Are Multi-Strategy Hedge Funds Able to Time Strategies?

Master's Thesis

Thesis presented by
Antoine GUSTIN

Supervisor
Rajna N. GIBSON BRANDON (UNIGE)

Supervisor
/

Reader
Alain DE CROMBRUGGHE

Academic year 2017-2018

In order to obtain the Joint Degree
Master 120 en Sciences économiques, Orientation générale, Finalité spécialisée (UCL/UNamur)
and
Maîtrise en sciences économiques (UNIGE)

Ecole d'économie de Louvain / UCL • Place Montesquieu 3 • 1348 Louvain-la-Neuve / Belgium
Département des Sciences économiques / UNamur • Rempart de la Vierge 8 • 5000 Namur / Belgium
University of Geneva / GSEM – Uni Mail, Bd du Pont d'Arve 40 – CH-1211 Genève / Suisse

Master's Thesis:
Are Multi-Strategy Hedge Funds Able to Time Strategies?

ANTOINE GUSTIN

THESIS SUPERVISOR: RAJNA N. GIBSON BRANDON

June 4, 2018

Abstract

We study multi-strategy hedge funds and their supposed ability to time strategies. Indeed, investors can expect these funds to exhibit significant skills in predicting their strategies' returns and consequently reallocating their capital across these strategies to generate additional performance. We evaluate the existence of such specific skills using a small sample of 90 multi-strategy funds from the TASS database, on the period spanning from January 1994 to December 2017. Building on the previous timing literature, we design two types of econometric models which both allow for the presence of *strategy timing* skills. Our results show that multi-strategy managers do not possess *true* strategy timing ability, i.e. spanning from superior information. However, they are able to react to publicly available signals to optimally reallocate capital across strategies. We conjecture that their flexibility in switching strategies ultimately constitutes their main advantage over funds-of-hedge-funds. We do observe some evidence of *true* strategy timing abilities, both at the aggregate and fund levels, but these are later shown to be the consequence of *luck* on the part of some managers. Our results are robust to various additional tests, and we conjecture that the presence of upward data biases in our sample eventually increases the external validity of this study.

Contents

Acknowledgements

I would first like to warmly thank Prof. Rajna N. Gibson Brandon for her support and help throughout the writing of this Master's thesis, as well as for her classes on Hedge Funds which inspired me to find this research question.

I also thank Mr. Sylvain Frénéat from Thomson Reuters, who generously spent his time to help me get the data needed for my empirical analyses.

Finally, I would like to thank Donatienne D'hoest for her precious help in improving the clarity of this thesis, as well as Luchos Hernandez, Nikos Manolopoulos, and Eduardos Diaz for their support and encouragements.

I. Introduction

Hedge funds are private collective investment vehicles which are loosely regulated – relatively to classic investment schemes such as mutual funds – and which thus have a lot of flexibility in the way they manage portfolios and interact with their investors. Over the last decades, the hedge fund industry has undergone an impressive development, which has attracted the attention of the academic research among others. In particular, researchers have sought to evaluate the risk-adjusted performance of these alternative funds, and many models have now been developed for this purpose, both for evaluation at the aggregate level or in a specific hedge fund style[1].

Indeed, hedge funds are generally divided in categories, which are based on their investment style or strategy. Among these, multi-strategy hedge funds represent a quite recent but nonetheless important segment, as they steadily gained in popularity since 2003 and now represent around 17% of assets under management in the industry[2]. In opposition to other hedge fund strategies, which generally define their focus market or trading scheme quite clearly[3], multi-strategy hedge funds trade in multiple markets and use multiple strategies, in an opportunistic manner and at the discretion of the manager. Note that this particularity is probably the main reason why the literature has very rarely focused on this hedge fund style, since it makes performance evaluation difficult and less insightful.

On top of diversification benefits, a multi-strategy fund can thus be expected to generate additional risk-adjusted performance by dynamically investing in different strategies, depending on their respective profitability over time. Indeed, if the goal of the multi-strategy framework is simply to provide diversification benefits to the investor, then it holds no real advantage over a typical fund-of-hedge-funds. As explained by Agarwal and Kale (2007), this supposed ability to time strategies – that is, to optimally switch between them as their respective profitability varies – is thus the main selling point of multi-strategy hedge funds, and can indeed be valuable since hedge fund strategies are known to exhibit time-varying returns. As an illustration, we provide in [Appendix A] some mission statements of multi-strategy funds which explicitly refer to this dynamic allocation of capital between strategies.

The main goal of this Master's thesis is thus to study whether or not multi-strategy funds indeed possess the ability to time strategies, and if so, how valuable is this ability. Through this research, we hope to contribute to the literature on multi-strategy funds, which until now remains quite scarce. We also hope to contribute to the timing literature in hedge funds, since to the best of our knowledge, this is the first paper to focus on the timing abilities of multi-strategy hedge funds. This question should also be of significant interest to hedge fund investors, since this supposed "strategy timing" ability accounts for an important part of the intrinsic attractiveness of multi-strategy funds.

[1]See, for example: Agarwal, Fung, Loon and Naik (2011), Agarwal and Naik (2000, 2004), Fung and Hsieh (1997, 2001, 2002, 2004, 2005), Mitchell and Pulvino (2001).

[2] Credit Suisse Hedge Fund Indexes, 2018.

[3] Common hedge fund categories are: long/short equity, equity market neutral, dedicated short bias, fixed income arbitrage, convertible arbitrage, global macro, event driven, managed futures and CTAs, emerging markets, and multi-strategy. Subcategories are also usually defined, such as trend following or statistical arbitrage for the global macro and managed futures categories, distressed or risk arbitrage for the event driven category.

We study this research question using a small sample of multi-strategy hedge funds from the Thomson-Reuters Lipper TASS database on the period 1994-2017. We build on the Treynor and Mazuy (1966) and the Henriksson and Merton (1981) frameworks to construct models which better accommodate strategy timing skills. We also design our own strategy timing factor, which is added to the Fung and Hsieh (2004) 8-factors model. We test for the presence of strategy timing abilities using both types of models, and conduct additional robustness tests, in particular to see whether managers also react to publicly available information, and if detected timing ability can be attributed to luck.

Overall, our results show that multi-strategy managers do not possess *true* strategy timing ability. However, they are able to react to publicly available signals, i.e. lagged information, to optimally adjust their exposures to strategies. We conjecture that their flexibility in switching strategies, e.g. in comparison to funds-of-hedge-funds, is in fact central in this process. We do observe some evidence of true strategy timing abilities, both at the aggregate and fund levels, but these are later shown to be the consequence of sampling variation, i.e. *luck* on the part of some managers. These results are robust to various additional tests, and we conjecture that the presence of upward data biases in our sample eventually increases the external validity of this study.

In the next section, we review the existing literature on timing abilities in the hedge fund industry. In section 3, we present the data, discuss its inherent biases, and conduct a first analysis of the investment styles of multi-strategy funds in our sample. In section 4, we present the underlying methodology of our empirical analysis by reviewing the theoretical models we use as well as their estimation procedure. In section 5, we present and discuss our empirical results at the aggregate level. In section 6, we conduct tests at the fund level. Section 7 will conclude.

II. Literature Review

The study of timing abilities initially began in the mutual funds industry, with the pioneering work of Treynor and Mazuy (1966). Market timing skill is then defined as the ability to outguess the market, that is, to correctly anticipate upward and downward movements in the stock market. Since then, several timing models have been developed[4], and overall academic research has shown that mutual funds do not possess market timing skills. More recently, researchers have applied the same models to hedge funds, since their peculiar characteristics and in particular, their investment flexibility should imply that they would be more easily able to profit from any timing ability they could possess. In addition, hedge fund managers often use these timing skills as marketing arguments.

On this subject, Chen (2006) conduct an overall study of the timing abilities of hedge funds in their focus markets, and find significant timing ability for several strategies. They also find evidence of selectivity skills, and conclude that the latter has a stronger presence among hedge funds than timing ability. Chen and Liang (2007) focus on the market return and volatility timing abilities of self-described market timing funds. They find evidence of market timing ability both at the aggregate and fund level, and especially in bear and volatile market states. Cao, Chen, Liang, and Lo (2013) study the liquidity timing abilities of hedge funds, and find strong evidence of liquidity timing at the fund level, the effect being both statistically and economically significant. These papers follow a similar approach in that they use the Treynor and Mazuy (1966) framework, and similar models such as the one developed by Henriksson and Merton (1981), as the basis for their empirical analysis. This thesis is in the continuity of this specific literature.

In a similar fashion, Fung, Xu, and Yau (2002) study the performance of global equity hedge funds and use the Henriksson-Merton timing model to distinguish stock selection ability from market timing ability. Some researchers have also studied timing abilities in the hedge fund industry by using alternative approaches. Agarwal, Ruenzi, and Weigert (2017) study hedge funds' exposure to tail risk and, among others, their ability to time it, using the Lehman Brothers bankruptcy as an exogenous event. Griffin and Xu (2009) study the performance of a sample of hedge funds using their disclosed positions through SEC 13F filings, including their ability to time the market, which they find is insignificant. Cai and Liang (2012) use a dynamic linear regression model to generate time-varying alphas and betas for their sample of hedge funds, and then study the correlation between the dynamic betas and their respective factors, finding positive and significant market timing, volatility timing and liquidity timing abilities for some hedge fund strategies in the process. Jiang (2003) develops a nonparametric test for market timing ability which is able to distinguish the quality of the timing information from the aggressiveness with which the manager reacts to such information.

A specific strand of the literature also focuses on existing issues in the classic market timing models. Jagannathan and Korajczyk (1986) argue that the presence of options in a managed portfolio will lead the Treynor and Mazuy (1966) and the Henriksson and Merton (1981) models to detect significant timing ability and a negative stock selection ability, resulting in the two being negatively correlated overall. They thus conclude that option trading should be controlled for when evaluating the timing skills of managers. Ferson and Schadt (1996) argue that true timing ability should rely on superior

[4] See, among others, Jensen (1972), Henriksson and Merton (1981), Admati, Bhattacharya, Pfleiderer, and Ross (1986), Jagannathan and Korajczyk (1986), Lehmann and Modest (1987), Ferson and Schadt (1996), Goetzmann, Ingersoll, and Ivkovich (2000).

information and that, as such, market timing models need to account for publicly available information. They derive "conditional" versions of the classic market timing models that are thus able to distinguish mere reaction to public signals from true timing skill. Goetzmann, Ingersoll, and Ivković (2000) study the efficiency of the Henriksson and Merton (1981) market timing measure when the frequency at which investment decisions are made is inferior to that at which performance is measured. They show through simulations that in this case the Henriksson-Merton measure is weak and strongly biased downward. They also propose a way to partially correct this problem using an instrumental variable estimator. Getmansky, Lo, and Makarov (2004) study the apparent autocorrelation in hedge fund returns, and argue that the phenomenon is mostly due to illiquid trading and return smoothing. They show that the latter will induce important biases in reported returns, in the form of lower variance and spurious market neutrality, and derive an econometric model of return smoothing which allows to partly correct the issue. Finally, Baquero, Ter Horst, and Verbeek (2005) study persistence in hedge fund returns at the quarterly level, and focus on the look-ahead bias which, as they show, may significantly alter the results of such an analysis. They model the conditional survival probabilities of hedge funds and subsequently derive a weighting scheme to control for the look-ahead bias, which also depends on unconditional survival probabilities.

As mentioned in the previous section, the literature specifically focusing on multi-strategy hedge funds remains quite scarce, if not non-existent. To the best of our knowledge, Agarwal and Kale (2007) provide the only published study of this hedge fund category to this day. The authors investigate the relative performance of multi-strategy hedge funds and funds-of-hedge-funds since, as they argue, both investment schemes should yield similar returns in well-performing markets as they essentially provide the same benefits to investors. They find that the former significantly and consistently outperform the latter, both on a net-of-fees and gross-of-fees basis. They observe that this outperformance remains after controlling for differences in agency risk and investment flexibility, thereby leading them to conclude that the self-selection of skilled managers into multi-strategy hedge funds must be the source of this superior relative performance.

III. The Data

We now present the data we use in our empirical tests. We first review well-known biases inherent to hedge fund return series, and then discuss their relevance for this work. We also explain our data filtering choices, and present summary statistics for the multi-strategy funds in our sample. Finally, we conduct an analysis of the funds' investment styles, in order to obtain insights on the various strategies they might use in practice.

Hedge funds returns data is known to exhibit several potentially important biases. Indeed, as hedge funds are often structured as (offshore) private investment vehicles, they are exempted from usual reporting obligations, and as such the data is generated on a voluntary basis. Most of the biases are a direct or indirect consequence of this fact. A thorough study of these biases can be found in Fung and Hsieh (2000), of which we hereunder provide a short summary.

