15}. To reproduce these levels of kurtosis in the simulations, we use the framework of Bekaert et al. (1998), and model the $N \times 1$ vector of unpredictable terms, $\varepsilon_{t+1}$, as a mixture of normal distributions\textsuperscript{16}. The details of our approach is given in the appendix. In Figure 6, we plot the difference between the additional kurtosis produced by the conditional strategy over the unconditional one when the asset returns have fat versus normal tails. This difference is computed for one and five

\textsuperscript{13}We also run the same test using developed market government bond indices. The results, available upon request, remain unchanged.

\textsuperscript{14}Excess kurtosis (compared with the normal distribution) is also found in emerging market indices, on which the conditional strategy can also be applied (Cavaglia et al., 1997; Harvey, 1994).

\textsuperscript{15}The upper bound is based on the estimates of Bekaert et al. (1998) for emerging market equity indices. They find that in more than 70% of the cases, the kurtosis is lower than 6 during the 80’s and the 90’s.

\textsuperscript{16}In a single-asset framework, Xu (2004) proposes another way to model fat-tails based on the student-$t$ distribution. The main issue with this approach is that there is an automatic link between the standard deviation and the kurtosis (through the degree of freedom parameter), which is not necessarily observed empirically.
risky assets under different levels of $R^2$ and the three sources of predictability $Tim$, $Mix$, and $Sel$.

In the unconstrained case (Panel A), the additional kurtosis produced by the conditional strategy over the unconditional one under fat tails is important for the single-asset strategy. It increases quickly as $R^2$ goes from 0% to 1%, and remains constant afterwards. A similar pattern is observed with five assets under $Tim$, although the increase is less pronounced. By contrast, we observe a slight decrease compared with the normal case under $Mix$ and $Sel$. One possible explanation for this result is the diversification effect on kurtosis produced by the opposite positions implied by selectivity signals. This kurtosis diversification effect is also observed in the hedge fund industry, as Lhabitant and Learned (2002) find that including an increasing number of hedge funds in the portfolio greatly reduces its kurtosis.

When short-selling constraints are imposed (Panel B), the additional kurtosis is moderate, and fairly similar regardless of the number of assets. To summarize our results, the investor’s exposure to extreme losses under fat-tail distributions is rather limited. For example, consider an increase in kurtosis equal to 2 for the conditional strategy under $Tim$ (the highest additional kurtosis under $Tim$ observed in Figure 6 is equal to 1.59). Compared with a normal distribution with the same excess mean and standard deviation as the conditional strategy (i.e., 8.6% and 13.2%, as in Table 3)), the Value-at-Risk at 5% would only rise by only 11.0%.

Conclusion

We examine the relation between the statistical and economic significance of predictability for levels of out-of-sample predictability $R^2$ close to zero. With an investment universe composed of equity indices, we find that an out-of-sample $R^2$ as small as 0.5% is sufficient to largely outperform the unconditional strategy. In addition, we observe that the performance for a given $R^2$ is magnified when the source of predictability comes from asset-selectivity. Contrary to factor-timing, the forecasts under asset-selectivity are uncorrelated, thus allowing the investor to take larger positions in response to predictive information. Regulatory constraints may prevent the investor from taking the large positions implied by the conditional strategy. To address this issue, we also examine the performance of the conditional strategy subject to short-selling constraints. Although the positive relation between $R^2$ and performance is less pronounced, the economic gains over the unconditional strategy remain positive. While our main focus is on equity indices, all of these results remain unchanged if the investment universe is formed with bond indices.
A further application of our approach consists in examining the trade-off between the number of assets and the achievable level of $R^2$. Under asset-selectivity, we find that the required $R^2$ for one asset has to be very high to achieve the same performance as the multi-asset conditional strategy. As a result, it is easier to improve performance by increasing the number of assets rather than the forecast accuracy for a single asset. By contrast, the reward for increasing the number of assets is less pronounced under factor-timing. Finally, using a wide range of values for $R^2$ in both single- and multi-asset strategies, we also find that the response of the Sharpe ratio to increase in $R^2$ is always positive. It implies that the Sharpe ratio is a reliable performance indicator, and that the condition underlying the theoretical argument of Dybvig and Ross (1985) is too stringent to have empirical implications.

