Available online at www.sciencedirect.com
ScienceDirect
Journal of Empirical Finance 14 (2007) 443 – 464
Journal of EMPIRICAL FINANCE
www.elsevier.com/locate/jempfin

International conditional asset allocation under specification uncertainty

Laurent Barras *

Swiss Finance Institute at HEC–University of Geneva, Boulevard du Pont d'Arve 40, 1211 Geneva 4, Switzerland

Received 15 August 2005; received in revised form 28 August 2006; accepted 27 September 2006
Available online 28 December 2006

Abstract

This paper examines the impact of specification uncertainty on the performance of international mean–variance conditional asset allocation. Specification uncertainty is defined as the uncertainty faced by the investor regarding the specification choices necessary to implement a conditional strategy. To assess the impact of this phenomenon, we measure the performance of a group of strategies that the investor could reasonably consider. The strong performance variability across the strategies indicates that the gains previously documented are overstated. Our findings provide an explanation to the apparent paradox between the economic and statistical significance of predictability, and are consistent with the semi-strong form of market efficiency.

JEL classification: G11
Keywords: Conditional asset allocation; Predictability; Specification uncertainty; Performance measurement

1. Introduction

Previous studies testing the profitability of international conditional asset allocation find that they greatly outperform unconditional and buy and hold strategies (Solnik, 1993; Harvey, 1994). This striking performance contributes to the apparent paradox between the economic and statistical significance of predictability. On the one hand, it is surprising that the generated profits are substantial, since the out-of-sample explanatory power of the predictive variables is generally quite low. On the other hand, Kandel and Stambaugh (1996) show that small levels of

* Tel.: +41 223798141; fax: +41 223798104.
E-mail address: barras@hec.unige.ch.

0927-5398/$ - see front matter
doi:10.1016/j.jempfin.2006.09.002

predictability are sufficient to significantly improve the investor's expected utility. Moreover, this superior performance challenges the efficient market hypothesis, since the conditional asset allocation is only based on public information. To implement the conditional asset allocation, the investor has to make many specification choices. For example, he must determine the set of predictive variables. These choices are difficult because there is no solid theoretical ground to justify them. As a result, the investor is subject to specification uncertainty since he does not know which specification will produce the best results in the future. In most studies, this problem is ignored, and the specification is exogenously defined by the researcher[1].

In this paper, we examine the impact of specification uncertainty on the performance of international mean–variance conditional asset allocation[2]. To model this uncertainty, we consider three important specification choices: the predictive variables, the estimation window length, and the portfolio weight constraints. By combining these different choices, the investor can potentially choose among a total of 112 strategies. All of them are implemented using a universe of 12 developed market equity indices between January 1990 and September 2004. They are rebalanced monthly by means of stock index futures to reduce transaction costs. Since this approach makes the currency exposure negligible, our framework is very general and applies to all investors regardless of their nationality.

Our results show that specification uncertainty has a strong impact on performance, and implies that the gains previously documented are overstated. First, performance varies greatly across the different specification choices, especially across the set of predictive variables and portfolio constraints. As a result, an unlucky investor could choose a perfectly reasonable specification which turns out to yield low performance. Second, we do not find any search technologies that can be used to select performing strategies. The best strategies do not share predictable characteristics, and their performance is sensitive to minor specification changes. Financial criteria such as the Sharpe ratio, the cumulative wealth, and the certainty equivalent measure are not helpful in picking performing strategies. We also find that implementations of conditional asset allocation based on the efficient portfolio optimization of Ferson and Siegel (2001) and the international CAPM do not improve performance. All these results are robust to changes in the investor's risk aversion coefficient.

By accounting for specification uncertainty, the apparent paradox between the statistical and economic significance of predictability disappears. When the statistical significance of predictability declines, as it is the case during the 90's (Bossaerts and Hillion, 1999), the proportion of unprofitable conditional strategies increases. Although some performing strategies still exist, specification uncertainty prevents the investor from reliably finding them. The economic significance of predictability declines, as the investor is more likely to choose unprofitable strategies. Over our sample period, the evidence suggests that international equity markets satisfy the semi-strong efficiency defined by Timmermann and Granger (2004)[3].

[1] This is the case for Solnik (1993), Harvey (1994), as well as many dynamic strategies tested in the US market (see, among others, Breen et al., 1989; Klemkosky and Bharati, 1995; Handa and Tiwari, 2006; Marquering and Verbeek, 2004; Xu, 2004).

[2] Few papers study the issue of specification uncertainty (Pesaran and Timmermann, 1995; Cooper and Gulen, 2006). They model dynamic strategies investing either 0% or 100% in a single risky asset. Contrary to them, we examine multiple-asset strategies which depend on the joint predictions across assets, as well as the inverstor's risk aversion.

[3] "A market is efficient with respect to the information set, Ω_t, search technologies, St, and forecasting models, Mt, if it is impossible to make economic profits by trading on the signals produced from a forecasting model in Mt defined over predictor variables in the information set Ω_t and selected using a search technology in St."

The remainder of the paper is as follows. Section 2 explains the design of the international conditional asset allocation. Section 3 defines the notion of specification uncertainty. Section 4 describes the investment universe and the performance measures. Section 5 contains the empirical results.

2. International conditional asset allocation

2.1. Using stock index futures

The investor forms his portfolio from a universe of N country equity indices and his reference currency riskfree asset. Since conditional asset allocation produces high turnover (the average monthly turnover across our set of 112 conditional strategies is around 80%), replicating the country indices with the underlying securities is too costly. To address this issue, we implement the conditional asset allocation with stock index futures.

Using stock index futures reduces transaction costs and facilitates short-selling. But apart from these two characteristics, it is equivalent to a replication strategy hedged against currency risk. To see this, note that the currency-hedged return R^H_{it+1} of the country i index (i=1,…, N) between time t and t+1 is equal to $R_{ft}+r^l_{it+1}+s_{it+1}\cdot r^l_{it+1}$ (see the Appendix). R_{ft} denotes the investor's reference currency riskfree rate set at time t. r^l_{it+1} stands for the return of the country i index in currency i minus the country i riskfree rate, and s_{it+1} represents the rate of return of one currency i in terms of the investor's reference currency. Since the cross-product $s_{it+1}\cdot r^l_{it+1}$ is small, we can neglect it[4]:

$$R^H_{it+1} \approx R_{ft} + r^l_{t+1}. \tag{1}$$

Consider the total return R_{pt+1} of the currency-hedged strategy investing a fraction w_{Ht} in the riskfree asset, and w_{it} in the country i index (i=1,…, N). Using Eq. (1) and the budget constraint, R_{pt+1} is given by:

$$R_{pt+1} = w_{Ht} \cdot R_{ft} + \sum_{i=1}^{N} w_{it} \cdot R^H_{it+1} \approx R_{ft} + \sum_{i=1}^{N} w_{it} \cdot r^l_{it+1}. \tag{2}$$

In the absence of arbitrage, the country i futures contract expiring at time t+1 can be replicated by a long position in the country i index, and a short position in the country i riskless asset. Neglecting $s_{it+1}\cdot r^l_{it+1}$, its return can be approximated by r^l_{it+1} (see the Appendix). From Eq. (2), it follows that another way to implement the strategy is to invest the initial wealth at the riskfree rate R_{ft}, and take positions in the N stock index futures markets[5].

We denote by t_c the roundtrip transaction cost in the futures market, expressed as a percentage of the futures price at time t+1. Following Sutcliffe (2006, 44–45), we set t_c equal to 10 basis points. Since positions in the stock index futures market are taken for one period only, the

[4] Taking the US dollar as the reference currency, the annualized average absolute value of the cross-products across the countries in our sample amounts to 0.18%. Considering that this term does not modify our results.

[5] We assume that the contract initial margin yields an income equal to R_{ft}. However, if there is no income, the effect on the portfolio excess return is weak, because the margin represents only 3% to 5% of the contract notional value (Sutcliffe, 2006). Note that the initial wealth can also be used to satisfy potential margin calls.

turnover of the portfolio at each rebalancing date is equal to one. As a result, the net excess return r_{pt+1}^{net} of the strategy is given by:

$$r_{pt+1}^{\mathrm{net}} = r_{pt+1} - t_c\left(1 + r_{pt+1}\right), \tag{3}$$

where r_{pt+1} denotes the excess return of the conditional strategy: $r_{pt+1} = \Sigma_{i=1}^{N} w_{it} \cdot r_{it+l}^{1}$. Since both r_{pt+1} and r_{pt+1}^{net} do not depend on the investor's reference currency, they are the same regardless of the investor's nationality.

2.2. Mean–variance optimization

The portfolio is initially formed at time t_0 and rebalanced at a monthly frequency. To determine the optimal portfolio weights w_{it} at each time $t(t=t_0,\ldots, T)$, we use the standard conditional mean–variance analysis proposed among others by Solnik (1993), Harvey (1994), and Handa and Tiwari (2006). Consistent with our previous discussion, the optimization is based on currency-hedged country index excess returns.