The first bias to take into consideration is the ex-ante selection bias[5], arising from the fact that hedge funds may choose whether or not to report to a database. If funds that do not report exhibit different characteristics than the ones that do, then performance measurement based on the observable sample of funds will be biased. However, whether this issue biases the overall returns upward or downward is unclear; funds that have good performance may report (to attract new investors), but may as well not report (to avoid attracting new investors, because capacity is limited).

A second important bias is the well-documented survivorship bias, which is a type of ex-post selection bias. Prior to 1994, funds that were liquidated or stopped reporting were deleted from the databases along with their return histories. As one can expect the performance of surviving funds to be better than that of defunct funds, the survivorship bias is positive. It is also particularly relevant for the hedge fund industry, where survival rates have been showed to be low relatively to more classic investments schemes[6]. Estimates of the survivorship bias for hedge funds usually range between 2-3% per year. An obvious way to correct the bias is to include defunct funds in the empirical analysis.

Another ex-post bias is the backfill or instant history bias, which arises when managers backfill the returns of their fund (since its creation) when first reporting to a database. This is problematic as hedge fund managers often run their fund for a "test period", that could be as long as 2 or 3 years, before opening the fund to investors. As such, the managers only goes on running his fund if the performance over the test period is satisfying; as such the backfilling process biases overall returns upward in the databases. Estimates of the bias magnitude are diverse; Fung and Hsieh (2000) estimate it to be about 1.4% per year, but some more recent studies estimate it to be as high as 5%. Authors usually correct the bias by deleting a fixed number of months at the start of each return series (typically 12 to 24 months), or delete all the months prior to the inception date, if the latter is available.

Other biases exist, but are usually attributed less magnitude. The end-of-life bias arises when funds which perform badly stop reporting some months prior to liquidation, so that the potentially worst returns are excluded from the database (as with LTCM in October 1997). The multi-period sampling bias arises when funds are required to have some minimum return history for research needs. The magnitude of this latter bias has

[5] Also sometimes referred to as "non-reporting bias".

[6] See, for example, Fung, Hsieh, Naik and Ramadorai (2008).

however been shown to be small enough so that it can be ignored (see Fung and Hsieh (2000)).

Our data is obtained from the Thomson-Reuters Lipper TASS database. The initial sample consists of 142 active multi-strategy funds, for which we have data over the period spanning from the 1st of January 1994 to the 31st of December 2017. Unfortunately, we were not able to access the full database, which leads to some important issues. First, the access to the "graveyard" database – that is, to data concerning defunct funds – was restricted, so that our sample only contains funds that were active on the 1st of January 2018. As of that date, the TASS database contained 1900 multi-strategy funds, of which 1504 were defunct and thus inaccessible to us. Obviously, this means our analysis will be prone to survivorship bias, which, as previously explained, can be quite substantial in the hedge fund industry. Second, some data items were also restricted, the most important of which being the funds' reported returns, usually net-of-fee. We thus use the funds' reported net asset value (hereafter NAV) over time to compute returns manually; which implies these are gross-of-fee returns. Moreover, as data on fees were incomplete in most cases, we were not able to deduct fees and expenses manually afterwards. This is also an important issue as it means we analyse the base performance of multi-strategy hedge funds, rather than the performance those funds are able to pass on to their investors. Another notable restricted data item is the funds' assets under management (hereafter AUM) over time, on which we have data for a subset of funds only.

Following the hedge fund literature, we delete the funds that are denominated in currencies other than the U.S. dollar, leaving us with the original sample of 142 funds. We also require funds to have a NAV history of at least 24 months to obtain satisfying statistical properties in future linear regressions[7]. We do not, however, impose the usual restriction that the funds' AUM must be at least USD 10 million, since we do not possess AUM data for all the funds. This should not, in theory, have a significant impact on our results[8]. Finally, because our sample size is already relatively small, we do not attempt to correct for the backfill bias by deleting the first two or three years of each fund's return history[9].

We perform an analysis of the funds' mission statements to detect possible further issues in the sample, as detailed in [Appendix B]. While doing so, we also analyse the disclosed focus markets and underlying strategies of the funds (results are discussed below). We find 22 groups of funds that are seemingly duplicates, for a total of 66 funds. These are deemed "duplicates" in that they share the same name, and probably belong to the same investment company or manager. As it may well be that these duplicates follow different strategies and thus exhibit distinguishable returns features, we perform a linear correlation analysis to identify "real" duplicates. The details of this process are presented in [Appendix C]. We end up deleting 42 funds that are deemed to be "real" duplicates. Applying all these filters together, our final sample consists of 90 funds.

---

[7] As authors in the hedge fund literature usually require either 24 or 36 months of minimum return history for funds in their sample, we later change this requirement to 36 months as an additional robustness test.

[8] This size criterion is usually imposed to mitigate possible biases with small funds (ex. backfill bias) (Chen (2006)). Moreover, authors often report that changing the AUM threshold or removing it does not affect their results in any significant way, see for example Chen (2006) or Cao, Chen, Liang, Lo (2013).

[9] Since we require funds to have a return history of at least 24 months, deleting the first 24 or 36 months of each return series would force us to drop additional funds from the sample.

Panel A of [table 1] contains summary statistics for the final sample of funds. Return statistics are computed for all funds, as well as for an equally-weighted portfolio of funds. We also simulate the value of this equally-weighted portfolio over the sample period and plot the result in [figure 1]. Surprisingly, the average fund in our sample earned a monthly gross-of-fee return of about 0.75%, which is actually not higher than what was reported in the previous literature[10]. To mitigate the possible impact of the 2008 "subprime" crisis, we exclude the period spanning from September to December 2008 from our return series and compute the average monthly gross-of-fee return once more. We obtain a second estimate of 0.82%, which is still consistent with what other authors have reported. It is thus possible that the survivorship bias is of relatively low importance in this particular sample. Still, the months following the Lehman Brothers bankruptcy seem to have been rough for multi-strategy funds; as an illustration, [figure 2] contains the returns series during that period.

Also, and as one might have expected, the overall return distribution exhibits fat tails, and positive skewness. We also find positive autocorrelation overall, hinting at illiquid trading among multi-strategy hedge funds, as explained by Getmansky, Lo, and Makarov (2004). Finally, the average fund in our sample has existed (and/or reported to the database) for around 9 years and 5 months.

Panel B of [table 1] contains summary statistics for the Fung and Hsieh (2004) model variables that we use in our empirical tests. We present this specific model and its associated variables in section 4.

[Table 2] presents the results of the analysis of the funds' mission statements. First, it seems that 23.3% of funds (21 funds) focus on one specific market or strategy only; which is unexpected as multi-strategy funds would typically invest in multiple markets or strategies, as explained above. Also, 6 funds state they invest either primarily or solely in other funds, which would make them funds-of-funds. This leads us to think the following: either the TASS classification scheme is imprecise[11], or these funds underwent some level of mission drift and did not update their mission statements in the meantime. To avoid making any arbitrary choice, we keep these funds into the sample[12]. Also, this suggests that conducting a mission drift analysis on multi-strategy funds would be interesting, but we leave this topic to future research.

[Table 2] also gives some insights about the focus markets and/or main strategies of the multi-strategy funds in our sample. We see that a relatively high proportion of funds declare investing in equity (43.3%) and fixed-income (33.3%) markets or strategies. The analysis, unfortunately, does not yield meaningful conclusions for other markets or strategies. We note, however, that the "commodities / CTAs", "currencies" and "statistical

[10] For example, Agarwal and Kale (2007) report an average monthly gross-of-fee return of about 1.21% for multi-strategy funds, and an average net-of-fee return of about 0.97%, over the period 1994-2004. Agarwal, Ruenzi, and Weigert (2017) report an average monthly, net-of-fee and excess return of about 0.52% over the period 1996-2012.

[11] Lipper TASS states that "The classification identifies those hedge funds that pursue a diversified investment strategy whereby a single investment process does not account for more than 75% of risk capital. [...] In many cases Multi-Strategy hedge funds are constructed as funds of hedge funds, combining different strategies/styles and also distributing the assets among various managers. Funds of hedge funds pursuing Multi-Strategy are identified with the usual "Fund of Funds" flag to maintain a distinction with single-or multi-manager Multi-Strategy hedge funds". This should imply that no fund-of-funds should be present in our sample, as we specifically excluded the funds labelled "FoHF/multi-strategy" when extracting the data.

[12] We will, however, remove them later on as a robustness test.

/ quantitative arbitrage" markets and/or strategies are cited slightly more often than the rest (respectively, 16.7%, 15.6%, and 13.3%). It is also interesting to note that a low number of funds declare using convertible arbitrage strategies (7.8%), while Lipper TASS stated that "many Multi-Strategy managers began as convertible arbitrage managers that diversified into other strategies". Overall, funds which refer to specific hedge fund strategies in their mission statement are quite scarce, which was to be expected.

In order to get further insights on the strategies in which our multi-strategy funds invest, we regress the equally-weighted portfolio's returns on the Credit Suisse/Tremont (CS) and Hedge Fund Research (HFR) hedge fund indices, as well as on the Fung and Hsieh (2004) eight-factor model. The results are presented in [table 3]. Note that the CS indices use a definition of hedge fund strategies which is very close to the one used by Lipper TASS. The HFR indices, however, are very strategy-specific and are thus aggregated into broader categories[13]. The Fung and Hsieh (2004) model is made of a set of "asset-based style" factors, that is, factors which aim to replicate the dynamic and option-like features of hedge fund returns by using standard and liquid asset classes. It has been shown to explain the returns of diversified hedge fund portfolios well; the details of this model are discussed in section 4 since we also use it for our empirical analyses.

The regression of the equally-weighted portfolio on CS indices (panel A of [table 3]) shows that the multi-strategy funds in our sample are mainly exposed to the risk premia of "convertible arbitrage" and "long/short equity" strategies. They also have medium exposure to the risk premium of "managed futures" strategies. The results remain similar when HFR indices are used instead (panel B): our multi-strategy funds are primarily exposed to the risk premia of "equity hedge", "macro", and "convertible arbitrage" strategies. The results are comparable in that the "equity hedge" and "macro" categories of HFR respectively encompass the "long/short equity" and "managed futures" categories of CS. The significant exposures to convertible arbitrage strategies are quite interesting, since this observation is more in line with the supposed origins of multi-strategy hedge funds as stated by Lipper TASS. Although these funds seem, indeed, to have diversified into different strategies, they also seem to keep on investing in convertible arbitrage strategies. The regression on the Fung and Hsieh (2004) factors (panel C) yields complementary results. Exposures to the equity and fixed-income markets are highly significant, and it is also the case for the emerging markets factor.

These results lead us to believe that multi-strategy funds, at least in this particular sample, mainly trade in equity, fixed-income, and convertible arbitrage markets. The exposures to "macro" and "managed futures" strategies revealed by the CS and HFR indices regressions are harder to interpret; one possible explanation is that the multi-strategy funds in our sample hold some exposure to the commodities markets, as global macro and managed futures[14] strategies are typically linked to commodities trading. Finally, our multi-strategy funds probably hold exposure to emerging markets as well.

[13] For example, long/short equity, equity market neutral, and dedicated short bias hedge funds all fall into the "equity hedge" category.

[14] Also called "commodities trading advisors" or "CTAs".

IV. Methodology

In this section, we provide an overview of the models which we use in our empirical analyses. We begin by reviewing the market timing models of Treynor and Mazuy (1966, hereafter TM) and Henriksson and Merton (1981, hereafter HM). We then define "strategy timing", and conjecture how multi-strategy managers would implement it in practice. Next, we develop extended versions of the TM and HM models that better accommodate the presence of strategy timing abilities. We then briefly review the performance evaluation model which we use, and discuss how our previously developed strategy timing frameworks can be applied to it. After briefly discussing possible issues inherent to the models we developed, we create a strategy timing factor which builds on our previous assumptions about this timing process, and discuss its characteristics. Finally, we detail the estimations procedures that we use in our empirical analyses.

a. *Market timing models*

Market timing models describe "how a fund manager possessing superior information about the market should adjust his portfolio's market exposure when he observes a signal about future market return", as defined by Chen and Liang (2006). Early market timing models generally build on the Capital Asset Pricing Model (CAPM) framework, as is the case with the TM and HM models. The underlying assumption is thus that the fund's returns are generated according to the following process:

$$r_{t+1} = \alpha + \beta r_{m,t+1} + \varepsilon_{t+1} \tag{1}$$

Where r_{t+1} is the fund's excess return over the next period, $r_{m,t+1}$ is the market portfolio's excess return over the next period, and ε_{t+1} is the error term. The fund is then characterized by its exposure to the market, β, and its abnormal return, α, which is Jensen's alpha.