The out-of-sample $R^2$ is the driving factor of the performance produced by the conditional strategy. Since achieving a positive $R^2$ is very difficult because of estimation risk and specification uncertainty, implementing conditional strategies seems to be a daunting task. Our paper contradicts this assertion, and conveys a positive message for active investors. It reveals that extremely small levels of predictability are sufficient to implement profitable conditional strategies. In particular, an investor with very limited selectivity information can still obtain substantial economic gains.
Appendix

Derivation of the Sharpe Ratio of the Conditional Strategy

The unconditional excess mean of the conditional strategy, $\mu_C$, can be derived by the law of iterated expectation (using the information set $I_t$) and the properties of the trace operator:

$$
\mu_C = E \left( E \left( \frac{1}{A} \Pi_1 \Sigma^{-1} r_{t+1} \bigg| I_t \right) \right) = \frac{1}{A} E \left( E \left( \Pi_1 \Sigma^{-1} \Pi_t \bigg| I_t \right) \right)
$$

$$
= \frac{1}{A} Tr \left( \Sigma^{-1} E \left( \Pi_t \Pi_t^\prime \right) \right) = \frac{1}{A} \left[ Tr \left( \Sigma^{-1} \Omega \right) + \mu \Sigma^{-1} \mu \right] \tag{20}
$$

In order to derive the unconditional variance of the conditional strategy, $\sigma^2_C$, we decompose the variance into two parts:

$$
\sigma^2_C = \frac{1}{A^2} \left[ Var \left( E \left( \Pi_1 \Sigma^{-1} r_{t+1} \bigg| I_t \right) \right) + E \left( Var \left( \Pi_1 \Sigma^{-1} r_{t+1} \bigg| I_t \right) \right) \right]
$$

$$
= \frac{1}{A^2} Var \left( \Pi_1 \Sigma^{-1} \Pi_t \right) + \frac{1}{A^2} E \left( \Pi_1 \Sigma^{-1} \Pi_t \right) \tag{21}
$$

The term $Var \left( \Pi_1 \Sigma^{-1} \Pi_t \right)$ corresponds to the variance of a quadratic form in normal variables. Magnus and Neudecker (2001) show that:

$$
Var \left( \Pi_1 \Sigma^{-1} \Pi_t \right) = 2 Tr(\Omega(\Sigma^{-1}\Omega)^2) + 4 \mu \Sigma^{-1} \Omega \Sigma^{-1} \mu
$$

Using Equation (20), we know that:

$$
E \left( \Pi_1 \Sigma^{-1} \Pi_t \right) = Tr \left( \Sigma^{-1} \Omega \right) + \mu \Sigma^{-1} \mu \tag{23}
$$

Using Equations (20), (22), and (23), the Sharpe ratio is equal to:

$$
S_C = \frac{Tr \left( \Sigma^{-1} \Omega \right) + \mu \Sigma^{-1} \mu}{\left[ 2 Tr(\Omega(\Sigma^{-1}\Omega)^2) + 4 \mu \Sigma^{-1} \Omega \Sigma^{-1} \mu + Tr \left( \Sigma^{-1} \Omega \right) + \mu \Sigma^{-1} \mu \right]^{\frac{1}{2}}} \tag{24}
$$
Determination of the Covariance Matrix $\Omega$ of the Predicted Values

Using Equation (5), the asset $i$ predicted value under factor-timing, $\pi_{it}^T$, is equal to $b_i^T \pi_{ft}$. It implies that $\sigma^2_{\pi_{it}}$ and $cov\left(\pi_{it}^T, \pi_{jt}^T\right)$ can be written as:

$$
\begin{align*}
\sigma^2_{\pi_{it}} &= b_i^T V_{\pi f} b_i, \\
cov\left(\pi_{it}^T, \pi_{jt}^T\right) &= b_i^T V_{\pi f} b_j
\end{align*}
$$

where $V_{\pi f}$ denotes the $K \times K$ diagonal covariance matrix of the factor predicted values. We assume that the explanatory power of the factors, denoted by $R^2_f$, is constant across the $K$ factors. Using the definition of $R^2_f$, we have $V_{\pi f} = R^2_f V_f$, where $V_f$ is the $K \times K$ diagonal factor covariance matrix. By replacing $V_{\pi f}$ by $R^2_f V_f$ in Equation (25), we obtain:

$$
\begin{align*}
\sigma^2_{\pi_{it}} &= R^2_f b_i^T V_f b_i, \\
cov\left(\pi_{it}^T, \pi_{jt}^T\right) &= R^2_f b_i^T V_f b_j
\end{align*}
$$

(26)