In the first step, we forecast the N country excess returns during the next month. These forecasts are based on a linear specification of the $(K+1)\times 1$ random vector z_t composed of one (for the constant term) and K predictive variables. For each country, the investor collects M observations of past realizations of r_{it+1}^{l} and z_t up to time t_0, and estimates the $(K+1)\times 1$ vector of regression coefficients $\widehat{\phi}_{ik}^{M}$ by the Ordinary Least Square method. Then, the vector $\widehat{\phi}_{ik}^{M}$ is jointly used with the realized values of the predictive variables at t_0 to obtain a step ahead estimate $\widehat{\pi}_{it_0}$ of the country i conditional expected excess return (in currency i):

$$\widehat{\pi}_{it_0} = z_{t_0}' \widehat{\phi}_{ik}^{M} \tag{4}$$

Each estimate is stacked to form the $N\times 1$ conditional expected excess return vector $\widehat{\prod}_{t_0}$:

$$\widehat{\prod}_{t_0} = [\widehat{\pi}_{1t_0}, \ldots, \widehat{\pi}_{Nt_0}]'. \tag{5}$$

We assume that the conditional covariance matrix $\boldsymbol{\Sigma}$ is constant. Estimating a time-varying conditional covariance matrix is extremely difficult as the number of assets becomes large. Besides these estimation issues, Cavaglia et al. (1997), Han (2005), Marquering and Verbeek (2004) find that the economic significance of volatility predictability is very low at a monthly frequency[6]. Moreover, Chopra and Ziemba (1993) show that covariance misspecification has a far lesser impact on the investor's expected utility than the one related to the expected return vector. To reduce the estimation error contained in the sample covariance matrix, we use the shrinkage estimator $\widehat{\boldsymbol{\Sigma}}$ proposed by Ledoit and Wolf (2004). The idea is to shrink the residual sample covariance matrix $\widehat{S}$ towards a matrix $\widehat{F}$ which can be precisely estimated:

$$\widehat{\Sigma} = \alpha^{*} \cdot \widehat{\mathrm{F}} + (1-\alpha^{*}) \cdot \widehat{\mathrm{S}}, \tag{6}$$

$\widehat{S}$ is computed from the N regression residuals vectors (of size $M\times 1$) computed in the first step. $\widehat{F}$ is the sample covariance matrix constrained to have constant pairwise correlation. α^{*} is

[6] Our discussion is well summarized by Cavaglia et al. (1997): "One of the methods that we employ, the multivariate Garch, is only feasible on a small number of regions or countries. Our results indicate, however, that the simpler rolling-variance–covariance model produce impressive results."

the optimal shrinkage intensity obtained by minimizing $E(\|\widehat{\Sigma}(\alpha)-\Sigma\|^2)$, where Σ denotes the true residual covariance matrix, and $\|\cdot\|^2$ the Frobenius norm[7].

In the second step, the investor uses the two inputs $\widehat{\prod}_{t_0}$ and $\widehat{\Sigma}$ to determine the optimal portfolio weights. The $N\times 1$ vector of country weights is written as $w_{t_0}=[w_{1t_0},\dots, w_{Nt_0}]'$. Replacing w_{Ht_0} with $1-w'_{t_0}1_N$, we maximize the mean–variance function:

$$\max_{w_{t_0}} \; w'_{t_0}\widehat{\prod}_{t_0}-\frac{1}{2}A\cdot w'_{t_0}\widehat{\Sigma}w_{t_0} \quad s.t.\; C(w_{t_0})\leqslant c, \tag{7}$$

where $C(w_{t_0})$ is the set of J bounding constraints, and c the $J\times 1$ vector of bounds. The investor's incentives to put these constraints are explained hereafter. Note that, since there is a riskfree asset, $w'_{t_0}\;1_N$ can be different from one. The parameter A defines the investor's risk aversion. This parameter can affect performance because a low risk aversion induces more aggressive bets. If $\widehat{\prod}_t$ is estimated accurately, this aggressive strategy generates higher performance. However when $\widehat{\prod}_t$ is imprecise, this strategy is more affected by the estimation errors contained in $\widehat{\prod}_t$ (Frost and Savarino, 1988). To address this issue, we choose three different values for A: 2 (low), 5 (medium), and 10 (high).

3. Specification uncertainty

To take advantage of potential predictability, the investor must precisely specify how the conditional asset allocation is implemented. In this paper, we focus on three major sources of specification uncertainty caused by the choices of predictive variables z_t, estimation window length M, and portfolio weight constraints $C(w_t)$.

3.1. Set of predictive variables

The number of potential predictive variables proposed in the literature is large, making the choice of predictive variables extremely difficult (see, among others, Keim and Stambaugh, 1986; Fama and French, 1988, 1989; Goyal and Welch, in press). In an international setting, this issue becomes even more intricate because of the distinction between Global and Local variables. First, their relative importance is likely to depend on the degree of country integration to the world market (Bekaert and Harvey, 1995). Second, Harvey (1991), Ferson and Harvey (1993) find that Local variables still have some predictive power on integrated market returns.

To estimate each country expected excess return, we consider Local and Global predictive variables. In the first case, the variables vary across countries. In the second case, the variables are common to all of them and reflect potential market integration. These Local and Global variables are then classified in three groups. The first group (denoted by Asset) contains variables related to asset prices, namely the dividend yield and the lagged stock index excess return. The second one (Interest) contains variables related to the evolution of interest rates. These are the short-term interest rate and the term spread. The third one (Combined) contains all predictive variables. Finally, we consider a regression model where all Local and Global variables are jointly used (All). Therefore, the set of predictive variables is composed of 7 elements.

[7] The Frobenius norm of the $N\times N$ matrix Z is defined as: $\|Z\|^2=\Sigma_{i=1}^{N}\Sigma_{j=1}^{N} z_{ij}^2$, where z_{ij} is the ith row-jth column element of Z. Although the optimal shrinkage intensity α^* depends on unknown parameters (such as the elements of Σ), it can be consistently estimated as shown by Ledoit and Wolf (2003).

Concerning the Local variables, the dividend yield and the lagged stock index excess return are computed from the Datastream country indices. Short-term interest rates are proxied by one-month euro-market interest rates, and the term spread is defined as the difference between the yield of long-term Government bonds provided by International Financial Statistics and the one-month euro-market interest rate. For the Global variables, the world dividend yield and the lagged world index excess return are computed from the Datastream US dollar world index. The global short-term interest rate and the term spread are proxied by US variables.

3.2. Set of estimation window lengths

The second difficult choice concerns the estimation window length. If the relation between the predictive variables and asset returns is stable through time, one should take a large number of observations to reduce the small sample bias of the regression coefficients (Nelson and Kim, 1993; Stambaugh, 1999). On the contrary, if this relation is subject to structural breaks (Paye and Timmermann, 2006), a model which rapidly reflects these changes is more appropriate.

To estimate the coefficients of the predictive regression for each country, we propose two estimation procedures. The first one uses an expanding window (denoted by Expanding). The initial length of the estimation window is set to 50 or 100 observations. The second one is based on a rolling window (Rolling) with two different fixed window lengths of 50 and 100 observations. As a result, the set of estimation window lengths contains 4 distinct elements.

3.3. Set of portfolio weight constraints

Mean–variance optimization is very sensitive to estimation risk, as shown by Jobson and Korkie (1981). Estimation errors affect $\widehat{\prod}_t$ because predictive regressions are very noisy, and subject to estimation bias and structural breaks. To avoid taking excessive positions, the investor can impose weight constraints. Frost and Savarino (1988) find that upper-bound and short-selling constraints reduce the impact of estimation risk. Adding constraints can, however, hurt performance if $\widehat{\prod}_t$ contains valuable predictive information. Since the overall effect of a weight constraint on performance depends on the unknown return generating process, it cannot a priori be determined (Jagannathan and Ma, 2003).

We assume that the investor uses portfolio constraints only to reduce estimation risk. We do not consider cases where statutory constraints are exogenously imposed to the investor[8]. Our set of constraints is composed of 4 cases. The first one (denoted by Free) does not impose any constraints on the individual country weights. The second one (Positivity) prohibits short-selling. The third one (Diversification) forces each country weight to lie between 0 and 20%. The fourth one (Variability) imposes that the absolute difference of each country weight at two consecutive rebalancing dates lies between 0 and 30%. For Free and Variability, we add another constraint which forces the sum of the squared weights to be less than 2. It eliminates extreme cases where the overall position in the N countries, defined as $|w_t'|\ 1_N$, is so large that it could not be taken in the futures market.

Table 1 summarizes the elements forming the three specification sets. By combining all these elements, the investor can potentially choose among 112 conditional strategies. This number is computed as the product of the number of elements in each specification set ($7\times4\times4$).

[8] This is the case for mutual and pension fund managers, who frequently face short-selling or upper-bound constraints. These cases are not problematic, as it is unlikely that these managers have the freedom required to implement the conditional asset allocation examined in this paper.

Table 1
Composition of the three specification sets

Set of predictive variables	
1. Local Asset	Dividend yield and lagged stock index excess return
2. Local Interest	Short-term interest rate and term spread
3. Local Combined	All Local variables
4. Global Asset	Dividend yield and lagged stock index excess return
5. Global Interest	Short-term interest rate and term spread
6. Global Combined	All Global variables
7. All	All Local and Global variables
Set of estimation window lengths	
1. Expanding 50	Expanding window with 50 initial observations
2. Expanding 100	Expanding window with 100 initial observations
3. Rolling 50	Rolling window with 50 observations
4. Rolling 100	Rolling window with 100 observations
Set of portfolio weight constraints	
1. Free	No constraints on w_{it}
2. Positivity	$0 \leq w_{it} \leq 1$ and $w_t' 1 \leq 1$
3. Diversification	$0 \leq w_{it} \leq 20\%$
4. Variability	$\lvert w_{it} - w_{it-1} \rvert \leq 30\%$

Description of the different elements forming the three specification sets (predictive variables, estimation window lengths, and portfolio weight constraints). The elements of these three sets are combined together to create a total of 112 conditional strategies.