The basic CAPM framework assumes that the market exposure is fixed through time. However, if the fund manager adjusts his exposure in an attempt to time the market based on superior information, then the model is subject to misspecification and the α measures neither selectivity skill nor market timing skill, as explained by Jensen (1972). In their pioneering work, Treynor and Mazuy (1966) address this issue by developing a new model in which the market exposure is allowed to vary through time:

$$r_{t+1} = \alpha + \beta_t r_{m,t+1} + \varepsilon_{t+1} \tag{2}$$

$$\text{where, } \beta_t = \beta_0 + \gamma(r_{m,t+1} + \eta_{t+1}) \tag{3}$$

Here, the manager is assumed to set the fund's market exposure at the beginning of the period, according to his forecast about the market excess return over the next period, $(r_{m,t+1} + \eta_{t+1})$. Note that this term depends on an error term so that the forecast may prove imprecise. The market exposure is defined by a base level, β_0, and evolves linearly with the predicted market excess return, depending on the coefficient γ. Note also that this assumption of a linear relationship between the market beta and the manager's forecast is precisely the central characteristic of the TM model. By substituting equation (3) into equation (2), we obtain the TM model:

$$r_{t+1} = \alpha + \beta_0 r_{m,t+1} + \gamma r_{m,t+1}^2 + \varepsilon'_{t+1} \tag{4}$$

Where the forecast error is integrated into the original error term to obtain the new error term ε'_{t+1}, which is assumed to be zero-mean and independent from the market excess return. Estimating this model allows to directly measure the market timing skills of the manager by looking at the estimate of the timing coefficient, γ.

Henriksson and Merton (1981) develop another variant of this framework, by assuming that the manager only predicts whether the market excess return will be positive or negative, and not the extent to which it will be. As such, instead of equation (3), the time-varying market beta is defined in the HM model as follows:

$$\beta_t = \beta_0 + \gamma\, I(r_{m,t+1} + \eta_{t+1} > 0) \quad (5)$$

Where I is and indicator variable which takes the value 1 if the predicted market excess return is positive, and the value 0 otherwise. As such, by substituting equation (5) into the fund's return equation (2), we obtain the HM model:

$$r_{t+1} = \alpha + \beta_0 r_{m,t+1} + \gamma \max(r_{m,t+1}, 0) + \varepsilon'_{t+1} \quad (6)$$

The timing variable is thus actually a call option[15] on the market excess return, with a strike price set at zero. The HM model can thus be viewed as a simplified version of the TM model, since it restricts the manager's forecast to concern the sign of the market excess return only[16]. Also, previous studies on timing skills in the hedge fund industry generally found little difference between the results obtained through the TM and HM models[17]. Nonetheless, we use both specifications in our analysis so that our results are hopefully not too dependent on these underlying assumptions.

b. Defining "strategy timing" ability

We now focus on what we call "strategy timing" abilities. As explained before, multi-strategy hedge fund managers can be expected to time their strategies over time. At this point, however, it is necessary to make some assumptions about how this process would take place.

We assume that the multi-strategy manager makes his decisions for each month in two main steps: first, he decides on how to allocate capital to each strategy for the following month, and second, he goes on making investment decisions inside each strategy. Of course, we do not assume that every multi-strategy hedge fund is led by one lone manager; it may well be that the "orchestra leader" – the main manager – makes the capital allocation decision, and then sub-managers go on running their strategy with the capital they have been allowed to use. In fact, most multi-strategy hedge funds seem to follow a "manager-of-managers" structure or a similar one[18]. The main point in this assumption is that the "capital allocation" decision is a separate process from the strategies' individual management. As such, when the "main" manager makes the capital allocation decision, he considers his available strategies as individual assets, and makes

[15] Alternatively, the variable can be defined as a put option. The assumption is then that the market beta is reduced when the market excess return is negative, that is, β_0 is the market beta when the market excess return is positive.
[16] Coggin, Fabozzi, and Rahman (1993) make this point as well, explaining that the HM model "provides no significant advantage over the Bhattacharya and Pfleiderer (1983) model", which they consider is a refined version of the TM model. Dybvig and Ross (1985) also consider it to be an important weakness of the HM model, since "it only tests whether the manager has special information, and does not test whether the manager uses the information correctly".
[17] See for example Chen (2006), Chen and Liang (2006).
[18] Lipper TASS states that "in many cases Multi-Strategy hedge funds are constructed as funds of hedge funds, combining different strategies/styles and also distributing the assets among various managers".

forecasts about each strategy's return over the next period. Consequently, we can use the TM and HM timing frameworks to estimate strategy timing abilities, as long as the factors we use represent the returns of different hedge fund strategies. Our model will thus have the following generic form:

$$r_{t+1} = \alpha + \sum_{k=1}^{K} \beta_{k,t} f_{k,t+1} + \varepsilon_{t+1} \tag{7}$$

Where r_{t+1} is the fund's excess return over the next period, $f_{k,t+1}$ the excess return of strategy $k = 1,2,\dots,K$ over the next period, and $\beta_{k,t}$ the fund's exposure to strategy k, which directly depends on how much capital does the "main" manager allocate to strategy k at the start of the period. Note that the beta is allowed to vary through time.

We now take the example of a perfect "strategy timer" to illustrate the concept. This talented manager is specialized in two hedge fund strategies, the excess returns of which we denote by f_1 and f_2, respectively. Each month, this manager is thus able to perfectly predict which strategy will yield the highest return, consequently, he invests all the fund's capital into the best-performing strategy, as shown in the table here.

A perfect strategy timer's returns

$f_{1,t+1}$	$f_{2,t+1}$	r_{t+1}
5%	7%	7%
4%	2%	4%
4%	8%	8%
6%	3%	6%
-10%	-2%	0%

An important implication is that the exposure to each strategy – that is, the capital allocated to it – ultimately depends not only on the predicted return of that strategy, but also on how it compares to the predicted return of the other strategy. Mathematically, we can write:

$$\beta_{k,t} = g(f_{1,t+1}, f_{2,t+1}), k = 1,2$$

Where g is an undefined function. Of course, in our perfect strategy timer example, this function is easily defined: it is an indicator variable which takes the value 1 if the excess return of strategy k is superior to that of the other strategy, and if it is positive, otherwise the variable takes the value 0.

$$\beta_{1,t} = I(f_{1,t+1} > f_{2,t+1}, f_{1,t+1} > 0)$$
$$\beta_{2,t} = I(f_{2,t+1} > f_{1,t+1}, f_{2,t+1} > 0)$$

This highlights a very important point: "strategy timing" as we defined it until now is actually the sum of two distinguishable timing processes. The first one relates to the first condition in the indicator variables above: it is the process of forecasting which strategy will yield the highest return, and trying to time it, by increasing exposure to the best-performing strategies, and decreasing exposure to the worst-performing ones. We will specifically refer to this as "relative strategy timing". The second timing process relates to the second condition in the indicator variables above: it is the process of forecasting whether or not the strategies will yield positive excess returns, and trying to time it, by switching to risk-free investments when the strategies yield negative excess returns. We hereafter refer to this specific process as "absolute strategy timing".

The example of the perfect strategy timer illustrates this distinction. In the first four rows of the table, the excess returns of both strategies are positive, so that the manager only engages in "relative strategy timing". In the last row, however, the two strategies yield negative excess returns and the manager thus engages in "absolute strategy timing".

Modelling this behaviour is in fact quite difficult. A first, and obvious choice is to apply the TM and HM frameworks presented earlier, by adding a timing coefficient on each strategy factor, in a fashion similar to Chan, Chen and Lakonishok (2002). We then obtain the following models for the fund returns:

$$r_{t+1} = \alpha + \sum_{k=1}^{K} \beta_{k,0} f_{k,t+1} + \sum_{k=1}^{K} \gamma_k f_{k,t+1}^2 + \varepsilon'_{t+1} \quad (8)$$

$$r_{t+1} = \alpha + \sum_{k=1}^{K} \beta_{k,0} f_{k,t+1} + \sum_{k=1}^{K} \gamma_k \max(f_{k,t+1}, 0) + \varepsilon'_{t+1} \quad (9)$$

Where $f_{k,t+1}$ denotes the return of strategy k in excess of the risk-free rate for the next period, and K the number of strategies available to the multi-strategy manager. Equation (8) builds on the TM model, while equation (9) is based on the HM model.

However, these models overlook the fact that a strategy timer would actually have to compare the forecasted returns of available strategies; instead, they assume that the manager considers each strategy separately, and decides on the exposure $\beta_{k,t}$ based uniquely on the forecasted excess return of strategy k. In other words, these two models assume that the manager only engages in "absolute strategy timing". This will lead to a model misspecification issue if the multi-strategy manager engages in relative strategy timing as well. This can be easily seen through the perfect strategy timer example above: assuming that either equation (8) or (9) holds, no values for $\beta_{k,0}$ and γ_k can be found so that the model is consistent. This issue raises the need for another model.

Before going on to the next sub-section, where we derive more adequate models, it seems useful to discuss and conjecture how strategy timing abilities are implemented by managers in practice. This will also serve as justification for some modelling choices we make later on.

Earlier, we took the example of a perfect timer, who could accurately forecast each strategy's future return, with certainty. In practice, of course, the manager's forecasts are subject to imprecision. As such, the manager, as long as he is risk-averse, will avoid investing all the fund's capital into one strategy; rather, he will simply increase (decrease) the fund's exposure to strategies that are predicted to perform well (badly)[19]. This will also allow him to retain the benefits of diversification which, as explained before, are another important selling argument for multi-strategy hedge funds. Furthermore, the manager may be hindered by "switching costs", i.e. costs incurred by the disinvestment process required to switch between strategies. These costs may in fact be quite relevant in determining the aggressiveness with which the manager can react to his forecasts, as hedge funds are known, and have been shown, to regularly invest in illiquid securities[20].

As previously explained, multi-strategy managers may try to time the overall level of profitability of their strategies, in addition to timing their relative profitability: they may engage in "absolute strategy timing" as well as in "relative strategy timing". They would then reallocate their capital to risk-free investments when all their strategies are predicted to generate negative excess returns[21]. On a theoretical point of view, there is no

[19] This is, of course, the intuition behind the model of Treynor and Mazuy (1966). The linear relationship between the forecast and the factor exposure can be derived mathematically under certain assumptions, see for example Jensen (1972).

[20] See for example Getmansky, Lo, and Makarov (2004). Trading in illiquid securities is without a doubt one of the main reasons why hedge fund managers impose lock-up periods, gates, and redemption fees onto their clients.

[21] This would most probably happen in the event of a market downturn, as hedge fund strategies are known to exhibit increased correlation with the market during this kind of event (as it is the

a priori reason to believe the manager only engages in one of the two processes: indeed, relative and absolute strategy timing both only require to forecast the excess returns of available strategies. As such, a rational manager would make use of his forecasts to the fullest and engage in both timing aspects. An adequate model should thus allow for both components of strategy timing ability as well.

Finally, it is actually unclear whether a typical multi-strategy manager adjusts his strategy exposures systematically and frequently, e.g. at the start of each month, or rather unfrequently, on an ad-hoc basis. The latter may be plausible if the switching costs we previously mentioned are high enough so that frequent, incremental adjustments in strategy exposures are in fact not desirable. In this case, a model which assumes that strategy timing takes place on a regular basis might be misspecified. On this subject, Goetzmann, Ingersoll, and Ivković (2000) show that timing ability will be underestimated if the model assumes that investment decisions are made on a monthly basis while in reality, they are made more frequently, e.g. on a daily basis. However, there should be no bias if the opposite is actually true, i.e. if the model assumes that decisions are taken on a more frequent basis that they are in reality. As such, assuming that strategy timing decisions are made at the start of each month should not undermine our capability in detecting existing strategy timing skills if such decisions are in fact made less frequently. Also, as the strategy timing process is essentially a capital allocation decision, it seems unlikely that it would happen on a relatively frequent basis, e.g. daily. We are thus confident that our results should remain unbiased if our models assume monthly strategy timing decisions.

c. *Strategy timing models*

We now build on the TM and HM frameworks to develop new models which specifically allow for strategy timing abilities, as previously defined.