$R^2_f$ must be compatible with the level of predictability $R^2$ initially defined in Equation (3):

$$
R^2 = \frac{1}{N} \sum_{i=1}^{N} R^2_{fi} = \frac{1}{N} \sum_{i=1}^{N} \sigma^2_{\pi_{it}} \Leftrightarrow R^2_f (R^2) = N \cdot R^2 \left( \sum_{i=1}^{N} \frac{b_i^T V_f b_i}{\sigma^2_{\pi_{it}}} \right)^{-1}
$$

(27)

Plugging $R^2_f (R^2)$ into Equation (26), we can determine $\sigma^2_{\pi_{it}} = \sigma^2_{\pi_{it}}$, as well as $cov\left(\pi_{it}, \pi_{jt}\right) = (1 - \eta) \cdot cov\left(\pi_{it}, \pi_{jt}\right)$. After collecting terms, we can write $\Omega$ as:

$$
\Omega (R^2, \eta) = (1 - \eta) \cdot R^2_f (R^2) \cdot \beta V_f \beta' + \eta V_{\pi}
$$

(28)

where $V_{\pi}$ denotes the $N \times N$ diagonal matrix containing the elements $\sigma^2_{\pi_{it}}$.

Modelling Fat-Tail Distributions

We define the $N \times 1$ vector of unpredictable terms, $\varepsilon_{t+1}$, as a mixture of normal distributions:

$$
\varepsilon_{t+1} = p \varepsilon_{t+1}^{(1)} + (1 - p) \varepsilon_{t+1}^{(2)}
$$

(29)

where $\varepsilon_{t+1}^{(1)} \sim N\left(0, \Sigma^{(1)}\right)$ and $\varepsilon_{t+1}^{(2)} \sim N\left(0, \Sigma^{(2)}\right)$. For each asset $i$ ($i = 1, \ldots, N$), we set $\varepsilon_{it+1}$ equal to $p \varepsilon_{it+1}^{(1)} + (1 - p) \varepsilon_{it+1}^{(2)}$, where $\varepsilon_{it+1}^{(1)} \sim N\left(0, \sigma_{\varepsilon_{it+1}^{(1)}}\right)$ and $\varepsilon_{it+1}^{(2)} \sim N\left(0, \sigma_{\varepsilon_{it+1}^{(2)}}\right)$. The relations
between the (known) parameters, $\sigma^2_{\epsilon_i}$ and $Kurt_{\epsilon_i}$, and the (unknown) variances under the two regimes, $\sigma^2_{\epsilon_i}^{(1)}$ and $\sigma^2_{\epsilon_i}^{(2)}$, are given by the following equations (see Bekaert et al., 1998):

$$\sigma^2_{\epsilon_i} = p\sigma^2_{\epsilon_i}^{(1)} + (1 - p)\sigma^2_{\epsilon_i}^{(2)},$$  
$$Kurt_{\epsilon_i} = p\left(\sigma^2_{\epsilon_i}^{(1)} + 3\sigma^4_{\epsilon_i}^{(1)}\right) + (1 - p)\left(\sigma^2_{\epsilon_i}^{(2)} + 3\sigma^4_{\epsilon_i}^{(2)}\right). \quad (30)$$

After setting arbitrarily $p$ equal to 0.5, we solve this system of two equations to determine the parameters $\sigma^2_{\epsilon_i}^{(1)}$ and $\sigma^2_{\epsilon_i}^{(2)}$. Then, we stack all standard deviations, $\sigma_{\epsilon_i}^{(r)}$, under a given regime $r$ ($r = 1, 2$) to form a $N \times 1$ vector $\sigma_{\epsilon_i}^{(r)} = \left[\sigma_{\epsilon_1}^{(r)}, ..., \sigma_{\epsilon_N}^{(r)}\right]'$. We make the additional assumption that the correlation between each pair of assets $i$ and $j$ under each regime (denoted by $c_{ij}$) is constant. $c_{ij}$ is defined such that it is compatible with the unconditional correlation $\rho_{ij}$:

$$\rho_{ij} = \frac{\text{cov} \left(\epsilon_{it+1}, \epsilon_{jt+1}\right)}{\sigma_{\epsilon_i}\sigma_{\epsilon_j}} = \frac{pc_{ij}\sigma_{\epsilon_i}^{(1)}\sigma_{\epsilon_j}^{(1)} + (1 - p)c_{ij}\sigma_{\epsilon_i}^{(2)}\sigma_{\epsilon_j}^{(2)}}{\sigma_{\epsilon_i}\sigma_{\epsilon_j}}. \quad (31)$$