Our modelling of specification uncertainty is conservative since there are many other specification choices to be made, such as the trading rule or the rebalancing horizon. But if we find that specification uncertainty affects performance across the 112 strategies, this conclusion is likely to be preserved under finer representations of specification uncertainty[9].

4. Data and performance measurement

4.1. Investment universe

Our investment universe is composed of 12 developed markets: Australia, Belgium, Canada, Denmark, France, Germany, Italy, Japan, the Netherlands, Switzerland, the United Kingdom, and the United States. Ideally, the country index excess return r_{it+1}^{l} should be computed with country indices on which commonly traded futures contracts are based[10]. Unfortunately, most of these time-series contain few observations, since they only start during the 90's. For this reason, we proxy r_{it+1}^{l} by the gross returns (price and dividends reinvested) of the 12 Datastream country indices minus their respective riskfree rates given by the one-month euro-market interest rates. This choice is reasonable because Datastream indices are very close to their local counterparts: computed with the longest possible time-series, the annualized average mean difference (in absolute value) amounts to 0.7%, and the average pairwise correlation is equal to 0.98. Except for

[9] To assess the robustness of our results, we considered different groupings of the predictive variables, different window sizes, different thresholds for Diversification and Variability, and an estimation technique with fixed regression coefficients. These results, available upon request, are similar to those exposed in Section 5.

[10] These are the following: S&P/ASX 200 (Australia), BEL 20 (Belgium), Toronto 35 Index (Canada), KFX (Denmark), CAC 40 (France), DAX (Germany), MIB 30 (Italy), S&P Topix (Japan), AE-Index (Netherlands), SMI (Switzerland), FT-SE 100 (UK), S&P 500 (USA).

Table 2
Descriptive statistics of the Datastream country indices

	Excess mean	Std deviation	Skewness	Kurtosis	Cross-correlation
Australia	4.7%	12.9%	−0.06	2.70	0.53
Belgium	3.7%	16.3%	−0.34	4.17	0.61
Canada	5.5%	14.3%	−0.45	4.58	0.60
Denmark	5.1%	18.3%	−0.23	2.95	0.60
France	5.0%	19.1%	−0.18	3.36	0.68
Germany	2.3%	19.6%	−0.44	3.61	0.67
Italy	2.0%	23.9%	0.53	3.65	0.53
Japan	−4.4%	20.7%	0.26	4.12	0.34
Netherlands	6.0%	17.4%	−0.50	5.06	0.71
Switzerland	7.9%	17.6%	−0.32	7.48	0.67
UK	2.8%	15.0%	−0.15	3.70	0.66
USA	7.2%	14.6%	−0.34	3.34	0.59

Descriptive statistics of the Datastream country index monthly excess returns over their respective riskfree rates (in local currency) between January 1990 and September 2004. For each country, the riskfree rate is proxied by the one-month euro-market interest rate. The excess mean and standard deviation are annualized. The annualization factor is set to 12 for the excess mean and $12^{1/2}$ for the standard deviation. Cross-correlation is computed for each country as the average return correlation with the 11 remaining countries.

the UK and the USA, the above countries have introduced stock index futures only in the late 80's (Sutcliffe, 2006). For this reason, all strategies considered in this paper start in January 1990. Some descriptive statistics on the excess return time-series are shown in Table 2.

4.2. Performance measures

To assess the economic significance of predictability, we measure the performance of the conditional strategy against various benchmark strategies which do not account for predictability. Our first measure is the unconditional Sharpe ratio differential, SD, estimated as:

$$\widehat{\mathrm{SD}} = \frac{\widehat{\mu}_p}{\widehat{\sigma}_p} - \frac{\widehat{\mu}_B}{\widehat{\sigma}_B}, \tag{8}$$

where $\widehat{\mu}_p$, $\widehat{\sigma}_p$ and $\widehat{\mu}_B$, $\widehat{\sigma}_B$ respectively denote the estimated excess mean and standard deviation of the conditional and the benchmark strategies. Our second measure is the unconditional portfolio weight measure proposed by Ferson and Khang (2002):

$$\mathrm{FK} = \sum_{i=1}^{N} \mathrm{cov}\left(w_{it} - w_{Bit}, r_{it+1}^{l}\right) = \sum_{i=1}^{N} \mathrm{cov}\left(w_{it}, r_{it+1}^{l}\right) - \sum_{i=1}^{N} \mathrm{cov}\left(w_{Bit}, r_{it+1}^{l}\right), \tag{9}$$

where w_{it} and w_{Bit} denote the weights of the country i at time t for the conditional and the benchmark strategies. This measure indicates whether the conditional strategy creates value over the benchmark by overweighting countries with higher conditional excess returns. Following Grinblatt and Titman (1993), it is estimated as:

$$\widehat{\mathrm{FK}} = \frac{1}{T_1} \sum_{t=k+1}^{T} \sum_{i=1}^{N} r_{it+1}^{l}(w_{it} - w_{it-k}) - r_{it+1}^{l}\left(w_{it}^{B} - w_{it-k}^{B}\right) = \frac{1}{T_1} \sum_{t=k+1}^{T} r_{t+1}^{0} - r_{Bt+1}^{0}, \tag{10}$$

where $T_1 = T - k + 1$, r_{t+1}^{0} is the return of a zero-cost portfolio formed with a long position w_{it} and a short position w_{it-k} in the country i, and r_{Bt+1}^{0} is defined in the same way. Following Grinblatt and

Titman (1993), we set k to 1 year. One advantage of SD and FK is that they are not based on an asset pricing model. As discussed by Solnik (1993), systematic risk is more difficult to define in an international setting. Another advantage is that these measures only depend on r_{pt+1}^{net} and r_{it+1}^{l}, and are, therefore, independent of the investor's nationality.

Dybvig and Ross (1985) show that the Sharpe ratio may not be a reliable indicator of the performance of mean–variance conditional asset allocation. They show that, if excess returns are conditionally normal, the optimization in Eq. (7) is equivalent to the maximization of the exponential utility function. In this case, the portfolio is on the conditional mean–variance frontier. However, it is not on the unconditional frontier, because the investor with an exponential utility function is ready to trade a lower Sharpe ratio for a higher skewness. With one risky asset, Dybvig and Ross (1985) find that SD is negative if the risky asset excess return, μ, is higher than its standard deviation, σ.

From an empirical point of view, observing a negative SD is unlikely, because the condition $\mu > \sigma$ is very stringent (In Table 2, none of the countries satisfy this condition)[11]. Nevertheless, we propose two ways to deal with this potential issue. The first one consists in using FK. The FK measure of the conditional strategy is positive if and only if the investor has predictive information (see the Appendix). Therefore, the sign of FK in the presence of predictive information is clearly identified. Second, we implement the efficient portfolio optimization technique proposed by Ferson and Siegel (2001). Since their strategy maximizes the unconditional Sharpe ratio in the presence of conditioning information, it is immune from the Dybvig and Ross critique.

4.3. Benchmark strategies

Our first benchmark strategy is the unconditional asset allocation. It is identical to the conditional strategy, except that $\widehat{\prod}_t$ and $\widehat{\Sigma}$ are replaced with the estimated unconditional moments, $\widehat{\mu}$ and $\widehat{V}$, computed using past returns. To account for specification uncertainty in this unconditional framework, we use the different estimation windows and portfolio constraints shown in Table 1. As a result, a total of 16 unconditional strategies are tested (4×4). Our second benchmark is a buy and hold strategy corresponding to the value-weighted portfolio (denoted by Market). Finally, we implement an equally-weighted portfolio (Equal), where the portfolio weights are rebalanced at the beginning of each month.

5. Empirical results

5.1. Average performance of the conditional asset allocation

5.1.1. Comparative analysis of the conditional and benchmark strategies

Table 3 compares the annualized excess mean, standard deviation, and Sharpe ratio averaged across the 112 conditional strategies with those generated by the 16 unconditional strategies, Market, and Equal. For each measure, we test the null hypothesis that the difference between the conditional and benchmark strategies is equal to zero, against the alternative that it is positive. To compute the p-values in parentheses, we use a bootstrap method explained in the Appendix. In our case, there are two advantages to use the bootstrap instead of asymptotic theory. First, the

[11] Further analysis is required to determine the sign of SD, as this condition is only sufficient (but not necessary), and does not extend to multiple assets. Assuming conditional normality, SD has a closed-form expression valid with one or multiple assets. Using different asset classes and predictability levels, Barras (2006) finds that SD is always positive.