The general idea is the following: the multi-strategy manager will take a diversified position in several strategies, and then periodically adjust each strategy exposure depending on that same strategy's forecasted return, relative to the forecasted overall return of all available strategies. We thus propose the following model:

$$r_{t+1} = \alpha + \textstyle\sum_{k=1}^{K} \beta_{k,t} f_{k,t+1} + \varepsilon_{t+1} \tag{10}$$

$$\text{where, } \beta_{k,t} = \beta_{k,0} + \gamma_k (\tilde{f}_{k,t+1} - \bar{\tilde{f}}_{t+1}) \tag{11}$$

$$\tilde{f}_{k,t+1} = f_{k,t+1} + \eta_{k,t+1} \tag{12}$$

$$\bar{\tilde{f}}_{t+1} = \frac{1}{K} \textstyle\sum_{k=1}^{K} \tilde{f}_{k,t+1} = \bar{f}_{t+1} + \bar{\eta}_{t+1} \tag{13}$$

Where $f_{k,t+1}$ is the return of strategy k in excess of the risk-free rate for the next period, while $\beta_{k,t}$ is the exposure to strategy k and is set at the start of each period, and depends on $\beta_{k,0}$, the "basic exposure", as well as on $(\tilde{f}_{k,t+1} - \bar{\tilde{f}}_{t+1})$, the forecasted "relative return" of strategy k. The latter is in fact the forecasted deviation of the excess return of strategy k from the arithmetic mean of all the available strategies' excess returns, as defined in equation (13). Note that a negative value for $\bar{f}_{t+1}$ implies that the average return of strategies for the next period is inferior to the risk-free rate. The forecasts are subjects to

case with most securities). As such, absolute strategy timing may in practice be very similar to market timing.

random errors, as in the original model (12). Substituting equations (11), (12) and (13) into (10), we obtain the full model:

$$r_{t+1} = \alpha + \sum_{k=1}^{K} \beta_{k,0} f_{k,t+1} + \sum_{k=1}^{K} \gamma_k \, f_{k,t+1}(f_{k,t+1} - \bar{f}_{t+1}) + \varepsilon'_{t+1} \tag{14}$$

Where the error term ε'_{t+1} incorporates the forecast errors from (12) and (13). We see that the model keeps the simplicity of the original TM model as well as its linearity assumption, while explicitly allowing for relative strategy timing [22]. However, the underlying assumption is that the manager does not engage in absolute strategy timing. Indeed, whether $\bar{f}_{t+1}$ is high or low does not affect the overall exposure to the strategies; only the deviations $(f_{k,t+1} - \bar{f}_{t+1})$ affect each strategy exposure separately. As such, even if all strategies are predicted to yield negative returns, the exposures $\beta_{k,t}$ may still remain close to the "basic exposures" $\beta_{k,0}$. As explained before, this assumption cannot be easily justified as a rational manager would use his forecasts $\tilde{f}_{k,t+1}$ to time the "absolute" returns of strategies as well. We thus introduce a second version of this model, where we allow for both aspects of strategy timing by modifying equation (11) as follows:

$$\beta_{k,t} = \beta_{k,0} + \gamma_k \left(\tilde{f}_{k,t+1} - \bar{\tilde{f}}_{t+1} \right) + \delta_k \bar{\tilde{f}}_{t+1} \tag{15}$$

Where δ_k measures how the manager adjusts the strategy exposures depending on how the available strategies are predicted to perform overall, i.e. the absolute strategy timing skill of the manager. Note that this specification is actually a generalized version of the TM extended model, since by introducing the restriction $\gamma_k = \delta_k$, we come back to equation (8). The downside, however, is that it requires two additional variables for each strategy factor on which strategy timing is evaluated; while it would be best to add as little variables as possible. We test the goodness-of-fit of these different specifications later on.

We also develop the HM counterparts of these models, by replacing the linear deviations from the mean by an indicator variable, as shown below.

$$\beta_{k,t} = \beta_{k,0} + \gamma_k I \left(\tilde{f}_{k,t+1} - \bar{\tilde{f}}_{t+1} > 0 \right) \tag{16}$$

$$\beta_{k,t} = \beta_{k,0} + \gamma_k I \left(\tilde{f}_{k,t+1} - \bar{\tilde{f}}_{t+1} > 0 \right) + \delta_k I(\bar{\tilde{f}}_{t+1} > 0) \tag{17}$$

As in the original HM model, the variable I takes the value 1 if the condition in parentheses is met, and 0 otherwise. Specification (16) thus allows for relative strategy timing only, while specification (17) allows for both aspects of strategy timing.

We will hereafter refer to models (8) and (9) as the "extended" TM and HM models, to specifications (11) and (16) as the "relative strategy timing" TM and HM models, and to specifications (15) and (17) as the "generalized strategy timing" TM and HM models, respectively.

[22] In fact, this specification can be mathematically related to the original TM market timing model. By taking raw returns instead of excess returns, and by considering that the manager chooses between two strategies, which are (1) investing in the market portfolio, and (2) investing in the risk-free asset, the original model is found and it can be shown that $\gamma_1 = 2\gamma$ exactly.

d. Applying the strategy timing framework to a benchmark model

The "models" which we presented earlier assumed that we could create factors which perfectly replicate the returns of various hedge funds strategies. In practice, of course, this is easier said than done. As hedge fund strategies have generally been shown to be dynamic and quite complex, it seems unreasonable to think that each of them can be entirely described by one time-varying factor. Still, some researchers have managed to describe specific strategies by using only a few factors. For example, Mitchell and Pulvino (2001) model the returns of event-driven risk arbitrage strategies by constructing a single factor based on mergers and acquisitions which occurred during their sample period. Fung and Hsieh (2001) explain the performance of trend-followers by constructing trend-following portfolios for specific markets, based on the returns of lookback straddles. Fung and Hsieh (2002) study fixed-income strategies and come to the conclusion that most of the funds in this category are mainly exposed to credit spreads. Fung and Hsieh (2005) also show that the returns of equity-oriented hedge fund can be explained to a certain extent by a "market" factor and a "size" factor, both of which are based on market indices. Agarwal, Fung, Loon, and Naik (2011) model the returns of convertible arbitrage hedge funds by constructing two factors, which represent the returns of buy-and-hedge and buy-and-hold strategies in that market. In the light of these advancements, it thus seems possible to model the returns of multi-strategy hedge funds in a way that would be very close to what we described in the previous section.

In our case, we use the model developed by Fung and Hsieh (2004), which has managed to impose itself as a reference benchmark model in the industry. The original version is made of seven factors, which aim to replicate the returns of long/short equity, fixed-income, and trend-following strategies. Recently, an eighth factor was added, which represents exposure to emerging markets strategies. The advantage of this model in our context is thus that factors are indeed supposed to represent the returns of several hedge fund strategies, which is what we need to apply the frameworks we developed in the previous section. We do not use other performance evaluation models for comparison, because none present the same characteristic.

Although the Fung and Hsieh (2004) model is often employed in the recent hedge fund literature, it is important to emphasize that it has some shortcomings, especially in the context of this study. First, although the two long/short equity factors and the two fixed-income factors seem to generally have statistically significant power in explaining the returns of their respective hedge fund strategies, they cannot capture the return characteristics of these strategies to the fullest. Indeed, Fung and Hsieh (2004) report an R^2 of 0.77 in a regression of the returns of long/short equity funds on the two equity factors, and an R^2 of 0.30 in a regression of the returns of fixed-income hedge funds on the two fixed-income factors. These factors thus constitute incomplete representations of their respective hedge fund strategies' returns, which may hinder our empirical analysis. Second, the three trend-following factors exhibit very low explanatory power in most studies, which hints at a possible issue in the modelling of this strategy. Third, the model theoretically encompasses long/short equity, fixed-income, trend-following, and emerging markets strategies only. Other hedge fund strategies, e.g. the event-driven strategy, are thus not meant to be described by the model. This limit may prove quite restrictive in our case if multi-strategy funds routinely manage strategies other than the four which are embedded in the model. In particular, multi-strategy funds in our sample have been shown to invest in convertible arbitrage strategies (cf. section 3), which are not embedded in the model either[23].

An alternative way of modelling the returns of multi-strategy funds would be to use hedge fund strategy indices as factors, e.g. the Credit Suisse and HFR indices we employed in section 3. Although this approach may seem attractive at first sight, it is in fact undermined by selection biases. Indeed, as data providers construct their indices based on the performance of hedge funds which do report to them, the indices are prone to potentially important selection bias, as explained by Fung and Hsieh (2000). Empirical results may then be dependent on the indices which were used, which is clearly undesirable.

We now briefly describe the Fung and Hsieh (2004, hereafter FH) eight-factor model. It assumes that hedge fund returns can be modelled in the following way:

$$r_{t+1} = \alpha + \beta_1 MKT_{t+1} + \beta_2 SIZE_{t+1} + \beta_3 BOND_{t+1} + \beta_4 CRDT_{t+1} + \beta_5 PTFS_{BD,t+1} + \beta_6 PTFS_{FX,t+1} + \beta_7 PTFS_{COM,t+1} + \beta_8 EMRG_{t+1} + \varepsilon_{t+1} \tag{18}$$

Where r_{t+1} is the fund's excess return observed at t+1, *MKT* is the total monthly return on the S&P 500 in excess of the risk-free rate, and proxies the market portfolio, *SIZE* is the total monthly return on the Russell 2000 index minus that of the S&P 500, and proxies the usual size factor as developed by Fama and French (1992), *BOND* is the monthly change in the 10-year constant maturity T-bond rate, *CRDT* is the monthly change on the spread between Moody's Investor Service Baa bond and the 10-year constant maturity T-bond, and proxies the overall level of credit spreads, $PTFS_{BD}$, $PTFS_{FX}$, and $PTFS_{COM}$ are diversified portfolios of lookback straddles, respectively in the bond, currencies, and commodities markets, and proxy the returns of trend-following strategies, and *EMRG* is the total monthly return on the MSCI Emerging Markets index. Panel B of [table 1] presents the usual summary statistics for these variables.

In order to measure strategy timing ability, we allow the coefficients of the *MKT*, *SIZE*, and *CRDT* factors to vary over time as described in the previous section. This choice is made following the observation that the multi-strategy funds in our sample seem to mainly have exposure to the stock and fixed-income markets. As *MKT* and *SIZE* may in fact represent two different strategies in the equity markets – respectively a buy-and-hold strategy on the market portfolio and an arbitrage strategy between small cap and large cap stocks – we allow the coefficients of both factors to vary independently. Regarding the exposure to fixed-income markets, we only allow the coefficient of the *CRDT* factor to vary over time, as Fung and Hsieh (2002) concluded that credit spreads should be the main driver for the return of fixed-income strategies. With this specification, we thus consider that multi-strategy managers only attempt to time these three strategies, as other coefficients are restricted to remain constant over time. The strategy timing model is thus written as follows:

[23] As a way to address this last issue, it would be interesting to update the factors developed by Mitchell and Pulvino (2001) and Agarwal, Fung, Loon, and Naik (2011) with recent data and add them to the Fung and Hsieh (2004) model. However, this goes beyond the scope of this Master's thesis.

$$r_{t+1} = \alpha + \beta_{t,1}MKT_{t+1} + \beta_{t,2}SIZE_{t+1} + \beta_3 BOND_{t+1} + \beta_{t,4}CRDT_{t+1} + \beta_5 PTFS_{BD,t+1} + \beta_6 PTFS_{FX,t+1} + \beta_7 PTFS_{COM,t+1} + \beta_8 EMRG_{t+1} + \varepsilon_{t+1} \tag{19}$$

And the time-varying coefficients have the following form for the TM version:

$$\beta_{k,t} = \beta_{k,0} + \gamma_k \left(\tilde{f}_{k,t+1} - \bar{\tilde{f}}_{t+1}\right) + \delta_k \bar{\tilde{f}}_{t+1} \qquad for\ k = 1,2, or\ 4.$$

Or, alternatively, the following form for the HM version of the model:

$$\beta_{k,t} = \beta_{k,0} + \gamma_k I\left(\tilde{f}_{k,t+1} - \bar{\tilde{f}}_{t+1} > 0\right) + \delta_k I(\bar{\tilde{f}}_{t+1} > 0) \qquad for\ k = 1,2, or\ 4.$$

As a reminder, $\beta_{k,0}$ is the "base" exposure, γ_k measures relative strategy timing, and δ_k measures absolute strategy timing. For the HM version of the model, I denotes the usual indicator variable which takes the value 1 if the condition stated in brackets is met, and 0 otherwise. By plugging the TM version of the coefficients' expressions into equation (19), we thus obtain the TM variant of the generalized model in its empirical form:

$$\begin{aligned} r_{t+1} = \alpha &+ \beta_1 MKT_{t+1} + \gamma_1 MKT_{t+1}\left(MKT_{t+1} - \bar{f}_{t+1}\right) + \delta_1 MKT_{t+1}\bar{f}_{t+1} + \beta_2 SIZE_{t+1} \\ &+ \gamma_2 SIZE_{t+1}\left(SIZE_{t+1} - \bar{f}_{t+1}\right) + \delta_2 SIZE_{t+1}\bar{f}_{t+1} + \beta_3 BOND_{t+1} \\ &+ \beta_4 CRDT_{t+1} + \gamma_4 CRDT_{t+1}\left(CRDT_{t+1} - \bar{f}_{t+1}\right) + \delta_4 CRDT_{t+1}\bar{f}_{t+1} \\ &+ \beta_5 PTFS_{BD,t+1} + \beta_6 PTFS_{FX,t+1} + \beta_7 PTFS_{COM,t+1} + \beta_8 EMRG_{t+1} + \varepsilon'_{t+1} \end{aligned} \tag{20}$$

Where:

$$\bar{f}_{t+1} = \frac{MKT_{t+1} + SIZE_{t+1} + CRDT_{t+1}}{3}$$

And ε'_{t+1} is an error term in which the forecast errors were incorporated. We obtain the relative strategy timing model by imposing the restriction $\delta_k = 0$ for k = 1, 2, and 4. We obtain the extended model by imposing the restriction $\delta_k = \gamma_k$ for k = 1, 2, and 4 so that the term $\bar{f}_{t+1}$ cancels out. Similarly, the HM version of this model is obtained by plugging the HM version of the coefficients' expressions into equation (19).