Using these coefficients $c_{ij}$, the covariance matrix under the regime $r$, $\Sigma^{(r)}$, is given by:

$$\Sigma^{(r)} = K^{(r)} \odot C, \quad (32)$$

where $K^{(r)} = \sigma_{\epsilon_i}^{(r)}\sigma_{\epsilon_i}^{(r)}$, $C$ is the $N \times N$ matrix containing the elements $c_{ij}$, and $\odot$ denotes the Hadamard product (corresponding to an element-by-element multiplication). For each iteration $q$ ($q = 1, ..., Q$), we draw one realization of $\Pi_t$, and one realization from the variable $z \sim U [0, 1]$. If $z < p$, we draw the unpredictable term from the distribution under the first regime. Otherwise, it is drawn from the distribution under the second regime. The excess return of the conditional strategy is computed as: $r_{Ct+1} = u_t'\left(\Pi_t + \epsilon_{it+1}^{(r)}\right)$. By repeating this procedure $Q$ times, we obtain a $Q \times 1$ vector of realizations of $r_{Ct+1}$, from which we can compute all performance measures ($Q$ is set to 10,000).
References


Han, Y., 2005, Can an Investor Profit from Return Predictability in Real Time? Working Paper, Tulane University


Table 1
Estimated Parameters of the Factor Structure

The parameter values are computed with a Principal Component Analysis (PCA) applied to the monthly excess returns (in local currency) of 12 developed market equity indices. The PCA is run over the full period from January 1980 to December 2005 (312 observations), as well as over two subperiods of equal length (January 1980-December 1992 and January 1993-December 2005). The average explanatory power denotes the average proportion of the country variance explained by one, two, and three factors, respectively. The factor excess mean and standard deviation are annualized. For each of the three factors, the average beta and its standard deviation are equal to the cross-sectional average and standard deviation of the betas across the 12 countries.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Average explanatory power</td>
<td>1 factor</td>
<td>63.7%</td>
<td>57.3%</td>
</tr>
<tr>
<td></td>
<td>2 factors</td>
<td>70.4%</td>
<td>64.2%</td>
</tr>
<tr>
<td></td>
<td>3 factors</td>
<td>76.6%</td>
<td>70.7%</td>
</tr>
<tr>
<td>Factor excess mean</td>
<td>$\mu_{f1}$</td>
<td>7.4%</td>
<td>5.8%</td>
</tr>
<tr>
<td>(Annualized)</td>
<td>$\mu_{f2}$</td>
<td>3.8%</td>
<td>-0.3%</td>
</tr>
<tr>
<td></td>
<td>$\mu_{f3}$</td>
<td>-4.0%</td>
<td>21.2%</td>
</tr>
<tr>
<td>Factor standard deviation</td>
<td>$\sigma_{f1}$</td>
<td>13.8%</td>
<td>13.4%</td>
</tr>
<tr>
<td>(Annualized)</td>
<td>$\sigma_{f2}$</td>
<td>43.4%</td>
<td>74.2%</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{f3}$</td>
<td>63.3%</td>
<td>94.7%</td>
</tr>
<tr>
<td>Average beta</td>
<td>$\mu_{b1}$</td>
<td>0.967</td>
<td>0.969</td>
</tr>
<tr>
<td></td>
<td>$\mu_{b2}$</td>
<td>0.017</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>$\mu_{b3}$</td>
<td>0.005</td>
<td>0.003</td>
</tr>
<tr>
<td>Standard deviation of beta</td>
<td>$\sigma_{b1}$</td>
<td>0.185</td>
<td>0.180</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{b2}$</td>
<td>0.138</td>
<td>0.094</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{b3}$</td>
<td>0.079</td>
<td>0.059</td>
</tr>
</tbody>
</table>
Table 2
Comparison of Actual and Simulated Country Equity Index Excess Returns

Panel A contains descriptive statistics of the monthly excess returns (in local currency) of 12 developed market equity indices between January 1980 and December 2005 (312 observations). The excess mean and standard deviation are annualized. Cross-correlation is the country average correlation with the other countries. Explanatory power denotes the proportion of the country variance explained by the three-factor model from the Principal Component Analysis. In Panel B, we compare the statistic averages across the 12 countries (Actual) with those obtained in five simulated sets of 12 hypothetical countries (Simulated). The parameters used for the simulation are shown in the first column of Table I.