Table 3
Comparative analysis of the conditional and benchmark strategies

	Excess mean	Standard deviation	Sharpe ratio
Low aversion (A = 2)			
Conditional versus	7.2%	28.5%	0.23
Unconditional	4.8% (0.22)	28.2% (0.31)	0.15 (0.22)
Market	2.2% (0.08)*	13.9% (0.00)*	0.16 (0.34)
Equal	2.7% (0.08)*	13.9% (0.00)*	0.20 (0.36)
Medium aversion (A = 5)			
Conditional versus	5.2%	22.2%	0.23
Unconditional	3.4% (0.24)	19.6% (0.00)*	0.16 (0.31)
Market	2.2% (0.17)	13.9% (0.00)*	0.16 (0.34)
Equal	2.7% (0.14)	13.9% (0.00)*	0.20 (0.40)
High aversion (A = 10)			
Conditional versus	3.9%	17.4%	0.22
Unconditional	2.3% (0.22)	12.7% (0.00)*	0.16 (0.33)
Market	2.2% (0.29)	13.9% (0.00)*	0.16 (0.39)
Equal	2.7% (0.31)	13.9% (0.00)*	0.20 (0.45)

This table compares the annualized excess mean, standard deviation, and Sharpe ratio averaged across the 112 conditional strategies with those generated by the benchmark strategies (Unconditional, Market, and Equal). For the unconditional strategies, the figures are averaged across the 16 strategies. Market and Equal denote the value-weighted and equally-weighted portfolios, respectively. The annualization factor is set to 12 for the excess mean and $12^{1/2}$ for the standard deviation as well as the Sharpe ratio. For each of the three measures, the figures in parentheses represent the bootstrapped *p*-values under the null hypothesis that there is no difference between the conditional and the benchmark strategies. An asterisk indicates that the difference is positive at the 10% significance level.

bootstrap is often more accurate in finite sample (Davison and Hinkley, 1997; Horowitz, 2001). Second, the difference in mean, standard deviation, or Sharpe ratio depends on the joint distribution of all strategies, making its asymptotic distribution hard to calculate.

The average excess mean and standard deviation of the conditional strategies are always higher than those of the benchmark strategies. While the excess mean differences are significant (at the 10% level) only 2 times out of 9, the standard deviation is significantly higher in all but one cases. Compared to the benchmarks, the average Sharpe ratio of the conditional strategies is always superior. However, the differences are not sufficiently high to reject the null hypothesis of equal performance. In particular, the Sharpe ratio of Equal (0.22) is very close to the one produced by the conditional strategies. Changing the risk aversion coefficient alters the risk-return profile of the conditional strategies. When A is low, the investor takes more aggressive bets (positive or negative) in the various countries. This portfolio generates high returns when the predictions are correct, but is also more risky. However, we see that these two effects offset each other, since the Sharpe ratio remains unchanged across the different values of A.

5.1.2. Percentage of conditional strategies with positive performance

The previous analysis shows that, on average, the conditional asset allocation does not significantly outperform the benchmark strategies. One obvious reason for this result is the predictability decline observed during the 90's (Bossaerts and Hillion, 1999; Goyal and Welch, in press). To determine whether some strategies still dominate the benchmarks despite this decline, we measure the proportion of the 112 conditional strategies yielding a positive performance

Table 4
Percentage of conditional strategies with positive performance

	U50	U75	Market	Equal
Panel A: Sharpe ratio differential				
Low aversion ($A=2$)	75.0% (0.00)*	66.9% (0.00)*	71.4% (0.01)*	65.2% (0.01)*
Medium aversion ($A=5$)	75.0% (0.00)*	58.9% (0.12)	67.8% (0.01)*	58.0% (0.08)*
High aversion ($A=10$)	78.6% (0.00)*	55.4% (0.34)	75.0% (0.00)*	64.3% (0.08)*
Panel B: Ferson–Khang measure				
Low aversion ($A=2$)	80.4% (0.07)*	79.5% (0.06)*	80.4% (0.08)*	82.2% (0.07)*
Medium aversion ($A=5$)	77.7% (0.09)*	74.1% (0.12)	83.0% (0.06)*	84.8% (0.05)*
High aversion ($A=10$)	82.1% (0.07)*	78.6% (0.09)*	83.9% (0.07)*	86.6% (0.06)*

This table contains the percentage of conditional strategies having a positive performance against four benchmark strategies. The results using the Sharpe ratio differential and the Ferson–Khang measure are given in Panels A and B, respectively. The two unconditional benchmarks (U50 and U75) are located at the 50 and 75% top quantiles of the ranking based on the Sharpe ratio of the 16 unconditional strategies. Market and Equal denote the value-weighted and equally-weighted portfolios, respectively. The figures in parentheses represent the bootstrapped p-values under the null hypothesis that the percentage is equal to zero. An asterisk denotes significance at the 10% level.

against four benchmark strategies. We choose two unconditional strategies (denoted by U50 and U75) located at the 50% and 75% top quantiles of the ranking based on the Sharpe ratio of the 16 unconditional strategies. The two other strategies are Market and Equal. To test the null hypothesis that the percentage is equal to zero, we compute the p-values with the bootstrap method explained in the Appendix.

The results using the Sharpe ratio differential are summarized in Panel A of Table 4. The percentages of conditional strategies which outperform U50 and Market are very high, as they range from 67.8% to 78.6% across the different values of A. Moreover, all percentages are strongly significant. The percentages computed against U75 and Equal are lower, but they are significant except in two cases. In Panel B, the proportion of strategies with a positive Ferson–Khang measure is also very high. The percentages are always superior to 70%, and only one is not significant. Our analysis shows that many conditional strategies generate a positive performance. Similarly to previous studies, it would therefore be easy to choose exogenously a conditional strategy which outperforms all four benchmark strategies. To assess the economic significance of predictability, we must go one step further and determine whether this potential performance is achievable once specification uncertainty is accounted for.

5.2. Impact of specification uncertainty on performance

We measure the performance variation across the different elements forming the three specification sets (predictive variables, estimation windows, and portfolio constraints). If the variability is important, the impact of specification uncertainty is high (and vice-versa). For each element, we compute the percentage of strategies yielding a positive performance against the four benchmark strategies previously mentioned (U50, U75, Market, and Equal). Since the performance variation is similar regardless of A, we only discuss the results for $A=5$ (the other results are available upon request).

5.2.1. The Sharpe ratio differential

Table 5 contains the percentages for the Sharpe ratio differential (SD). The bootstrapped p-values under the null hypothesis that the percentage is equal to zero are given in parentheses. First, the

Table 5
Impact of specification uncertainty on the Sharpe ratio differential

		#	U50	U75	Market	Equal
Predictive variables						
Local	Asset	16	43.7% (0.50)	18.7% (0.95)	37.5% (0.74)	18.7% (0.98)
	Interest	16	75.0% (0.01)*	50.0% (0.32)	62.5% (0.07)*	50.0% (0.34)
	Combined	16	81.2% (0.00)*	62.5% (0.11)	75.0% (0.01)*	56.2% (0.13)
Global	Asset	16	100.0% (0.00)*	81.2% (0.01)*	93.7% (0.00)*	81.2% (0.00)*
	Interest	16	50.0% (0.47)	31.2% (0.97)	37.5% (0.82)	31.2% (0.98)
	Combined	16	93.7% (0.00)*	87.5% (0.00)*	87.5% (0.00)*	87.5% (0.00)*
All		16	81.2% (0.00)*	81.2% (0.01)*	81.2% (0.00)*	81.2% (0.00)*
Estimation windows						
Expanding	50	28	71.4% (0.00)*	60.7% (0.08)*	64.3% (0.02)*	57.1% (0.09)*
	100	28	71.4% (0.01)*	60.7% (0.13)	67.8% (0.02)*	60.7% (0.06)*
Rolling	50	28	75.0% (0.00)*	64.3% (0.11)	71.4% (0.02)*	64.2% (0.07)*
	100	28	82.1% (0.00)*	50.0% (0.22)	67.8% (0.03)*	50.0% (0.19)
Portfolio constraints						
Free		28	71.4% (0.01)*	50.0% (0.23)	60.7% (0.12)	50.0% (0.24)
Positivity		28	57.4% (0.33)	35.7% (0.87)	46.4% (0.70)	32.1% (0.99)
Diversification		28	92.8% (0.00)*	85.7% (0.01)*	92.8% (0.00)*	85.7% (0.00)*
Variability		28	78.6% (0.00)*	64.3% (0.07)*	71.4% (0.00)*	64.2% (0.06)*

This table contains the percentage of conditional strategies with a positive Sharpe ratio differential against four benchmarks across each element of the three specification sets (predictive variables, estimation windows, and portfolio constraints). The risk aversion coefficient A is equal to 5. The two unconditional benchmarks (50B, 75B) are located at the 50 and 75% top quantiles of the ranking based on the Sharpe ratios of the 16 unconditional strategies. Market and Equal denote the value-weighted and equally-weighted portfolios, respectively. The figures in parentheses represent the bootstrapped p-values under the null hypothesis that the percentage is equal to zero. An asterisk denotes significance at the 10% level.

performance variation across the different predictive variables is important. For instance, the difference between the lowest and highest percentages against Market amounts to 56.2%! Global Combined obtains the highest percentages against all benchmarks. This superior performance only stems from the informational content of Global Asset, since Global Interest yields low percentages. The performance declines when Local instead of Global variables are considered. The percentages are generally lower, and none of the Local specifications yield significant percentages against U75 and Equal. These results suggest that the 12 countries are integrated to the world market. Second, the choice of the estimation window length does not have a strong impact on performance. The maximum difference between the lowest and highest percentages for a given benchmark is a low 14.3% (against U75). Paye and Timmermann (2006) find that Germany, Japan, and the USA are subject to structural breaks during the 90's. Our results indicate that these breaks do not represent an important source of specification uncertainty.[12] Third, the percentage variation across the different portfolio constraints is strong. Diversification obtains the best results, followed by Variability. The percentages produced by Free are low, and highlight the need to add constraints to reduce estimation risk. Preventing short-selling (Positivity) produces the weakest performance, which is surprising in light of the results of Frost and Savarino (1988).