To avoid making any arbitrary choice regarding which factors do or do not have time-varying coefficients, we test a set of alternative specifications. We also use this test to compare the goodness-of-fit of the extended, relative, and generalized strategy timing models, as well as their TM and HM versions. This test is actually quite important since by allowing on the exposures to the *MKT*, *SIZE*, and *CRDT* factors to vary over time, we make the assumption that multi-strategy managers only engage in strategy timing concerning equity and fixed-income strategies. Any unjustified choice on this level should thus raise doubts about the validity of our future results. For simplicity, however, we detail the procedure we use as well as the results in [Appendix D]. In short, we select the generalized strategy timing model, with time-varying exposures on the *MKT*, *SIZE*, and *CRDT* factors as in equation (19), as this specification exhibits a fairly satisfying goodness-of-fit, while not having the highest number of added variables compared to the original FH model. In addition, the model is consistent with our conclusions regarding the main markets in which multi-strategy managers execute their strategies. The TM version of the modelled is preferred, but we use the HM variant as well in our empirical tests.

Instead of assuming that multi-strategy managers only attempt to time equity and fixed-income strategies, we could assume that all strategies are in fact subject to strategy timing, and purposely omit some of the timing coefficients to avoid having too many explanatory variables in the final model. We would then compute the monthly average strategy return – as defined in equation (13) – using all of the FH factors' returns. The advantage of this alternative approach lies in that no restrictive assumption is made concerning which strategies are timed by managers. However, the model is then explicitly subject to an omitted variables bias, which may affect strategy timing estimates if the timing variables are correlated with each other. To avoid testing too many different models, we thus stick to our initial assumptions.

[Table 5] presents summary statistics for the timing variables in this model. By "timing variables", we refer to the six additional variables in equation (20), in comparison with the original FH model. As an example, for the MKT factor, MKT_{ga} measures relative strategy timing ability, while MKT_{de} measures absolute strategy timing ability. The same terminology is employed for the $SIZE$ and $CRDT$ factors.

Panel A presents the summary statistics of the six linear strategy timing variables – those of the TM version of our model, as in equation (20). Interestingly, their distributions exhibit high kurtosis – fat tails – as well as positive skewness, aside from the $CRDT_{de}$ variable. This implies that the strategies' returns sometimes strongly deviate from the monthly mean, although commonly observed deviations may be smaller. Also, the two timing coefficients on the MKT factor exhibit the highest percentages of positive values, while the timing coefficients related to the $CRDT$ factor exhibit the lowest. This means that the MKT factor yields returns which are above the monthly mean on average, and the opposite conclusion can be made for the $CRDT$ factor. This is, as expected, in line with the average returns of each of those three factors, which follow the same ranking.

Panel B presents the summary statistics for the timing variables in the HM version of our model, which are constructed using the indicator variable. The analysis is fairly similar, so we skip it for the sake of simplicity.

[Table 6] present the correlation matrices for the two sets of factors. We notice that some correlation coefficients are quite high, especially between related timing variables, e.g. MKT_{ga} and MKT_{de}. This raises the issue of multicollinearity for future estimation procedures: indeed, if explanatory variables are too highly correlated, then our coefficient estimates will be imprecise.

e. *Another way to evaluate strategy timing abilities empirically*

The obvious problem when estimating a model such as equation (20) is that the high number of explanatory variables may hinder the overall precision of estimates. Indeed, as we are in the presence of sometimes quite small samples – some funds in our sample have less than 30 months of return history – the available return series may not prove large enough to estimate the coefficients of the six timing variables with sufficient precision. Another downside to this model is that strategy timing ability is measured by multiple variables jointly. Indeed, in opposition to a simple market timing model where a timing coefficient γ entirely describes the studied timing skill, in our model – as defined by equation (20) – the six timing coefficients $\gamma_1, \gamma_2, \gamma_4, \delta_1, \delta_2, \delta_4$ theoretically all need to be positive and significant so that one can conclude that a significant strategy timing ability is found in the sample. Interpretation may thus prove difficult in other cases, e.g. if some timing coefficients are positive and some others are negative. Also, this increases the probability of type II errors in detecting strategy timing. Indeed, if strategy timing was measured by one coefficient only, then our null hypothesis could be stated as:

$$H0: \theta_{ST} = 0$$

Where θ_{ST} is the coefficient which measures, alone, strategy timing ability. As $\hat{\theta}_{ST}$, its sample estimate, is a random variable, there is always a risk of type II error, i.e. not being able to reject the null while the alternative hypothesis is actually true. The issue in having multiple variables to jointly measure strategy timing ability is that each timing coefficient estimate is subject to this sampling variation, and thus the risk of not being able to conclude about timing ability, because not all the coefficient estimates are positive and significant, is higher than in the single-variable case. The null hypothesis in the multiple-timing-variables case is not as easily written as before, but we can state the alternative:

$$Ha: \gamma_1, \gamma_2, \gamma_4, \delta_1, \delta_2, \delta_4 > 0$$

It is easy to see that if any of these coefficients is not found to be statistically different from zero and/or positive, then we have to "reject" the alternative hypothesis, i.e. we cannot reject the null which is loosely defined as "no strategy timing". This regression-based test for detecting strategy timing abilities may thus have, in fact, low statistical power. In addition, as we previously showed, multicollinearity may well be an issue in this particular model due to high linear correlation between some of the added variables.

It would thus be far better to measure strategy timing skills through one coefficient only in our regression model. We enable this by creating a new "strategy timing" factor which we add to the original FH 8-factors model. The idea is the following: in the relative strategy timing models, the exposure to a strategy is allowed to increase and decrease proportionally to that strategy's relative performance, i.e. the deviation of that strategy's predicted return from the average predicted return of available strategies, as in equation (11). The manager is able to enhance his performance by engaging in strategy timing, but the magnitude of that enhancement ultimately depends on how much the strategies' returns differ from one another. Indeed, if all of the manager's strategies are predicted to yield the same return over the next period, then the value of any relative strategy timing ability for the next period is zero. On the opposite, if the strategies are predicted to generate very different levels of performance, then the manager, provided that he possess strategy timing skills, will adjust his exposures more aggressively, and increase his return by doing so.

Therefore, the ex-post performance of a manager who can time strategies (in the "relative" sense) will be linked to the dispersion of his strategies' returns. The more dispersed they are, the higher his performance. Our strategy timing factor is thus defined as the cross-sectional variance of the strategies' returns in each month, as shown below.

$$STX_t = \frac{1}{K}\sum_{k=1}^{K}(f_{k,t} - \bar{f}_t)^2 \tag{21}$$

Where STX_t is the value of our strategy timing factor in month t, and $f_{k,t}, \bar{f}_t$ denote as usual the excess return of strategy k, and the strategies' mean return, respectively, in month t. The STX factor is thus added as an explanatory variable to the FH 8-factors model (cf. equation (18)), and its coefficient measures strategy timing ability. An important distinction with the models we developed in sub-section (c) is that this factor is simply meant to measure the ex-post effect of strategy timing on the fund's performance, whereas the previous frameworks modelled the ex-ante relation between the two by linking the manager's forecasts to his time-varying exposures. Although we gain a significant advantage in terms of model simplicity and statistical power by using the FH model along with this new factor instead of the generalized strategy timing model, the latter provides more detailed insights on timing abilities and is more elegant theoretically. Still, combining the two approaches appears as a good compromise.

We use the variance as the dispersion measure because it is most common, but also because its properties are in line with our assumptions on strategy timing abilities. As a reminder, we discussed at the end of sub-section (b) the possible existence of "switching costs" which would make it undesirable for the manager to adjust his exposures to strategies if these strategies' returns do not vary enough from one period to the next. As such, in practice, the manager would only adjust exposures in the cases where the strategies' performance levels differ strongly from one another, i.e. when dispersion is high enough. This leads us to think that a convex relation between dispersion in the strategies' returns and the fund's performance is what is most representative of a strategy timer's behaviour in practice, hence the choice for the variance as a dispersion measure, since it places more weight on very large deviations. Nevertheless, we consider the possibility for this specific relation to be linear, and thus use the average absolute deviation from the mean as an alternative dispersion measure. We label this variant as the STX_L factor.

$$STX_{L,t} = \frac{1}{K}\sum_{k=1}^{K} |f_{k,t} - \bar{f}_t| \tag{22}$$

Following our previous choices, we compute the *STX* variable assuming that only the *MKT*, *SIZE*, and *CRDT* factors are subject to strategy timing. This enables us to compare the results obtained from the previously selected models since they rely on the same assumption. Nevertheless, we test alternatives ways of computing the *STX* factor to avoid making any arbitrary choice. We do so, as we did earlier, by regressing the returns of the equally-weighted portfolio of funds on each set of factors, and then comparing the goodness-of-fit. It appears that changing the way the *STX* is computed does not affect inference on the *STX* itself or on other variables in any significant manner, and the same conclusion is drawn for the STX_L. As such, we keep our initial assumption for the sake of simplicity. The results of this comparison are reported in [table 10][24].

[Table 7] presents the usual summary statistics for the *STX* factor and its linear counterpart, STX_L. As the STX_L factor is constructed as the average absolute deviation from the mean, it is expressed in percentage – as the returns series – while the *STX*, being a variance, is expressed in squared percentage. A comparison between the mean values or the standard deviations of the two factors is thus meaningless. However, the skewness and kurtosis estimates show that the *STX* factor has fatter tails, which was expected since the variance exaggerates the importance of large deviations. Also, panel B shows that the *STX* and STX_L factors exhibit relatively low correlations with the FH model factors, which rules out any multicollinearity issues. The highest correlation in absolute value is found to be with the *CRDT* factor in both cases, the coefficient being 0.293 and 0.231, respectively.

[24] As this table also presents regression results for the *STX* factor, we analyse it in full later on, in section 5.

f. Estimation procedures

We now explain very briefly how we estimate our models and how we conduct statistical inference. As our two models – the generalized strategy timing model and the *STX* model – are to be estimated for one fund at a time, and as all variables are actually computed on the same period, i.e. there are no lagged variables, we can estimate them very easily through ordinary least squares (hereafter OLS). The OLS estimates will be unbiased as long as independent variables are strictly exogenous[25]. Violations of this assumption usually come from omitted variables, measurement errors, or unaccounted lagged effects in the model. There is no obvious case of such a violation within our two frameworks, so we assume that strict exogeneity is observed.

However, even if OLS estimates are unbiased, inference might not be valid if the error term is subject to heteroscedasticity and/or serial correlation. The latter is especially relevant in our case as significant serial correlation was found in the returns series (cf. [table 1]). We thus test for the presence of these issues in aggregate-level regressions with both models. The homoscedasticity assumption is tested using a standard White (1980) test, and we use Durbin's (1970) alternative statistic[26] to check for the presence of serial correlation. Results are presented in [table 8]. The homoscedasticity assumption is rejected for the *STX* models and is close to being so for the TM version of the generalized strategy timing model. The assumption of no serial correlation is rejected for all models, although at different confidence levels. In the light of these results, we cannot boldly assume that the basic OLS inference will remain valid; hence, we use the standard errors of Newey and West (1987) in all our regressions, which are robust to unspecified forms of both heteroscedasticity and serial correlation (abbreviated HAC robust). Following the rule of thumb proposed in their original paper, we set the lag parameter as follows:

$$lags = 4\left(\frac{n}{100}\right)^{2/9}$$

Where n is the number of available observations, or time periods in our case. For aggregate-level regressions, n=287 and as such, the lag parameter is set to 5. By using these HAC robust standard errors, we ensure that the inference on coefficients remains consistent. It is known, however, that such standard errors may be quite inefficient in small samples, which might well be the case for us; but we prefer this alternative to simply using biased OLS estimates. Indeed, if serial correlation is positive – which is the case for our return series, see [table 1] – OLS standard errors will be biased downward, which may overestimate the statistical significance of some coefficients and thus increase the risk of making type I errors. Note, also, that in the presence of serial correlation the R^2 statistics may be biased; but we still report them for the sake of simplicity.

[25] Independent variables are strictly exogenous if the expectation of the error term in each period is zero, conditionally on the explanatory variables in all time periods. Mathematically, we write:

$$E(u_t|\mathbf{X}) = 0, t = 1,2, \ldots, n.$$

[26] This statistic has the advantage of being valid even if some explanatory variables happen not to be strictly exogenous.