Panel A: Country Index Excess Returns

<table>
<thead>
<tr>
<th></th>
<th>Excess mean</th>
<th>Std deviation</th>
<th>Cross-correlation</th>
<th>Explanatory Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>5.6%</td>
<td>18.3%</td>
<td>0.40</td>
<td>83.9%</td>
</tr>
<tr>
<td>Belgium</td>
<td>7.6%</td>
<td>17.2%</td>
<td>0.47</td>
<td>66.8%</td>
</tr>
<tr>
<td>Canada</td>
<td>4.9%</td>
<td>15.5%</td>
<td>0.49</td>
<td>71.7%</td>
</tr>
<tr>
<td>Denmark</td>
<td>9.2%</td>
<td>18.5%</td>
<td>0.43</td>
<td>54.1%</td>
</tr>
<tr>
<td>France</td>
<td>8.0%</td>
<td>20.3%</td>
<td>0.52</td>
<td>71.3%</td>
</tr>
<tr>
<td>Germany</td>
<td>6.2%</td>
<td>18.6%</td>
<td>0.53</td>
<td>77.7%</td>
</tr>
<tr>
<td>Italy</td>
<td>7.9%</td>
<td>25.3%</td>
<td>0.39</td>
<td>99.3%</td>
</tr>
<tr>
<td>Japan</td>
<td>4.7%</td>
<td>18.6%</td>
<td>0.30</td>
<td>97.4%</td>
</tr>
<tr>
<td>Netherlands</td>
<td>9.7%</td>
<td>16.9%</td>
<td>0.58</td>
<td>81.2%</td>
</tr>
<tr>
<td>Switzerland</td>
<td>9.1%</td>
<td>16.4%</td>
<td>0.55</td>
<td>76.8%</td>
</tr>
<tr>
<td>UK</td>
<td>6.8%</td>
<td>15.9%</td>
<td>0.54</td>
<td>70.7%</td>
</tr>
<tr>
<td>USA</td>
<td>7.4%</td>
<td>14.6%</td>
<td>0.51</td>
<td>68.0%</td>
</tr>
</tbody>
</table>

Panel B: Actual versus Simulated Excess Returns

<table>
<thead>
<tr>
<th>Average</th>
<th>Excess mean</th>
<th>Average Std deviation</th>
<th>Average Cross-correlation</th>
<th>Average Explanatory Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td>7.3%</td>
<td>18.0%</td>
<td>0.48</td>
<td>76.6%</td>
</tr>
<tr>
<td>Simulated 1</td>
<td>7.4%</td>
<td>18.4%</td>
<td>0.54</td>
<td>70.4%</td>
</tr>
<tr>
<td>Simulated 2</td>
<td>7.5%</td>
<td>17.8%</td>
<td>0.60</td>
<td>80.0%</td>
</tr>
<tr>
<td>Simulated 3</td>
<td>7.3%</td>
<td>19.0%</td>
<td>0.53</td>
<td>72.3%</td>
</tr>
<tr>
<td>Simulated 4</td>
<td>6.9%</td>
<td>17.8%</td>
<td>0.48</td>
<td>73.1%</td>
</tr>
<tr>
<td>Simulated 5</td>
<td>7.1%</td>
<td>18.2%</td>
<td>0.53</td>
<td>72.6%</td>
</tr>
</tbody>
</table>
### Table 3
**Statistical and Economic Significance of Predictability**
**Unconstrained Case**

We compare the performance of the conditional and the unconditional (Uncond) strategies under different levels and sources of predictability, $R^2$ and $\eta$. $Tim (\eta = 0)$ denotes pure factor-timing, $Mix (\eta = 0.5)$ refers to mixed factor-timing and asset-selectivity, and $Sel (\eta = 1)$ denotes pure asset-selectivity. Panel A considers an investment universe of one risky asset, and Panel B extends the analysis to five risky assets. In both Panels, we compute the (annualized) Sharpe ratio (along with the (annualized) excess mean and standard deviation), as well as the skewness and the kurtosis of the strategies.

#### Panel A: One Risky Asset

<table>
<thead>
<tr>
<th>Source $\eta$</th>
<th>Uncond</th>
<th>$Tim (\eta = 0)$</th>
<th>$Mix (\eta = 0.5)$</th>
<th>$Sel (\eta = 1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level $R^2$</td>
<td></td>
<td>0.5% 1.0% 1.5%</td>
<td>0.5% 1.0% 1.5%</td>
<td>0.5% 1.0% 1.5%</td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.41</td>
<td>0.47 0.53 0.58</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>-Excess mean</td>
<td>3.4%</td>
<td>4.6% 5.8% 7.1%</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>-Std deviation</td>
<td>8.1%</td>
<td>9.6% 10.9% 12.1%</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.00</td>
<td>0.37 0.61 0.75</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.00</td>
<td>5.87 7.12 7.90</td>
<td>- - -</td>
<td>- - -</td>
</tr>
</tbody>
</table>