[12] Pettenuzzo and Timmermann (2005) find that accounting for past and potential future breaks can have a large effect on the proportion invested in US stocks. One possible explanation for this result is the length of their sample (1926–2003). Over such a long period, coefficient changes are likely to be much larger than in our case.

The wide percentage variation across the sets of predictive variables and portfolio constraints clearly indicates that specification uncertainty has an important impact on performance. For instance, if the investor uses Global Combined, he has a high probability of beating most benchmark strategies. This would not be the case with Local Asset. Similarly, 85.7% of the Diversification strategies dominate U75 and Equal. On the contrary, only 50% of the Free strategies can outperform these two benchmarks.

5.2.2. The Ferson–Khang measure

The results obtained with the Ferson–Khang measure (FK) are shown in Table 6. We see that all percentages are fairly similar across the four benchmarks. This is not surprising, since these strategies do not use potential predictability, making the last term in Eq. (9) close to zero. First, the performance variation across the different predictive variables is very strong. Contrary to Local variables, the percentages produced by Global Asset and Global Combined are always high and significant. Second, the performance across the different estimation window lengths is fairly constant. Apart from Expanding 100, the other elements produce similar percentages. Third, the results for the portfolio constraints show a clear difference between the elements Free, Positivity and the other two, Diversification and Variability. While the first group does not produce any significant percentages against all benchmarks, this is exactly the opposite for the second one.

Table 6
Impact of specification uncertainty on the Ferson–Khang measure

		#	U50	U75	Market	Equal
Predictive variables						
Local	Asset	16	68.7% (0.34)	56.2% (0.45)	68.7% (0.35)	62.5% (0.35)
	Interest	16	68.7% (0.36)	68.7% (0.34)	75.0% (0.22)	75.0% (0.23)
	Combined	16	50.0% (0.48)	43.7% (0.55)	56.2% (0.42)	68.7% (0.31)
Global	Asset	16	100.0% (0.00)*	100.0% (0.00)*	100.0% (0.00)*	100.0% (0.00)*
	Interest	16	81.2% (0.19)	75.0% (0.22)	87.5% (0.11)	87.5% (0.11)
	Combined	16	93.7% (0.06)*	93.7% (0.06)*	100.0% (0.00)*	100.0% (0.00)*
All		16	81.2% (0.13)	81.2% (0.12)	93.7% (0.05)*	93.7% (0.05)*
Estimation windows						
Expanding	50	28	89.3% (0.00)*	85.7% (0.05)*	89.3% (0.02)*	89.3% (0.02)*
	100	28	67.8% (0.28)	64.7% (0.36)	71.4% (0.26)	75.0% (0.22)
Rolling	50	28	71.4% (0.20)	67.8% (0.24)	82.1% (0.09)*	85.7% (0.07)*
	100	28	82.1% (0.09)*	78.6% (0.12)	89.3% (0.04)*	89.3% (0.05)*
Portfolio constraints						
Free		28	67.8% (0.29)	64.3% (0.30)	75.0% (0.21)	78.6% (0.18)
Positivity		28	67.8% (0.22)	60.7% (0.36)	78.6% (0.19)	78.6% (0.18)
Diversification		28	85.7% (0.08)*	82.1% (0.09)*	89.3% (0.05)*	89.3% (0.05)*
Variability		28	89.3% (0.06)*	85.7% (0.08)*	89.3% (0.06)*	85.7% (0.07)*

This table contains the percentage of conditional strategies with a positive Ferson–Khang measure against four benchmarks across each element of the three specification sets (predictive variables, estimation windows, and portfolio constraints). The risk aversion coefficient A is equal to 5. The two unconditional benchmarks (50B, 75B) are located at the 50 and 75% top quantiles of the ranking based on the Sharpe ratios of the 16 unconditional strategies. Market and Equal denote the value-weighted and equally-weighted portfolios, respectively. The figures in parentheses represent the bootstrapped p-values under the null hypothesis that the percentage is equal to zero. An asterisk denotes significance at the 10% level.

Interestingly, FK and SD both assign a positive performance to the same strategies (Global Asset, Global All, Diversification, Variability). It means that the strategies which are able to outweigh the performing countries are also those with higher Sharpe ratios. The main difference between the two measures comes from the higher percentage level required by FK to reject the null hypothesis. For example, although 81.2% of All strategies yield a positive measure against U50, the *p*-value is slightly above 0.10. One reason is the positive correlation between the returns of the zero-cost portfolio r^0_{t+1} across the conditional strategies, and the low correlation between r^0_{t+1} and r^0_{Bt+1} (see Eq. (10))[13]. If $\widehat{\mathrm{FK}}$ is higher than usual for a given strategy, the dependence structure implies that it will also be the case for other strategies. Therefore, even if the true FK is equal to zero for all strategies, it is not rare to observe a high percentage of strategies with a positive $\widehat{\mathrm{FK}}$.

5.3. Detecting the best conditional strategies

The performance variation measured across the different conditional strategies reveals the important impact of specification uncertainty. This impact could be reduced, however, if the investor could pick the best strategies. To this end, we present tests of different search technologies that could be used to detect these strategies ex-ante.

5.3.1. Do the best strategies share common characteristics?

An obvious search technology is to determine common characteristics shared by the best strategies. We rank the 112 conditional strategies according to their Sharpe ratios, and compare the specification choices implied by the 20 worst and best strategies. The results displayed in Table 7 first show that the predictive variables Local Combined, Global Combined, and All appear more often in the best strategies. Second, no differences emerge from the set of estimation windows, as each specification is equally spread among the best and worst strategies. Third, Diversification and Variability constraints yield the best performance, since they constitute at least 65% of the best strategies. On the contrary, Free and Positivity are present in at least 70% of the worst strategies.

Consistent with the previous results, we find that characteristics such as Global Combined or Diversification appear more often in the best strategies. However, it is not certain that the investor can use this piece of information to detect the best strategies ex-ante. First of all, it is difficult to find strong theoretical reasons to motivate these choices. We could argue that a predictive regression based on Global variables would a priori be a reasonable model, since we focus on developed markets. However, Global Interest produces an important proportion of the worst models. By the same token, adding constraints seems a reasonable ex-ante decision to reduce estimation risk. But it is not clear why Diversification or Variability is better than Positivity. Second, the performance of the best strategies is sensitive to minor specification changes. As a result, the investor's specification choice has to be very close to those of the best strategies. For instance, suppose that the investor with $A=5$ thinks that Global Combined predicts excess return with the highest accuracy. With Expanding 50 and Positivity, the Sharpe ratio amounts to 0.55. But if he uses Rolling 50 instead, the Sharpe ratio shrinks to 0.08. Similarly, the investor with $A=2$ may believe that Variability constraints are optimal. Associated with All and Rolling 50,

[13] The average correlation between r^0_{t+1} across the 112 strategies is equal to 0.42. Taking U75 as the benchmark, the average correlation between r^0_{t+1} and r^0_{Bt+1} is three times lower.

Table 7
Specification choices implied by the worst and best conditional strategies

		Risk aversion					
		Low ($A=2$)		Medium ($A=5$)		High ($A=10$)	
		Worst	Best	Worst	Best	Worst	Best
Predictive variables							
Local	Asset	30%	0%	20%	0%	25%	5%
	Interest	10%	10%	20%	5%	15%	20%
	Combined	10%	20%	5%	25%	5%	15%
Global	Asset	10%	15%	0%	10%	0%	10%
	Interest	20%	10%	35%	10%	45%	5%
	Combined	5%	20%	5%	15%	5%	20%
All		15%	25%	15%	35%	5%	25%
		100%	100%	100%	100%	100%	100%
Estimation windows							
Expanding	50	25%	25%	30%	35%	25%	30%
	100	25%	20%	30%	20%	35%	20%
Rolling	50	30%	30%	20%	25%	15%	35%
	100	20%	25%	20%	20%	25%	15%
		100%	100%	100%	100%	100%	100%
Portfolio constraints							
Free		25%	35%	30%	15%	45%	10%
Positivity		55%	0%	50%	5%	25%	15%
Diversification		0%	10%	0%	40%	5%	40%
Variability		20%	55%	20%	40%	25%	35%
		100%	100%	100%	100%	100%	100%

This table contains the proportion of the worst and best strategies containing a given element of the three specification sets (predictive variables, estimation windows, and portfolio constraints). To determine the worst and best strategies, we rank the 112 conditional strategies according to their Sharpe ratios. The 20 bottom (top) strategies are referred to as the worst (best) strategies.

this choice generates a Sharpe ratio equal to 0.50. However, changing All for Global Asset yields a Sharpe ratio of 0.01.