V. Empirical results at the aggregate level

We now present our empirical results at the aggregate level, i.e. using the equally-weighted portfolio of funds. We first study the estimates of the generalized strategy timing model. We then study the *STX* model and its variants. Finally, we test the robustness of the latter to more complete specifications, where we control for reaction to publicly available information, market timing, and both simultaneously.

As we are not able to construct a value-weighted portfolio of funds due to incomplete assets-under-management data in our sample, these results should be taken with caution since they may be dependent on the way our aggregate portfolio is constructed. Still, studies on hedge funds which conduct an aggregate-level analysis often find that using either methodology does not affect the results in any significant way[27]. Other authors simply prefer to conduct tests at the fund-level only[28]. Still, we run tests at the aggregate level, one of the reasons being that it easily provides a first look at the goodness-of-fit of our various models. Empirical results at the fund level are presented in section 6.

a. *Generalized strategy timing models*

We estimate the TM version of the generalized strategy timing model, as described by equation (20), using the returns of the equally-weighted portfolio of funds as the dependent variable. We estimate the HM counterpart of this model as well. Results are presented in [table 9].

Overall, the model does not perform well. This was to be expected given the high number of variables in these specifications as well as the sometimes strong linear correlations between them. None of the strategy timing variables exhibit significant coefficients in the TM version of the model, even at the 10% confidence level. The results are slightly different for the HM version of the model, where two of the three absolute strategy timing coefficients are positive and statistically significant (MKT_{de} and $SIZE_{de}$). As other timing coefficients stay statistically insignificant, this could hint at multi-strategy managers possessing some market timing ability, but no strategy timing ability[29]. Still, as multicollinearity issues between timing coefficients are quite important in this model, we cannot conclude with certainty about the existence of any timing ability. Rather, these results seem to suggest the opposite.

Earlier, we selected this particular model specifications on the basis of its high adjusted R^2 statistic (cf. [table 4]). One might argue that this approach is prone to bias, as the R^2 is not always an adequate measure of goodness-of-fit[30]. As such, we also run aggregate-level regressions using the other model specifications, and find qualitatively similar results. These are unreported for the sake of simplicity, but are available on request.

As no significant results are obtained regarding strategy timing ability in this model, we do not conduct further tests or robustness checks. However, we do run the model at the fund level in section 5.

---

[27] Examples related to our context include Chen (2006) and Chen and Liang (2006).

[28] See for example Fung, Xu, and Yau (2002), Cao, Chen, Liang, and Lo (2013), Agarwal, Ruenzi, and Weigert (2017).

[29] As previously mentioned, absolute strategy timing may be assimilated to a form of market timing.

[30] In particular, the R^2 might be biased under the presence of serial correlation in residuals.

b. STX factors and robustness to alternative specifications

We now estimate the *STX* models, i.e. the FH 8-factors model to which we add the *STX* variable, using the returns of the equally-weighted portfolio of funds as the dependent variable. [Table 10] reports the regression results.

Panel A presents the coefficient estimates as well as their *t*-statistics for the *STX* factor and its linear counterpart, the STX_L. Both model specifications perform surprisingly well; indeed, coefficient estimates on the FH factors exhibit values and significance levels which are quite close to those of the original model. Both variants of the *STX* factor exhibit a positive coefficient which is significant at the 2% confidence level, although the STX_L coefficient is somewhat less significant than the other. In addition, the alpha (i.e. the intercept) for both specifications is reduced compared to the original FH model. Part of the initial abnormal return could thus be explained by the positive value of strategy timing abilities. These results are thus so far quite consistent with the claims some multi-strategy managers make about being able to time their strategies.

As the *STX* variable is actually expressed in squared percentages, we cannot directly interpret its estimated coefficient. However, the STX_L offers this possibility since its unit is the percentage point. As such, given our estimations, an increase of one percentage point in the average absolute deviation of the strategies' returns from the mean return increases the equally-weighted portfolio's return by 0.22 points of percentage. To illustrate the magnitude of this estimate, we compute the predicted return of the portfolio of funds when all factors take their average value, as shown in the table below:

EWP	α	STX_L	*MKT*	*SIZE*	*BOND*	*CRDT*	$PTFS_{BD}$	$PTFS_{FX}$	$PTFS_{COM}$	*EMRG*
0.43	0.23	0.28	0.65	0.02	-0.01	0.00	-1.85	-0.90	-0.55	0.67

Note that values are expressed in percent. As such, given the values taken by the *MKT*, *SIZE*, and *CRDT* factors, the STX_L factor is equal to 0.28%, which translates as an increase of about 0.06 (0.22 x 0.28) points of percentage in the portfolio return. The STX_L factor thus accounts for around 15% (0.06/0.43) of the portfolio's performance in this case, which is equal to the contribution of the *MKT* factor. The contribution of strategy timing abilities to the performance of the portfolio is thus not negligible in this case.

Interestingly, the *STX* factor appears to have higher statistical significance than the STX_L. This supports our previous conjecture that because of "switching costs", multi-strategy managers may only engage in strategy timing when the strategies' dispersion is sufficiently high. We also notice that the alpha is still positive and significant after controlling for strategy timing with the *STX* factor, while it becomes insignificant when the STX_L factor is used. This observation, however, is difficult to explain. This is potentially due to the *STX* factor being close to zero quite often, while the STX_L factor, being linear in construction, has a higher average value (as previously reported in [table 7]). As such, the STX_L would automatically capture a higher portion of the alpha. Nevertheless, we believe that the *STX* is better specified, and as such we keep it as our main factor for future tests.

Panel B reports coefficient estimates for the *STX* factor when alternative data filters are used to construct our sample. Specifically, we require funds to have a minimum return history of 36 months in the first regression, and we remove the suspected funds-of-hedge-funds from the sample in the second regression (cf. section 3). The results appear to be quite close to the ones obtained when using the original sample (cf. panel A). This was to be expected, since these two alternative samples respectively contain 81 and 84 funds, while the base sample contains 90 funds. As such, removing such a small numbers of funds from the sample should have a minimal effect at the aggregate level. Still, the difference is surprisingly small in this case; we thus skip these robustness tests in other tests and consider that our data filtering choices should have no significant impact on overall results.

Panel C presents coefficient estimates for alternatively computed *STX* and STX_L factors. Specifically, we test different assumptions about which FH factors are subject to strategy timing. As discussed in section 4, the results of these alternative specifications are very similar to those of the original *STX* factors. We thus keep the *STX* and STX_L factors in their initial form, so that they rely on the same assumption about strategy timing than in the generalized model.

In the next sub-sections, we conduct additional robustness tests for the *STX* model. Specifically, we control for reaction to public information, and for the presence of market timing skills.

c. *Controlling for publicly available information*

Ferson and Schadt (1996) highlight an important flaw of classic market timing models, in that they do not take publicly available information into account when evaluating timing skills. Indeed, if the market return can be predicted using public information, then all rational managers should adjust their market exposure depending on that information; which, as the authors argue, cannot be considered to be true market timing ability. Still, classic market timing models such as the TM and HM frameworks assume that forecasts of the market return rely entirely on superior – private – information; as such, they will overestimate true timing skill if public information is indeed predictive, at least partly, of the market return. Ferson and Schadt (1996) thus derive "conditional" versions of the TM and HM models, which account for publicly available information and are thus able to distinguish true market timing from mere reaction to public signals. Their conditional performance evaluation framework is also able to explain negative timing abilities which are sometimes obtained through the classic models.

Our *STX* factor and its linear variant both happen to be serially correlated, with *AR(1)* coefficients of around 0.37 and 0.32, respectively, as reported in [table 7]. If the dispersion in the strategies' returns tends to stay at similar levels from one period to the next, then we can conjecture that well-performing strategies in one period will also perform well in the next period, and inversely for strategies which perform badly. As such, managers may use the lagged returns of strategies to predict their future relative returns and then adjust their exposures accordingly; we cannot consider this process to be true strategy timing skill as it does not rely on superior forecasts, but rather on public information. Following Cao, Chen, Liang, and Lo (2013), we thus use the residuals of an AR(1) regression as the unpredictable component of the *STX* factor, and the lagged value of the *STX* as the "public information" variable. Our conditional model is thus as follows:

$$r_{t+1} = \alpha + \gamma STX_{t+1}^{r} + \varphi STX_t + \sum_{k=1}^{K} \beta_k f_{k,t+1} + \varepsilon_{t+1} \qquad (23)$$

Where STX_{t+1}^{r} is the unpredictable part of the *STX* factor, and STX_t is the lagged value of the latter. The terms $f_{k,t+1}$ denote the factors of the original FH model, and ε_{t+1} is the usual error term. In this conditional model, γ measures true strategy timing skill, while φ measures reaction to public information. If multi-strategy managers only react to lagged information, then we expect the γ coefficient not to be significant once we estimate this model.

We estimate equation (23) using the equally-weighted portfolio's return as the dependent variable. Results are reported in panel A of [table 11]. The coefficient estimate for the STX^r variable is positive and significant, meaning that the strategy timing ability we previously detected is distinct from simple reaction to public information. Note, however, that this coefficient exhibits slightly lower statistical significance than the *STX* coefficient in the original regression. Also, the coefficient estimate on the lagged *STX* factor is positive and strongly significant, which implies that multi-strategy managers do react to publicly available information as well in order to "time" their strategies.

d. *Controlling for market timing*

As our *STX* factor is constructed to measure the average deviation between the strategies' returns, it proxies the value of relative strategy timing abilities only. Indeed, its value is not affected by the overall level of the strategies' returns, as such, the previous models we estimated do not take absolute strategy timing into account. Still, we can relate the latter to market timing ability, since both concepts essentially define whether or not a manager is able to switch to risk-free investments before his focus markets turn bearish. Building on this similarity, we control for the presence of market timing abilities in the *STX* model by adding the squared market return as a variable, in the spirit of the original TM model. This enables us to test the robustness of previous results when controlling for market timing, or similarly, when controlling for absolute strategy timing.

Panel B of [table 11] reports the results of this regression. The market timing coefficient estimate takes a negative value, and is statistically insignificant. Other coefficient estimates, as well as their significance, remain largely unaffected in comparison to the basic *STX* model. Relative strategy timing ability, as measured by the *STX* factor, is thus distinct from market timing ability, or similarly, from absolute strategy timing ability. Also, these results imply that multi-strategy managers would not possess the latter.

The fact that we find no significant market timing ability for the funds in our sample is actually consistent with the results obtained by Cai and Liang (2012), who conclude that multi-strategy hedge funds do not possess any ability in timing the stock, bond, commodity, or currency markets. However, these authors do find some evidence of multi-strategy managers being able to time market volatility.

Although the market timing coefficient is statistically insignificant, it is of negative sign, which may appear as a troubling result. Indeed, perverse timing ability makes no sense on a theoretical point of view[31]. However, the timing literature has generally shown that in the classic market timing models, estimates of selectivity skills (as measured by Jensen's alpha) and of market timing abilities tend to be negatively correlated[32]. A possible explanation is provided by Jagannathan and Korajczyk (1986) who argue that this observed negative covariance may be due to unaccounted options trading[33]. Controlling for this possibility in our regression appears irrelevant since the market timing coefficient we obtain is statistically insignificant, but one should remain aware

[31] See for example Henriksson and Merton (1981), Ferson and Schadt (1996).

[32] See Coggin, Fabozzi, and Rahman (1993) for a review of the relevant literature on this subject.

[33] The intuition is quite simple. The Henriksson and Merton (1981) model assimilates market timing ability to a free call option on the market portfolio. If a fund manager does not engage in market timing, but does hold call options, then his returns will resemble that of a true market timer. However, since the manager pays the options' prices, his performance will be lower to that of a true market timer; hence, classic market timing models will detect positive market timing ability, but a negative alpha.

that the negative sign of the estimate may simply be the consequence of this unaccounted issue.

e. Controlling for public information and market timing jointly

As a final robustness test for these results at the aggregate level, we control for reaction to public information and market timing jointly in our *STX* model. To do so, we simply add the squared market return as an explanatory variable in equation (23). All the coefficients thus keep their respective interpretation. Results are presented in panel C of [table 11]. Our previous conclusions remain unchanged, as coefficients estimates are unaffected overall, and the same is observed about their significance levels.

VI. Empirical results at the fund level

We now turn to fund-level analyses. We estimate the generalized strategy timing models as well as the *STX* models for each fund, and then summarize the cross-section of *t*-statistics. We then run a bootstrap analysis on these *t*-statistics, which allows us to assess their statistical significance in a valid manner, but also to determine if high observed values may simply be due to managers being lucky. Finally, we control, again, for reaction to public information as well as for the presence of market timing abilities.

a. Generalized strategy timing models

We first run fund-level regressions using the generalized strategy timing models. For each fund, we estimate equation (20), and its HM counterpart, compute the HAC robust *t*-statistics and store them. We thus obtain the cross-section of *t*-statistics for each of the six strategy timing coefficients, and for both model specifications. Following Cao, Chen, Liang and Lo (2013), we summarize this cross-section in [table 12] by reporting the percentage of *t*-statistics which exceed standard cut-off values. Panel A reports results for the TM version of the model, while panel B reports results for the HM counterpart of the model.