#### Panel B: Five Risky Assets

<table>
<thead>
<tr>
<th>Source $\eta$</th>
<th>Uncond</th>
<th>$Tim (\eta = 0)$</th>
<th>$Mix (\eta = 0.5)$</th>
<th>$Sel (\eta = 1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level $R^2$</td>
<td></td>
<td>0.5% 1.0% 1.5%</td>
<td>0.5% 1.0% 1.5%</td>
<td>0.5% 1.0% 1.5%</td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.52</td>
<td>0.65 0.76 0.85</td>
<td>0.88 1.12 1.32</td>
<td>1.05 1.38 1.63</td>
</tr>
<tr>
<td>-Excess mean</td>
<td>5.4%</td>
<td>8.6% 11.8% 15.1%</td>
<td>15.8% 26.4% 37.3%</td>
<td>23.1% 41.9% 61.6%</td>
</tr>
<tr>
<td>-Std deviation</td>
<td>10.4%</td>
<td>13.2% 15.6% 17.8%</td>
<td>17.9% 23.4% 28.2%</td>
<td>21.9% 30.1% 37.4%</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.00</td>
<td>0.30 0.46 0.55</td>
<td>0.25 0.38 0.50</td>
<td>0.36 0.58 0.77</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.00</td>
<td>4.67 5.25 5.48</td>
<td>3.99 4.31 4.57</td>
<td>4.35 4.85 5.31</td>
</tr>
</tbody>
</table>
We compute descriptive statistics on the total weight invested in risky assets by the conditional and unconditional (Uncond) strategies under different levels and sources of predictability, $R^2$ and $\eta$. *Tim* ($\eta = 0$) denotes pure factor-timing, *Mix* ($\eta = 0.5$) refers to mixed factor-timing and asset-selectivity, and *Sel* ($\eta = 1$) denotes pure asset-selectivity. Panel A considers an investment universe of one risky asset, and Panel B extends the analysis to five risky assets. In both Panels, Mean and Std deviation denote the average and the standard deviation of the total weight invested in risky assets at each rebalancing date. They are computed using each draw $q$ ($q = 1, \ldots, Q$) of the simulation. We further decompose the mean weight into the average total long and short positions invested in risky assets.

### Panel A: One Risky Asset

<table>
<thead>
<tr>
<th>Source $\eta$</th>
<th>Uncond</th>
<th>Tim ($\eta = 0$)</th>
<th>Mix ($\eta = 0.5$)</th>
<th>Sel ($\eta = 1$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level $R^2$</td>
<td></td>
<td>0.5%  1.0%  1.5%</td>
<td>0.5%  1.0%  1.5%</td>
<td>0.5%  1.0%  1.5%</td>
</tr>
<tr>
<td>Mean</td>
<td>46%</td>
<td>46%  46%  47%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Std deviation</td>
<td>0%</td>
<td>28%  40%  49%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Long position</td>
<td>46%</td>
<td>47%  49%  51%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Short position</td>
<td>0%</td>
<td>-1%  -3%  -5%</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

### Panel B: Five Risky Asset

<table>
<thead>
<tr>
<th>Source $\eta$</th>
<th>Uncond</th>
<th>Tim ($\eta = 0$)</th>
<th>Mix ($\eta = 0.5$)</th>
<th>Sel ($\eta = 1$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level $R^2$</td>
<td></td>
<td>0.5%  1.0%  1.5%</td>
<td>0.5%  1.0%  1.5%</td>
<td>0.5%  1.0%  1.5%</td>
</tr>
<tr>
<td>Mean</td>
<td>69%</td>
<td>70%  70%  71%</td>
<td>70%  70%  70%</td>
<td>69%  70%  70%</td>
</tr>
<tr>
<td>Std deviation</td>
<td>0%</td>
<td>38%  55%  67%</td>
<td>40%  57%  71%</td>
<td>41%  59%  74%</td>
</tr>
<tr>
<td>Long position</td>
<td>71%</td>
<td>88%  101% 113%</td>
<td>177% 236% 283%</td>
<td>230% 317% 390%</td>
</tr>
<tr>
<td>Short position</td>
<td>-2%</td>
<td>-18% -31% -43%</td>
<td>-107% -166% -213%</td>
<td>-161% -247% -320%</td>
</tr>
</tbody>
</table>
Table 5
Statistical and Economic Significance of Predictability
No-Short Selling Case

We compare the performance of the conditional and the unconditional (Uncond) strategies under different levels and sources of predictability, $R^2$ and $\eta$. $Tim(\eta = 0)$ denotes pure factor-timing, $Mix(\eta = 0.5)$ refers to mixed factor-timing and asset-selectivity, and $Sel(\eta = 1)$ denotes pure asset-selectivity. Panel A considers an investment universe of one risky asset, and Panel B extends the analysis to five risky assets. In both Panels, we compute the (annualized) Sharpe ratio (along with the (annualized) excess mean and standard deviation), as well as the skewness and the kurtosis of the strategies.