5.3.2. *Are financial criteria useful?*

A second search technology consists in using a financial criterion to select the strategy (Pesaran and Timmermann, 1995). At each rebalancing date t, the criterion is computed for the 112 conditional strategies with data collected up to time t. Then, the strategy with the highest criterion is chosen for the next period. The same procedure is repeated until the end of the investment period. The first criterion we use is the Sharpe ratio, SR_t, and the second one is the cumulative wealth, W_t. The third one is the certainty equivalent, CE_t, defined as $\mathrm{CE}=\mathrm{E}(R_{pt+1})-(1/2)\ A\cdot var\ (R_{pt+1})$. Contrary to SR_t, the criteria W_t and CE_t depend on the portfolio total return R_{pt+1} and, therefore, on the investor's riskfree rate R_{ft}. We take the US dollar interest rate as a proxy for R_{ft}, but the results are similar for other currencies. We use the first 2 years of observations to compute the criterion initial value for each strategy.

These three criteria lead to very risky strategies. Regardless of A, the risk is approximately 10% higher than the average standard deviation shown in Table 3. This high volatility has a strong negative impact on the Sharpe ratio. For $A=2$, the latter is equal to 0.03 and 0.10 for the SR_t- and

W_t-strategies. For $A=10$, these figures are even negative. In all cases, the strategies are beaten by the unconditional strategies (their average Sharpe ratios are around 0.35), Market (0.34), and Equal (0.35). Using the third criterion, CE_t, increases the excess mean and the Sharpe ratio. Apart from $A=2$, the Sharpe ratio is slightly higher than the one produced by the benchmarks. However, the differences are not significant. In a study on the US market, Pesaran and Timmermann (1995) find that SR_t- and W_t-strategies outperform the S&P 500 between 1960–1992. However, our results are not necessarily contradictory, since their strategies are beaten during two subperiods (1960–1969 and 1980–1989 with low transaction costs). Using these criteria may therefore be useful in specific periods, but not during the 90's characterized by a strong predicability decline.

5.4. Comparison with other conditional strategies

5.4.1. Conditional strategies based on the Ferson–Siegel optimization

The optimization rule proposed by Ferson and Siegel (2001) consists in maximizing the unconditional Sharpe ratio based on the information set I_t. The conditional strategy is determined by maximizing the following quadratic utility function conditionally on I_t:

$$\max E\left[W_{t+1}-\frac{a}{2}\cdot W_{t+1}^2|I_t\right] \Longleftrightarrow \max E\left[b\cdot R_{pt+1}-\frac{c}{2}\cdot R_{pt+1}^2|It\right], \tag{11}$$

where W_{t+1} is the investor's wealth at time $t+1$, and the coefficients b and c are defined as follows: $b=1-c$, $c=a\cdot W_t$. The investor's risk aversion A is equal to: $\frac{U''(W_t)\cdot W_t}{U'(W_t)}=\frac{a\cdot W_t}{1-a\cdot W_t}$. Using this result, we replace b by $\frac{1}{1+A}$ and c by $\frac{A}{1+A}$ in Eq. (11). Since the term $(A\cdot R_{ft}\cdot w_t'\widehat{\prod}_t)$ is very low, we neglect it[14] to obtain the equivalent maximization function:

$$\max_{w_t} w_t'\widehat{\prod_t}-\frac{1}{2}A\cdot w_t'\left(\widehat{\Sigma}+\widehat{\prod_t}\widehat{\prod_t}'\right)w_t \quad s.t. \quad C(w_t)\leqslant c, \tag{12}$$

where $C(w_t)$ is the set of J bounding constraints and c the $J\times 1$ vector of bounds. This optimization is run at each time t across the different elements of the three specification sets. If there is no estimation risk, and if $\prod_t$ and $\boldsymbol{\Sigma}$ are correctly specified, the unconstrained Ferson–Siegel (FS) optimization must produce the highest unconditional Sharpe ratio. But since $\widehat{\prod}_t$ contains estimation errors, this is not necessarily true anymore[15]. As shown by Ferson and Siegel (2001), the FS weights are more robust to estimation risk, since they are bounded functions of the information signal I_t. However, this robustness is not sufficient to offset the effects of estimation error. For $A=5$, we find that the average Sharpe ratio across the 28 possible strategies (considering the 7 predictive variables and the 4 estimation windows) is only equal to 0.10, and the average standard deviation amounts to 67%! In light of this result, we decide to add the bounding constraints $C(w_t)$ to the FS optimization.

In Table 8, we compare the performance of the constrained FS strategies with the one produced by the standard (S) strategies examined so far. We compute the differences between the annualized

[14] Taking the maximum value of $A\cdot R_{ft}$ (equal to 0.07 for $A=10$), and multiplying it by the average excess return of the conditional strategy (equal to 0.3% per month; see Table 3) produces a value of only 0.02%. By discarding this term, we make the optimization independent of the investor's nationality.

[15] This problem is similar to the one examined by Ferson, Siegel and Xu (2006). In theory, their time-varying weight mimicking portfolio must yield the highest correlation with the underlying factor. However, specification and estimation errors prevent their mimicking portfolio to fully deliver its potential performance.

Table 8
Comparison with other conditional strategies

		Ferson–Siegel			International CAPM		
		Mean	Std deviation	Sharpe ratio	Mean	Std deviation	Sharpe ratio
Overall							
		−2.2%*	−4.4%*	−0.06*	−1.4%*	+0.2%	−0.06
Predictive variables							
Local	Asset	−0.5%	−2.4%*	−0.03	−0.3%	+1.3%*	−0.02
	Interest	−2.2%	−4.7%*	−0.05	+0.8%	−0.4%	+0.04
	Combined	−3.0%*	−4.7%*	−0.08*	−2.2%	+2.1%*	−0.09
Global	Asset	−0.7%	−2.3%*	−0.01	−3.7%	+1.5%	−0.20*
	Interest	−2.6%	−5.6%*	−0.07	+1.9%	−2.8%*	+0.11
	Combined	−2.7%	−5.1%*	−0.06	−2.7%	+1.8%*	−0.11
All		−4.0%*	−6.1%*	−0.09	−4.4%	+4.6%*	−0.16*
Estimation windows							
Expanding	50	−1.8%	−3.6%*	−0.05	−1.8%	−2.5%*	−0.09
	100	−1.6%	−3.1%*	−0.05*	+0.2%	+1.9%*	+0.00
Rolling	50	−2.2%	−5.6%*	−0.03	−4.4%*	+2.0%*	−0.16*
	100	−3.3%*	−5.2%*	−0.09*	+0.4%	−0.2%*	+0.03
Portfolio constraints							
Free		−3.5%*	−6.3%*	−0.09*	−3.6%*	+1.3%*	−0.13*
Positivity		+0.0%	−1.0%*	+0.01	−0.2%	−0.5%*	−0.01
Diversification		−1.6%	−4.1%*	−0.05*	−0.5%	−0.5%*	−0.02
Variability		−3.8%*	−6.2%*	−0.09*	−1.3%	+0.9%	−0.06

This table contains the difference in annualized average mean, standard deviation, and Sharpe ratio between the Ferson–Siegel and standard strategies, and between the international CAPM and standard strategies. These differences are computed across all strategies (Overall), as well as each element of the three specification sets (predictive variables, estimation windows, and portfolio constraints). The risk aversion coefficient A is equal to 5. The Ferson–Siegel strategy maximizes the unconditional Sharpe ratio based on predictive information. The international CAPM strategy uses the international CAPM to compute the conditional expected excess return vector. The annualization factor is set to 12 for the excess mean and $12^{1/2}$ for the standard deviation as well as the Sharpe ratio. For each of the three measures, an asterisk indicates that the difference is significant at the 10% level.

average excess mean, standard deviation, and Sharpe ratio of the *FS* and *S* strategies for $A=5$. These differences are computed over all strategies and across the three specification sets. The constrained *FS* strategies conserve the robustness property, as their standard deviations are always significantly lower. However, the *FS* strategies also yield a lower excess mean in all but one cases. These two effects lead to a higher Sharpe ratio for the *S* strategies in all cases, except for Positivity. The Sharpe ratio difference is significant at the overall level (−0.06), as well as for Local Combined (−0.08), Global Combined (−0.09), and Rolling 100 (−0.09). This is also the case for Free (−0.09) and Variability (−0.09), which allow the *S* weights to differ from the FS weights to a greater extent.

5.4.2. Conditional strategies based on the international CAPM

Ferson and Harvey (1993) show that much of the predictability observed in developed countries is compatible with a rational asset pricing model with time-varying betas and risk premia. Estimating conditional expected returns using this theoretical approach represents an alternative to the linear regression. Applied to currency-hedged returns, the conditional international CAPM proposed by Solnik (1974) can be written as:

$$\pi_{it} = \beta_{it} \cdot \pi_{wt}, \tag{13}$$

where β_{it} is the country i conditional beta, and π_{wt} is the world market risk premium proxied by the value-weighted portfolio of the 12 countries (Market). β_{it} is approximated as a linear function of the $(L+1)\times$ 1 vector z_{it} composed of one (for the constant term) and L predictive Local variables: $\beta_{\mathrm{it}}=z_{it}'\delta_i$. π_{wt} is modelled as a linear function of the $(Q+1)\times$ 1 vector z_{gt} composed of one (for the constant term) and Q predictive Global variables: $\pi_{wt}=\mathrm{z}_{gt}'\gamma$. Defining the error terms u_{wt} and u_{it} ($i=1,\ldots,$ 12) as:

$$\begin{aligned} u_{wt} &= r_{wt} - z_{gt}'\gamma, \\ u_{it} &= r_{it} - (z_{it}'\,\delta_i) \cdot z_{gt}'\gamma, \end{aligned} \tag{14}$$

we can use the orthogonality conditions $E(u_{wt} \cdot z_{gt})$ and $E(u_{it} \cdot z_{it})$ to estimate the coefficients $\{\gamma, \delta_1,\ldots, \delta_{12}\}$. Using Eq. (13), we compute $\widehat{\prod}_t$ across the different Local variables (Asset, Interest, Combined), Global variables (Asset, Interest, Combined), and estimation windows (Expanding, Rolling). Then, we maximize the mean–variance function in Eq. (7). The differences between the international CAPM and S conditional strategies are shown in Table 8. Similarly to Fletcher and Hillier (2003), we find that the CAPM strategies do not perform well. This result can be due to the additional structure imposed to the relation between predictive variables and excess returns. At the overall level, the CAPM Sharpe ratio is lower (−0.06). This decrease also occurs in 11 out of the 15 specification choices, and is particularly strong for Global Asset (−0.2), All (−0.16), Rolling 50 (−0.16), and Free (−0.13).