Overall, even if some *t*-statistics appear to be positive and significant, these results raise doubts about the existence of strategy timing skills at the fund level. Indeed, even if the right tail of the distribution is thicker for four coefficients out of six, the difference with the left tail remains faint. Moreover, some funds seem to exhibit negative timing ability for some strategies. The HM version of the model exhibits a lower percentage of *t*-statistics exceeding cut-off values overall.

As explained before, the six strategy timing coefficients need to be simultaneously positive and significant so that we can conclude about the presence of significant strategy timing ability regarding a given fund. As [table 12] reports the distribution of *t*-statistics for each coefficient individually, it does not allow to analyse the joint distribution of *t*-statistics for the six timing coefficients. In other words, it would be useful to know if some funds do exhibit positive and significant coefficients on the six timing variables simultaneously, and if so, how many. We thus look at the full cross-section of *t*-statistics in detail to answer this question. It appears that no fund finds itself in such a situation; funds which exhibit large and positive *t*-statistics on some of the timing variables tend to show large and negative *t*-statistics on the rest of the timing variables. This confirms once and for all our intuition about multicollinearity issues in this model, as some coefficients systematically exhibit opposite values.

These results are in line with the aggregate-level analysis using the same models. Still, some funds are found to have seemingly significant strategy timing ability, even if it appears negative for some components. This may, in fact, simply be due to some of the fund managers being lucky. If a manager randomly allocates his capital to available strategies at each period, then he may sometimes "time" some of the strategies by pure luck; in turn, this could also explain why some of the timing coefficients exhibit negative and significant values, as such a manager may also turn out to be unlucky. Also, in [table 12], we use cut-off values which correspond to thresholds from the normal distribution. Assuming that these asymptotic values are valid in our case is wrong, as we are in the presence of very small sample periods for some of the funds. Indeed, in small samples, even the Student distribution with adequate degrees of freedom may not approximate correctly the true distribution of *t*-statistics. In the next section, we thus run a bootstrap analysis which allows us to bypass both of these issues.

b. Bootstrap procedure for t-statistics

We run a bootstrap procedure on the cross-section of t-statistics, following Cao, Chen, Liang and Lo (2013). Chen and Liang (2006), among others, use the same approach. The technique actually originates from Efron (1979). As explained by Cao, Chen, Liang and Lo (2013), this bootstrap procedure is justified by some specific issues. First, it is well known that hedge fund returns do not follow normal distributions, and are rather often described as "dynamic" and "option-like". The issue lies in the fact that t-statistics are distributed as a Student law only if coefficient estimates are normally distributed; this happens either because the dependent variable itself is normally distributed, or because the sample size is large enough so that the law of large numbers takes effect. In our case, of course, neither assumption is justified, and as such, t-statistics do not follow the Student distribution. We thus need another way of evaluating their statistical significance. Also, as we already explained, some funds might exhibit statistically significant timing ability by random chance, what these authors call a "multiple comparison problem". Finally, "if fund returns within a strategy are correlated, the timing measures might not be independent across funds" and so we cannot assume that they may be described by independent distributions. The bootstrap procedure we use provides a solution to these issues. The basic idea is the following: by randomly resampling the data, and more specifically the regression residuals, we create fictive funds that, by construction, have the same exposures to strategies as the real funds, but have no strategy timing ability. By comparing the true cross-section of t-statistics with the fictive one, we are then able to determine the significance level of the true t-statistics. As we generate the fictive cross-sectional distribution of t-statistics from the true distribution, we do not make any assumption on the nature of that distribution, and so the resulting inference is valid even in small samples, or when fund returns are correlated with each other.

We now detail the bootstrap process, which is divided in five main steps. This explanation builds on Cao, Chen, Liang and Lo (2013), who manage to provide a simple, yet detailed description of this complex procedure.

1. Estimate the relevant model for fund p. In this case, it is the generalized strategy timing model as described by equation (20)[34]:

$$\begin{aligned} r_{p,t+1} = \alpha_p &+ \gamma_{p,1} MKT_{t+1}\left(MKT_{t+1} - \bar{f}_{t+1}\right) + \delta_{p,1} MKT_{t+1}\bar{f}_{t+1} \\ &+ \gamma_{p,2} SIZE_{t+1}\left(SIZE_{t+1} - \bar{f}_{t+1}\right) + \delta_{p,2} SIZE_{t+1}\bar{f}_{t+1} \\ &+ \gamma_{p,4} CRDT_{t+1}\left(CRDT_{t+1} - \bar{f}_{t+1}\right) + \delta_{p,4} CRDT_{t+1}\bar{f}_{t+1} + \sum\nolimits_{k}^{K} \beta_k f_{k,t+1} \\ &+ {\varepsilon'}_{p,t+1} \end{aligned}$$

   Then, store the coefficient estimates $(\hat{\alpha}_p, \hat{\gamma}_{p,1}, \hat{\delta}_{p,1}, \dots, \hat{\beta}_{p,1}, \hat{\beta}_{p,2}, \dots)$ as well as the time-series of residuals $(\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_{Tp})$. Note that Tp denotes the length of the return history of fund p, and may thus vary from fund to fund.

2. Randomly resample the residuals with replacement to create a new fictive residuals series $\hat{\varepsilon}^b_{p,t+1}$ where $b = 1, 2, \dots, B$ is the bootstrap iteration index. Then, construct the fictive fund's returns by plugging the fictive residuals series back into the model, and using previously estimated coefficients as exposures.

[34] For the sake of simplicity, we put the FH factors and their respective exposures at the end of the equation in reduced form. Here, we thus have *K=8*, and *k=1* for the *MKT* factor, *k=2* for the *SIZE* factor, and *k=4* for the *CRDT* factor.

However, set the timing coefficients to zero, so that the fictive fund does not actually possess any timing ability by construction:

$$r_{p,t+1}^{b} = \hat{\alpha}_p + \sum_{k}^{K} \hat{\beta}_{p,k} f_{k,t+1} + \hat{\varepsilon}_{p,t+1}^{b}$$

3. Estimate the relevant strategy timing model (in this case, equation (20)) again, this time with the fictive fund's returns constructed in step 2 as the dependent variable; and store the coefficient estimates as well as their t-statistics. Any nonzero t-statistic on a timing coefficient is then due to sampling variation, since by construction the fictive fund does not have any timing skills.
4. Repeat steps 1 to 3 for the entire sample of funds, so that the cross-sectional distribution of t-statistics on the timing coefficients is obtained for the sample of fictive funds. Then, estimate the relevant cross-sectional statistics for this fictive sample. In our case, and for each strategy timing coefficient, we compute the bottom as well as the top 1%, 3%, 5% and 10% quantiles for the distribution of t-statistics.
5. Repeat steps 1 to 4 for B iterations to generate the empirical distributions of selected cross-sectional statistics (e.g. the t-statistics quantiles for each timing coefficient). These empirical distributions are asymptotically valid estimations of the true distributions, as such, they may be used to conduct inference as long as B is large enough. Finally, for a given cross-sectional statistic, compute its empirical *p-value* as the percentage of bootstrapped cross-sectional statistics (i.e. which are estimated on fictive fund samples) which exceed the observed value in the original sample. For example, if the top 1% t-statistic on $\hat{\gamma}_{p,1}$ is equal to 4 in the original sample, then compute the number of bootstrapped top 1% t-statistics on $\hat{\gamma}_{p,1}$ which exceed the value 4, and divide this number by B. The obtained value is the empirical *p-value* for the top 1% t-statistic on $\hat{\gamma}_{p,1}$.

As such, by comparing the selected cross-sectional statistics with their empirical distributions, which are generated under the null hypothesis of no strategy timing abilities, we are able to determine if high observed values for t-statistics in the original sample are due to sampling variation, or to the effective existence of these timing abilities.

Note that in their paper, Cao, Chen, Liang and Lo (2013) set the number of bootstrap iterations to 10,000. As we were limited in computational power, we set this parameter to 5,000, but this choice should not have a significant impact on our conclusions. Note, also, that we use HAC robust t-statistics in this bootstrap procedure as well, i.e. we use the t-statistics computed from the Newey and West (1987) standard errors.

The results of this bootstrap procedure for the TM version of the generalized strategy timing model are presented in [table 13]. For each of the six timing coefficients, we report different quantiles of the distribution of t-statistics, as well as their empirical *p-values*. It appears that none of these cross-sectional statistics are statistically significant, even at the 10% confidence level. Any large t-statistic value that we previously observed in this model was thus entirely due to sampling variation, or "luck". Note that because our sample contains less than 100 funds, the top 1% t-statistic for a given timing coefficient is actually the t-statistic of the "best timer" regarding the associated variable. There is thus no possibility that even the best timers exhibit significant timing ability in our case.

As we previously observed that the HM version of this model detected a lower number of t-statistics exceeding the standard cut-off values, we do not run the bootstrap procedure again for this alternative specification, as it would surely lead to the exact same conclusion.

So far, we found no significant strategy timing ability at the fund level. However, this result was expected, since the generalized models did not detect such an ability at the aggregate level either. We now apply the same analyses to the *STX* models.

c. *STX factors*

We follow the same procedure as with the generalized models. We begin by estimating the FH model along with the *STX* factor for each fund, and store the *t*-statistic of the *STX* coefficient. We then compute the percentages of *t*-statistics which exceed the standard cut-off values, and report the results in [table 14]. We do so as well for the STX_L factor, as well as for the *STX* factor when alternative samples are used.

The results are quite interesting. Panel A shows that the percentage of funds for which the *STX* *t*-statistic is above 1.645 is around 15.5%, which is a higher proportion than the ones we obtained for most of the six timing coefficients in the generalized strategy timing models (cf. sub-section (a)). This is consistent with the *STX* models detecting strategy timing ability at the aggregate level, while the generalized models did not. However, the left tail of the *t*-statistics distribution is thicker than the right tail for the *STX* factor, which is surprising given the results obtained at the portfolio level. Indeed, around 22% of funds exhibit *t*-statistics on the *STX* coefficient which are below -1.645. Still, this observation may simply be the consequence of sampling variation; we discuss this possibility later on when we run the bootstrap procedure on these *t*-statistics.

We also notice that the STX_L factor exhibits large *t*-statistics less often than the *STX* factor. This is consistent with our previous observation that the STX_L factor had slightly less statistical significance at the aggregate level. Finally, as shown in panel B, the *t*-statistics distribution for the *STX* factor does not change in any significant way when alternative samples are employed, meaning that these results are once again robust to different data filtering rules.

d. *Bootstrapped t-statistics for STX factors*

We now run the bootstrap procedure, as detailed in sub-section (b), on the *STX* models. The only difference is thus that we do not use equation (20), but rather the FH model along with the *STX* factor. When we construct the fictive funds' returns from the resampled residuals, we thus set the coefficient on the *STX* factor to zero, so that these pseudo-funds possess no strategy timing ability by construction.

We run the procedure for both the *STX* and STX_L factors, and report the *t*-statistics quantiles as well as their empirical *p-values* in [table 15]. While most of the quantiles for the *STX* factor remain significant at the 5% level, it is quite the opposite for the STX_L factor. This is consistent with our previous intuition that the *STX* factor was more adequately specified than its linear counterpart; as such, we conjecture that multi-strategy managers, if they engage in strategy timing at all, react more aggressively to large deviations within their strategies' returns, and less aggressively to small ones. As explained before, this behaviour might originate from the existence of "switching costs", i.e. disinvestment costs incurred when switching strategies.

Concerning both factors, we notice that the *t*-statistics quantiles lose statistical significance as they lie further into the distribution's tails, i.e. the empirical *p-values* for the top and bottom 1% quantiles are higher than the *p-values* of the top and bottom 10%

quantiles, respectively. This result may appear quite illogical at first sight; but it is in fact completely possible to observe this pattern since the different t-statistics quantiles are in fact distinct cross-sectional statistics, which each have their own empirical distribution. More specifically, when looking at the quantiles on the *STX* factor, we observe that the top 3% and 1% quantiles are statistically insignificant, while the top 10% and 5% quantiles are significant at the 1% and 5% confidence levels, respectively. Although this pattern is theoretically observable, one would expect all the upper quantiles of the t-statistics distribution to be statistically significant if strategy timing ability was truly present at the fund level. Moreover, the bottom quantiles are all significant at the 5% confidence level, which implies that previously detected perverse timing ability is not due to sampling variation. This may seem quite troubling; however, we showed in aggregate-level tests that multi-strategy managers also react to publicly available information to reallocate capital between their strategies. Since we do not take this possibility into account in these fund-level tests, these unexpected results might actually come from the current model being misspecified. This explanation is quite plausible, as Ferson and Schadt (1996) showed that accounting for public information generally removed the evidence about perverse timing abilities, i.e. "unconditional" models tend to detect significant negative timing skills. As such, in the next sub-section, we control for managers reacting to public information and then run the bootstrap procedure again.