### Panel A: One Risky Asset

<table>
<thead>
<tr>
<th>Source $\eta$</th>
<th>Uncond</th>
<th>$Tim (\eta = 0)$</th>
<th>$Mix (\eta = 0.5)$</th>
<th>$Sel (\eta = 1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level $R^2$</td>
<td>-</td>
<td>0.5% 1.0% 1.5%</td>
<td>0.5% 1.0% 1.5%</td>
<td>0.5% 1.0% 1.5%</td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.41</td>
<td>0.47 0.52 0.56</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>-Excess mean</td>
<td>3.4%</td>
<td>4.5% 5.3% 6.0%</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>-Std deviation</td>
<td>8.1%</td>
<td>9.4% 10.1% 10.5%</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.00</td>
<td>0.33 0.47 0.54</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.00</td>
<td>5.53 6.14 6.27</td>
<td>- - -</td>
<td>- - -</td>
</tr>
</tbody>
</table>

### Panel B: Five Risky Assets

<table>
<thead>
<tr>
<th>Source $\eta$</th>
<th>Uncond</th>
<th>$Tim (\eta = 0)$</th>
<th>$Mix (\eta = 0.5)$</th>
<th>$Sel (\eta = 1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level $R^2$</td>
<td>-</td>
<td>0.5% 1.0% 1.5%</td>
<td>0.5% 1.0% 1.5%</td>
<td>0.5% 1.0% 1.5%</td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.52</td>
<td>0.62 0.70 0.75</td>
<td>0.71 0.80 0.88</td>
<td>0.75 0.87 0.96</td>
</tr>
<tr>
<td>-Excess mean</td>
<td>5.4%</td>
<td>7.4% 8.8% 9.7%</td>
<td>9.4% 11.2% 12.9%</td>
<td>10.6% 13.3% 15.1%</td>
</tr>
<tr>
<td>-Std deviation</td>
<td>10.4%</td>
<td>11.9% 12.5% 12.9%</td>
<td>13.3% 14.1% 14.6%</td>
<td>14.2% 15.2% 15.7%</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.00</td>
<td>0.22 0.30 0.34</td>
<td>0.14 0.18 0.20</td>
<td>0.08 0.10 0.11</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.00</td>
<td>4.16 4.35 4.44</td>
<td>3.62 3.70 3.72</td>
<td>3.34 3.37 3.38</td>
</tr>
</tbody>
</table>
We compute the (annualized) Sharpe ratio along with the skewness of the conditional strategy for high levels of $R^2$ and different sources of predictability $\eta$. $Tim (\eta = 0)$ denotes pure factor-timing, $Mix (\eta = 0.5)$ refers to mixed factor-timing and asset-selectivity, and $Sel (\eta = 1)$ denotes pure asset-selectivity. We consider in turn an investment universe of one, five, and ten risky assets. The results in the unconstrained and no-short selling cases are shown in Panel A and B, respectively.