6. Conclusion

We examine the impact of specification uncertainty on the performance of international conditional asset allocation. Accounting for specification uncertainty greatly reduces the important gains previously documented for such strategies. First, different specification choices that could reasonably be made by the investor can generate very different performance. In particular, the performance variability is strong across the predictive variables and portfolio constraints. Second, the performance of the best strategies is very sensitive to minor specification changes. The conditional strategies selected with different financial criteria (the Sharpe ratio, the cumulative wealth, and the certainty equivalent) cannot outperform benchmarks that ignore conditioning information. We also find that using the efficient portfolio optimization of Ferson and Siegel (2001) and the international CAPM does not improve the performance.

Over our sample period, the statistical significance of predictability in international markets has declined (Bossaerts and Hillion, 1999), thus increasing the proportion of unprofitable versus profitable conditional strategies. Because of specification uncertainty, the investor is unable to pick the best strategies and is more likely to choose an unprofitable one. The paradox between the economic and statistical significance of predictability disappears, since the poor performance of conditional asset allocation corresponds to the low explanatory power of the predictive variables. Over our sample period, we therefore find that international equity markets satisfy the semi-strong efficiency defined by Timmermann and Granger (2004).

Acknowledgments

I am grateful to Fabien Couderc, Michel Dubois, Georges Gatopoulos, Dušan Isakov, Christophe Pérignon, Olivier Scaillet, Frédéric Sonney, René Stulz, and participants to the 2004 AFFI Meeting in Paris for their helpful comments. I especially thank Wayne Ferson (the editor), and two anonymous referees for their helpful comments. Financial support by the National Centre of Competence in Research "Financial Valuation and Risk Management" (NCCR FINRISK) is gratefully acknowledged.

Appendix A

A.1. The currency-hedged return of the country index

The total return R_{it+1} of the country i index in the investor's reference currency is equal to:

$$R_{it+1} = R^l_{it+1} + s_{it+1} + s_{it+1} \cdot R^l_{it+1}, \tag{15}$$

where s_{it+1} represents the rate of return of one currency i in terms of the reference currency, and R^l_{it+1} denotes the return of the country i index in currency i. In the absence of arbitrage, the return R^{CF}_{it+1} of a long position in the country i currency futures contract expiring at time $t+1$ is equal to:

$$R^{\mathrm{CF}}_{it+1} = \frac{R^l_{fit} - R_{ft}}{1 + R^l_{fit}} + s_{it+1}, \tag{16}$$

where R^l_{fit} denotes the country i riskfree rate, and R_{ft} the investor's reference currency riskfree rate. The currency-hedged return R^H_{it+1} of the country i index is equal to $R_{it+1} - h \cdot R^{CF}_{it+1}$, where h is the hedge ratio. Assuming that $\mathrm{cov}(R^l_{it+1}, s_{it+1}) = 0$, the hedge ratio h^* that minimizes the variance of the return differential, $R^H_{it+1} - R^l_{it+1}$, is equal to $h^* = 1 + R^l_{fit}$ (Solnik and McLeavey, 2004). Writing $r^l_{it+1} = R^l_{it+1} - R^l_{fit}$ and using the optimal ratio, R^H_{it+1} is equal to:

$$R^H_{it+1} = R_{it+1} - h^* \cdot R^{CF}_{it+1} = R_{ft} + r^l_{it+1} + s_{it+1} \cdot r^l_{it+1}. \tag{17}$$

A.2. The return of the country stock index futures contract

In absence of arbitrage, the country i index futures contract expiring at time $t+1$ can be replicated by a long position in the country i index and a short position in the country i riskfree asset. The return of the long position is given by Eq. (15). The return R_{fit+1} of the short position is given by:

$$R_{fit+1} = R^l_{fit} + s_{it+1} + s_{it+1} \cdot R^l_{fit}. \tag{18}$$

Using Eqs. (15), (18) and neglecting the cross-product $s_{it+1} \cdot r^l_{it+1}$, the return R^{IF}_{it+1} of a long position in the country i futures contract in the reference currency can be approximated by r^l_{it+1}:

$$R^{IF}_{it+1} = R_{it+1} - R_{fit+1} \approx r^l_{it+1}. \tag{19}$$

A.3. The Ferson–Khang measure in the presence of predictive information

Proposition. *The Ferson–Khang measure (FK) of the conditional asset allocation is strictly positive if and only if the investor has predictive information.*

Proof. From Eq. (9), FK is equal to $\Sigma_{i=1}^{N}$ cov(w_{it}, r_{it+1}) if the benchmark strategy does not account for predictability. Using the weight vector $w_t = \frac{1}{A}\cdot\Sigma^{-1}\prod_t$ obtained from the unconstrained optimization in Eq. (7), FK is given by:

$$\mathrm{FK} = \frac{1}{A}\mathrm{Tr}\left(\Sigma^{-1}\Omega\right), \tag{20}$$

where Ω is the $N\times N$ covariance matrix of $\prod_t$, and Tr denotes the trace operator. (If) If the investor has some predictive information, $\Omega \neq 0$. As Ω is symmetric and non-negative definite, it can be written as $\Omega=\Omega^{1/2}\Omega^{1/2}$, where $\Omega^{1/2}$ is symmetric and non-negative definite. We have $x'\Omega^{1/2}\Sigma^{-1}\Omega^{1/2}x=y'\Sigma^{-1}y$, where $y=\Omega^{1/2}x$. This expression is always strictly positive, except for cases where $y=0$. Therefore, $\Omega^{1/2}\Sigma^{-1}\Omega^{1/2}$ is non-negative definite and has either positive or zero eigenvalues. Since $\Omega\neq 0$, at least one of them is positive. Therefore, $\mathrm{Tr}(\Omega^{1/2}\Sigma^{-1}\Omega^{1/2})=\mathrm{Tr}(\Sigma^{-1}\Omega)>0$. (Only If) If FK is strictly positive, it implies that $\mathrm{Tr}(\Sigma^{-1}\Omega)>0$ and that $\Omega\neq 0$. Therefore, the investor has predictive information.

A similar result can be found in Grinblatt and Titman (1989, proposition 4). But contrary to them, conditional normality is not required, because the conditional asset allocation is determined according to Eq. (7), regardless of the return distribution. □

A.4. Computation of the bootstrapped p-values

The estimators of the difference in mean, θ_μ, standard deviation, θ_σ, and Sharpe ratio, $\theta_{\mu/\sigma}$, between K_1 conditional and K_2 benchmark strategies are equal to:

$$\begin{aligned}
\theta_\mu &= \frac{1}{K_1}\sum_{k=1}^{K_1} m_{k1} - \frac{1}{K_2}\sum_{k=1}^{K_2} m_{k1},\\
\theta_\sigma &= \frac{1}{K_1}\sum_{l=1}^{K_1}\left(m_{k2}-m_{k1}^2\right)^{\frac{1}{2}} - \frac{1}{K_2}\sum_{l=1}^{K_2}\left(m_{k2}-m_{k1}^2\right)^{\frac{1}{2}},\\
\theta_{\mu/\sigma} &= \frac{1}{K_1}\sum_{l=1}^{K_1}\frac{m_{k1}}{\left(m_{k2}-m_{k1}^2\right)^{\frac{1}{2}}} - \frac{1}{K_2}\sum_{k=1}^{K_2}\frac{m_{k1}}{\left(m_{k2}-m_{k1}^2\right)^{\frac{1}{2}}},
\end{aligned} \tag{21}$$

where $m_{k1} = \frac{1}{T}\Sigma_{t=1}^{T} r_{kt+1}, m_{k2} = \frac{1}{T}\Sigma_{t=1}^{T} r_{kt+1}^2$, and r_{kt+1} is the excess return of the strategy $k(k=1,\ldots, K_1+K_2)$. Since the estimators are smooth functions of the sample moments of the $(K_1+K_2)\times 1$ excess return vector r_{t+1}, the bootstrap is consistent (Horowitz, 2001). The distribution of each estimator, $T^{1/2}(\theta - E(\theta))$, can be approximated by the bootstrap distribution, $T^{1/2}(\theta^* - (\hat{\theta}))$, where θ^* is the bootstrap estimator, and $\hat{\theta}$ is computed with the original sample data. To compute the bootstrap distribution, we resample the $T\times(K_1+K_2)$ matrix R of the strategy return time-series[16]. For each iteration ($q=1,\ldots, Q$), we draw with replacement from the rows $\{r_{t+1}\}$ of R. From these resampled rows $\{r_{t+1}^q\}$, we create a

[16] An important feature of this approach is that we directly bootstrap R. We do not need to recompute the regression coefficient estimates at each bootstrap iteration (see White, 2000, 1105).

new matrix R^q, from which θ^q is calculated using one of the equations (Eq. (21)). This approach allows preservation of the cross-sectional correlation between the strategies[17]. To compute the p-values under the null $E(\theta)=0$, we sort all θ^q $(q=1,\ldots, Q)$ in ascending order, and determine M such that $(\theta^M-\hat{\theta})<\hat{\theta}<(\theta^{M+1}-\hat{\theta})$. The p-value is then equal to $1-\frac{M}{Q}$. In all tests, Q is set to 1000.