**Table 6**

Robustness of the Relation between $R^2$ and the Sharpe ratio

<table>
<thead>
<tr>
<th>Source $\eta$</th>
<th>Uncond</th>
<th>$Tim (\eta = 0)$</th>
<th>$Mix (\eta = 0.5)$</th>
<th>$Sel (\eta = 1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level $R^2$</td>
<td></td>
<td>3% 5% 7%</td>
<td>3% 5% 7%</td>
<td>3% 5% 7%</td>
</tr>
<tr>
<td>1 Risky Asset</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.41</td>
<td>0.71 0.84 0.95</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.00</td>
<td>1.08 1.34 1.48</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>5 Risky Assets</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.52</td>
<td>1.08 1.31 1.51</td>
<td>1.75 2.16 2.48</td>
<td>2.16 2.63 3.04</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.00</td>
<td>0.73 0.90 1.06</td>
<td>0.71 0.89 1.07</td>
<td>1.12 1.42 1.68</td>
</tr>
<tr>
<td>10 Risky Assets</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.53</td>
<td>1.18 1.45 1.67</td>
<td>2.56 3.18 3.63</td>
<td>3.27 3.89 4.10</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.00</td>
<td>0.71 0.84 0.95</td>
<td>0.63 0.81 0.98</td>
<td>0.97 1.30 1.63</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source $\eta$</th>
<th>Uncond</th>
<th>$Tim (\eta = 0)$</th>
<th>$Mix (\eta = 0.5)$</th>
<th>$Sel (\eta = 1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level $R^2$</td>
<td></td>
<td>3% 5% 7%</td>
<td>3% 5% 7%</td>
<td>3% 5% 7%</td>
</tr>
<tr>
<td>1 Risky Asset</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.41</td>
<td>0.66 0.73 0.82</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.00</td>
<td>0.66 0.75 0.82</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>5 Risky Assets</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.52</td>
<td>0.90 1.03 1.14</td>
<td>1.06 1.26 1.42</td>
<td>1.17 1.41 1.55</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.00</td>
<td>0.43 0.50 0.54</td>
<td>0.23 0.24 0.27</td>
<td>0.12 0.13 0.16</td>
</tr>
<tr>
<td>10 Risky Assets</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.53</td>
<td>0.96 1.11 1.25</td>
<td>1.22 1.43 1.61</td>
<td>1.40 1.69 1.90</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.00</td>
<td>0.36 0.39 0.44</td>
<td>0.12 0.16 0.17</td>
<td>0.07 0.08 0.08</td>
</tr>
</tbody>
</table>
Figure 1
Additional Expected Return over the Unconditional Strategy
Unconstrained Case

The graph represents the (annualized) additional expected return provided by the conditional strategy for the same level of standard deviation as the unconditional strategy. This additional gain is computed for one and five risky assets under different levels and sources of predictability, $R^2$ and $\eta$. $Tim (\eta = 0)$ denotes pure factor-timing, $Mix (\eta = 0.5)$ refers to mixed factor-timing and asset-selectivity, and $Sel (\eta = 1)$ denotes pure asset-selectivity.
The different figures show the distributions of the monthly excess returns of the conditional and unconditional strategies formed with one and five risky assets, respectively. \textit{Tim} ($\eta = 0$) denotes pure factor-timing, \textit{Mix} ($\eta = 0.5$) refers to mixed factor-timing and asset-selectivity, and \textit{Sel} ($\eta = 1$) denotes pure asset-selectivity. The distributions are computed using a standard Gaussian Kernel.
Figure 3
Additional Expected Return over the Unconditional Strategy
No-Short Selling Case

The graph represents the (annualized) additional expected return provided by the conditional strategy for the same level of standard deviation as the unconditional strategy. This additional gain is computed for one and five risky assets under different levels and sources of predictability, $R^2$ and $\eta$. $Tim (\eta = 0)$ denotes pure factor-timing, $Mix (\eta = 0.5)$ refers to mixed factor-timing and asset-selectivity, and $Sel (\eta = 1)$ denotes pure asset-selectivity.
Figure 4
Distributions of the Conditional and Unconditional Strategies
No-Short Selling Case with $R^2 = 0.5\%$

The different figures show the distributions of the monthly excess returns of the conditional and unconditional strategies formed with one and five risky assets, respectively. $Tim (\eta = 0)$ denotes pure factor-timing, $Mix (\eta = 0.5)$ refers to mixed factor-timing and asset-selectivity, and $Sel (\eta = 1)$ denotes pure asset-selectivity. The distributions are computed using a standard Gaussian Kernel.
We determine the different combinations of breadth (the number of assets) and skill (the out-of-sample level of $R^2$) that produce the same performance measured by the Sharpe ratio differential between the conditional and unconditional strategies. We determine the level of $R^2$ for a single asset required to yield the same performance as the five-asset conditional strategy under different levels of $R^2$ (for the five assets), and different sources of predictability, $\eta$. $Tim$ ($\eta = 0$) denotes pure factor-timing, $Mix$ ($\eta = 0.5$) refers to mixed factor-timing and asset-selectivity, and $Sel$ ($\eta = 1$) denotes pure asset-selectivity. The results in the unconstrained and no-short selling cases are shown in Panel A and B, respectively.
We measure the difference between the additional kurtosis produced by the conditional strategy over the unconditional one when the asset returns have fat versus normal tails. This difference is computed under different levels and sources of predictability, $R^2$ and $\eta$. $Tim (\eta = 0)$ denotes pure factor-timing, $Mix (\eta = 0.5)$ refers to mixed factor-timing and asset-selectivity, and $Sel (\eta = 1)$ denotes pure asset-selectivity. The results in the unconstrained and no-short selling cases are shown in Panel A and B, respectively.