To compute the bootstrap distribution of the proportion P_{SD} of strategies with a positive Sharpe ratio differential, we follow the approach proposed by White (2000). The estimator of the Sharpe ratio differential SD_k of the conditional strategy k $(k=1,\ldots, K_1)$ against a benchmark B is equal to:

$$\mathrm{SD}_k = \frac{m_{k1}}{\left(m_{k2}-m_{k1}^2\right)^{\frac{1}{2}}} - \frac{m_{B1}}{\left(m_{B2}-m_{B1}^2\right)^{\frac{1}{2}}}. \tag{22}$$

Using the indicator function I_k equal to 1 if $\mathrm{SD}_k>0$ and zero otherwise, P_{SD} is equal to:

$$P_{\mathrm{SD}} = \frac{1}{K_1}\sum_{k=1}^{K_1} I_k. \tag{23}$$

For each iteration $(q=1,\ldots, Q)$, we create a new matrix R^q, and compute SD_k^q for each strategy k. Denoting by $\widehat{\mathrm{SD}}_k$ the estimate computed with the original sample data, we set I_k^q equal to 1 if $\mathrm{SD}_k^q-\widehat{\mathrm{SD}}_k>0$, and zero otherwise. P_{SD}^q is then calculated by using I_k^q in Eq. (23). To obtain the p-value under the null $E(P_{\mathrm{SD}})=0$, we sort all P_{SD}^q $(q=1,\ldots, Q)$ in ascending order, and determine M such that $P_{\mathrm{SD}}^M<\hat{P}_{\mathrm{SD}}<P_{\mathrm{SD}}^{M+1}$. The p-value is then equal to $1-\frac{M}{Q}$. Q is set to 1000. For the Ferson–Khang measure, we simply replace SD_k with $\mathrm{FK}_k=m_{k1}^0-m_{B1}^0$, where m_{k1}^0 and m_{B1}^0 denote the average value of r_{kt+1}^0 and r_{Bt+1}^0 (Eq. (10)).

References

Barras, L., 2006. Understanding the relation between the statistical and economic significance of predictability. SFI Working Paper. University of Geneva.

Bekaert, G., Harvey, C.R., 1995. Time-varying world market integration. Journal of Finance 50, 403–444.

Bossaerts, P., Hillion, P., 1999. Implementing statistical criteria to select return forecasting models: what do we learn? Review of Financial Studies 12, 405–428.

Breen, W., Glosten, L.R., Jagannathan, R., 1989. Economic significance of predictable variations in stock index returns. Journal of Finance 44, 1177–1190.

Cavaglia, S.M.F.G., Dahlquist, M., Harvey, C.R., Nieuwland, F., Rathjens, P.L., Wilcox, J.W., 1997. Emerging/Developed Market Portfolio Mixes, Emerging Market Quarterly Winter, pp. 47–62.

Chopra, V.K., Ziemba, W.T., 1993. The effect of errors in means, variances, and covariances on optimal portfolio choices. Journal of Portfolio Management 20, 6–11.

Cooper, M., Gulen, H., 2006. Is time-series based predictability evident in real time? Journal of Business 79, 1263–1292.

Davison, A.C., Hinkley, D.V., 1997. Bootstrap Methods and Their Application. Cambridge University Press, London.

Dybvig, P.H., Ross, S.A., 1985. Differential information and performance measurement using a security market line. Journal of Finance 40, 383–399.

Fama, E.F., French, K.R., 1988. Dividend yields and expected stock returns. Journal of Financial Economics 22, 3–25.

Fama, E.F., French, K.R., 1989. Business conditions and expected returns on stocks and bonds. Journal of Financial Economics 25, 23–49.

Ferson, W., Harvey, C.R., 1993. The risk and predictability of international equity returns. Review of Financial Studies 6, 527–566.

[17] Blockbootstrap can be used to account for return autocorrelation. The block length is set to $n^{1/5}$ for one-sided tests (Hall et al., 1995). We implemented this methodology, but the difference is negligible.

Ferson, W., Khang, K., 2002. Conditional performance measurement using portfolio weights: evidence for pension funds. Journal of Financial Economics 65, 249–282.
Ferson, W., Siegel, A.F., 2001. The efficient use of conditioning information in portfolios. Journal of Finance 56, 967–982.
Ferson, W., Siegel, A.F., Xu, P.T., 2006. Mimicking portfolios with conditioning information. Journal of Financial and Quantitative Analysis 41, 607–635.
Fletcher, J., Hillier, J., 2003. An examination of linear factor models in country equity asset allocation strategies. Quarterly Review of Economics and Finance 209, 1–16.
Frost, P.A., Savarino, J.E., 1988. For better performance: constrain portfolio weights. Journal of Portfolio Management 14, 29–34.
Goyal, A., Welch, I., in press. A comprehensive look at the empirical performance of equity premium prediction. Review of Financial Studies.
Grinblatt, M., Titman, S., 1989. Portfolio performance evaluation: old issues and new insights. Review of Financial Studies 3, 393–413.
Grinblatt, M., Titman, S., 1993. Performance measurement without benchmarks: an examination of mutual fund returns. Journal of Business 66, 47–68.
Hall, P., Horowitz, J.L., Jing, B.-Y., 1995. On blocking rules for the bootstrap with dependent data. Biometrika 82, 561–574.
Han, Y., 2005. Can an investor profit from return predictability in real time? Working Paper. Tulane University.
Handa, P., Tiwari, A., 2006. Does stock return predictability imply improved asset allocation and performance? Journal of Business 79, 2423–2468.
Harvey, C.R., 1991. The world price of covariance risk. Journal of Finance 46, 111–157.
Harvey, C.R., 1994. Conditional asset allocation in emerging markets. Working Paper 4623, NBER.
Horowitz, J.L., 2001. The bootstrap. In: Heckman, J.J., Leamer, E.E. (Eds.), Handbook of Econometrics, vol. 5. North-Holland, Amsterdam.
Jagannathan, R., Ma, T., 2003. Risk reduction in large portfolios: why imposing the wrong constraints helps. Journal of Finance 58, 1651–1683.
Jobson, J.D., Korkie, B., 1981. Putting Markowitz Theory to work. Journal of Portfolio Management 7, 70–74.
Kandel, S., Stambaugh, R.F., 1996. On the predictability of stock returns: an asset-allocation perspective. Journal of Finance 17, 357–390.
Keim, D.B., Stambaugh, R.F., 1986. Predicting returns in the stock and bond makets. Journal of Financial Economics 17, 357–390.
Klemkosky, R.C., Bharati, R., 1995. Time-varying expected returns and asset allocation. Journal of Portfolio Management 21, 80–87.
Ledoit, O., Wolf, M., 2003. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. Journal of Empirical Finance 10, 603–621.
Ledoit, O., Wolf, M., 2004. Honey, I shrunk the sample covariance matrix. Journal of Portfolio Management 30, 110–119.
Marquering, W., Verbeek, M., 2004. The economic value of predicting stock index returns and volatility. Journal of Financial and Quantitative Analysis 39, 407–429.
Nelson, C.R., Kim, M.J., 1993. Predictable stock returns: the role of small sample bias. Journal of Finance 48, 641–661.
Paye, B.S., Timmermann, A., 2006. Instability of return prediction models. Journal of Empirical Finance 13, 274–315.
Pesaran, M.H., Timmermann, A., 1995. Predictability of stock returns: robustness and economic significance. Journal of Finance 50, 1201–1228.
Pettenuzzo, D., Timmermann, A., 2005. Predictability of stock returns and asset allocation under structural breaks. Working Paper. University of California San Diego.
Solnik, B., 1974. An equilibrium model of the international capital market. Journal of Economic Theory 8, 500–524.
Solnik, B., 1993. The performance of international allocation strategies using conditioning information. Journal of Empirical Finance 1, 33–55.
Solnik, B., McLeavey, D., 2004. International Investments5th edition. Addison-Wesley, New-York.
Stambaugh, R.F., 1999. Predictive regressions. Journal of Financial Economics 54, 375–421.
Sutcliffe, C.M.S., 2006. Stock Index Futures, 3nd edition. Ashgate, Hampshire.
Timmermann, A., Granger, C., 2004. Efficient market hypothesis and forecasting. International Journal of Forecasting 20, 15–27.
White, H., 2000. A reality check for data snooping. Econometrica 68, 1097–1126.
Xu, Y., 2004. Small levels of predictability and large economic gains. Journal of Empirical Finance 11, 247–275.