As an illustration of the unexpected results depicted in [table 15], [figure 3] plots the kernel density distributions of bootstrapped quantiles, along with their observed value in the original sample, for the *STX* factor. We note that these empirical distributions are obviously non-normal, which justifies the use of the bootstrap procedure to conduct statistical inference.

e. Controlling for publicly available information

We now control for reaction to public information in the *STX* models. We employ the same methodology to do so as in the aggregate-level tests: we estimate an AR(1) model on the *STX* factor and use the resulting residuals as the unpredictable component of the *STX* variable, which we label the STX^r. We then add the lagged *STX* factor to the model, which represents publicly available information. The model thus corresponds to equation (23). As usual, we estimate this model for each fund and then bootstrap the t-statistics with 5,000 iterations. Results of this process are presented in [table 16].

Quite interestingly, the upper quantiles of the STX^r factor t-statistics are all statistically insignificant, while those of the lagged *STX* factor are all significant at the 1% confidence level. In aggregate-level tests, we concluded that multi-strategy managers do use public information to adjust their exposures to strategies, but that they also possess true strategy timing ability. It appears here that while the former is indeed true, the latter is not, and previously detected strategy timing ability was the result of sampling variation, or *luck* on the part of some managers.

We note, however, that a remaining issue undermines this conclusion. Indeed, the lower quantiles of the STX^r factor t-statistics are all significant at the 5% confidence level, implying that this second model specification is still detecting significant perverse timing ability. Theoretically, negative timing ability makes no sense, as a manager who sets his exposures randomly at the start of each period would perform better.

In aggregate-level tests, we also controlled for the presence of market timing ability, which we consider is similar to absolute strategy timing ability. In unreported results, we do so as well at the fund level. Bootstrapped t-statistics then indicate that neither relative strategy timing ability – as measured by the *STX* factor – nor market timing ability are significant at the fund level. More specifically, controlling for market timing results in the lower quantiles of the *STX* factor *t*-statistics all being statistically insignificant as well.

As such, the perverse strategy timing ability we observe when controlling for public information may in fact be due to not controlling for market timing as well. As we did previously in aggregate-level tests, we thus control for publicly available information and market timing jointly in the next sub-section.

f. Controlling for public information and market timing jointly

We run the bootstrap procedure as usual, this time using the *STX* model where we control for both public information and market timing. As a reminder, the model is thus written as follows:

$$r_{t+1} = \alpha + \gamma STX_{t+1}^{r} + \varphi STX_t + \gamma_{MT} MKT_{t+1}^2 + \sum_{k=1}^{K} \beta_k f_{k,t+1} + \varepsilon_{t+1} \quad (24)$$

Where γ measures true relative strategy timing ability, φ measures reaction to lagged signals, i.e. public information, and γ_{MT} measures market timing ability, which we relate to absolute strategy timing ability.

Results of the bootstrap process are shown in [table 17]. The perverse strategy timing ability we previously detected disappears completely, while "positive" reaction to public information remains statistically significant. Indeed, the upper quantiles of the lagged *STX* factor *t*-statistics are the only ones to exhibit *p-values* that are inferior to the conventional significance levels. Meanwhile, none of the quantiles of the STX^r factor nor of the market timing factor are significant at any standard confidence level.

The conclusion is thus quite obvious: while multi-strategy managers are capable of using publicly available information to partly predict their strategies' returns, they have no superior ability or private information on this level, i.e. they do not possess "true" strategy timing skills. As such, it is unlikely that some multi-strategy managers consistently outperform their peers in timing strategies, as they all solely rely on public information to do so. Furthermore, multi-strategy managers do not seem to possess market timing skills either.

This result is actually in opposition with the conclusions of Agarwal and Kale (2007), who study and compare the performance of multi-strategy funds and funds-of-hedge-funds. Indeed, as both may be considered similar investment schemes in that they aim to provide the same core benefits to investors – mainly diversification across strategies – both types of funds should yield similar returns in well-functioning markets. They observe, however, that multi-strategy funds consistently outperform funds-of-hedge-funds on average. They provide four possible explanations to this result: (1) agency risk is higher for multi-strategy funds, thus investors require higher returns of their managers; (2) multi-strategy managers have higher investment flexibility, in particular, they can easily switch between strategies, while funds-of-hedge-funds are hindered by the lockup and restriction periods imposed by the funds they invest in; (3) funds-of-hedge-funds ultimately charge double-layered fees, thereby leading to lower net-of-fee performance; and (4) talented managers should rationally self-select in becoming multi-strategy managers, and as such these managers perform better on average. Since the authors control for the first three explanations in their regressions, and still observe superior performance on the part of multi-strategy managers, they conclude that "*the single explanation of self-selection [should be] the driving force behind the superior* *performance of the MS funds relative to FOFs*[35]*. This, in turn, suggests that MS fund managers may possess superior ability compared to FOFs. For example, one can think of MS fund managers having greater timing ability in allocating capital across different hedge fund strategies. Self-selection by better ability managers into MS funds is consistent with there being fewer MS funds because there is a likely to be a limited supply of better ability managers*"[36].

As such, these authors conjecture that multi-strategy managers should possess "true" strategy timing ability if they indeed self-select in managing multi-strategy funds. However, our results show that multi-strategy managers do not, in fact, possess such an ability, and rather merely react to public information to predict their strategies' returns. Although the explanation of self-selection stays valid, since managing multiple strategies simultaneously without trying to time them may still require superior skill, we believe that the differences in *flexibility* they discuss might, in fact, also significantly explain the superior performance of multi-strategy funds relative to funds-of-hedge-funds. Indeed, even if "switching costs" are substantial, multi-strategy managers would still experience less difficulty in reallocating capital across strategies, since funds-of-hedge-funds are legally limited by lock-up and restriction periods, or similar schemes. One might imagine that if the two types of managers observe the same public strategy timing signal, e.g. one specific strategy has performed badly over the last two months, the multi-strategy manager will be able to reallocate capital almost immediately – although it may prove costly – while the fund-of-funds manager will have to notice the relevant funds that he wishes to disinvest, and then wait several months before finally being able to do so. The point is that even if managers time strategies solely based on public information, differences in *investment flexibility* between the two types of funds may still explain the observed differences in performance.

In their paper, Agarwal and Kale (2007) argue that they control for *investment flexibility* in their multivariate regressions, and thus ultimately reject this explanation. However, they do so by adding each fund's lock-up and restriction period as control variables, which, in fact, does not allow to control for the difference we just discussed. Indeed, the true difference lies in the fact that funds-of-hedge-funds are limited by the lock-up and restriction periods *of the funds they invest in*, and so adding their own lock-up and restriction periods as control variables does not proxy this reality. In particular, the summary statistics they provide on their sample of funds show that funds-of-hedge-funds do not impose longer lock-up and/or restriction periods than multi-strategy funds, which should be the case if these variables proxied correctly the phenomenon we are discussing[37]. As such, we believe that their results are not sufficient to reject the possibility of the *investment flexibility* being a driver behind the superior performance of multi-strategy funds relative to funds-of-hedge funds. Taking this point into account, our results are no longer inconsistent with theirs.

---

[35] "MS" is an abbreviation for "multi-strategy", and "FOFs" for "funds-of-hedge-funds".

[36] Quoted from Agarwal and Kale (2007), pages 18 to 19.

[37] In fact, funds-of-hedge-funds generally hold significant amounts of cash so as to be able to redeem shares if requested by investors without having to impose extremely large restriction and lock-up periods. Using cash holdings as an additional explanatory variable may thus be an effective way to control for funds-of-hedge-funds being hindered by the restriction and lock-up periods of the funds they invest in. However, this observed behaviour on the part of funds-of-hedge-funds supports our claim that their own restriction and lock-up periods do not proxy correctly their limited investment flexibility.

Earlier (cf. section 1), we conjectured that strategy timing ability should represent the main advantage of multi-strategy funds over funds-of-hedge-funds. As our results show that *true* strategy timing ability does not exist among multi-strategy managers, we are now able to refine this statement: the *capability* that multi-strategy funds have in dynamically adjusting their exposures to strategies **based on the use of public information** represents their main advantage over funds-of-hedge-funds. By *capability*, we thus do not mean *skill*, but rather *possibility*. Indeed, even if funds-of-hedge-funds know how to react, as well, to public information concerning the performance of strategies, they *may not* do so, since they are legally limited by the lock-up and restriction periods of the funds they invest in.

Also, and as we mentioned earlier, it is unlikely that some multi-strategy managers persistently outsmart their peers in timing strategies, as they all rely on publicly available information to do so. Still, our results seem to indicate the opposite: as shown in [table 17], the top 1% *t*-statistic quantile on the STX_{lag} factor is more than twice as high in value as the corresponding top 10% *t*-statistic quantile. Some managers thus appear to react to public information *more effectively* than others, leading us to wonder whether hedge funds investors could profit from this observation by selectively investing in funds that best react to these public signals. To answer this question, we conduct an out-of-sample persistence analysis for this "reaction skill". We start in March 2003 by estimating equation (24) for each available fund using the past 36 months of the return series. We then assign each fund to a decile portfolio based on its estimated coefficient for the STX_{lag} factor – that is, its ability to react to public signals which are predictive of the strategies' returns. Portfolio 1 contains the best "timers" – that is, funds which react to public information most effectively – while portfolio 10 contains the bottom "timers". The decile portfolios are held for a subsequent period of either 3, 6, 9, or 12 months, and the process is then repeated. Finally, we estimate the original FH 8-factors model using each portfolio's return series and report the out-of-sample alphas in [table 18]. If some multi-strategy managers were persistently able to react more effectively to public information when adjusting their exposures to strategies, we would expect the upper decile portfolios to exhibit higher alphas than the lower decile portfolios. However, results clearly show that it is not the case. This confirms our initial intuition: as multi-strategy managers use publicly available information to reallocate capital across strategies, none is able to persistently outperform others in doing so. This in turn implies that hedge fund investors should not attempt to select multi-strategy funds by evaluating their strategy timing abilities, whether based on private or public information.

Finally, concerning the existence of "switching costs", note that we proxy public information in our model by using the lagged value of the *STX* factor. As such, this variable is constructed in the same manner as the original *STX* factor, and thus models a *convex* relationship between the lagged dispersion of strategies' returns and the fund performance. We thus conjecture that although multi-strategy managers react to public timing signals concerning their strategies, they do so more aggressively when their strategies' returns are predicted to strongly diverge, and may overlook predictions about small divergences. This behaviour may once again be explained by the presence of these "switching costs".

g. *The impact of data biases*

As explained in section 3, our data is subject to potentially substantial biases which may affect our results. As our sample is constituted of live funds only, it is prone to survivorship bias. Also, we use gross-of-fee returns, which means that we analyse the base investment performance of multi-strategy funds, and not the performance they pass on to investors. Finally, we do not attempt to correct for the backfill bias.

Nevertheless, these issues eventually reinforce our previous conclusions. Indeed, our results show that the multi-strategy funds in our sample possess neither strategy timing skills nor market timing skills once we account for luck and public information. The consequence of the survivorship bias is that only the best-performing funds, i.e. the survivors, find themselves into the final sample. The backfill bias has a similar impact, as backfilled returns would logically have to illustrate impressive performance for the process to have some interest. Finally, net-of-fees returns will, by construction, always be lower than gross-of-fees returns.

The fact that we find no significant timing skills in our sample, while we know the latter is affected by those three significant *upward* biases, thus raises serious doubts about multi-strategy hedge funds being able to time strategies overall. Indeed, if even the best-performing funds in this category, while having their returns look as good as possible, do not exhibit significant superior abilities on this level, then it seems rather unlikely that other funds may do so. In addition, our results have been shown to be robust to alternative data filtering rules, as well as to alternative model specifications. The external validity of our results should thus not be underestimated, despite *ex-ante* data issues.

VII. Conclusion

In this Master's thesis, we analyse an important component of the overall attractiveness of multi-strategy hedge funds, which is their supposed ability to time strategies. Indeed, these funds can be expected to exhibit superior skill in predicting their strategies' returns and consequently increasing their exposure to the best-performing ones, as generic hedge fund strategies have been shown to generate time-varying performance. We study this question using a small sample of 90 multi-strategy funds from the Thomson-Reuters TASS database, on the period spanning from January 1994 to December 2017. We conduct analyses at both the aggregate and fund level, and use two types of evaluation frameworks. Our results ultimately show that even the best-performing multi-strategy managers in our sample do not possess *true* strategy timing skills.

In particular, we find that multi-strategy managers do react to publicly available information which may help them predict their strategies' returns, but do not possess superior – i.e. private – information on this level. As such, any strategy timing ability they exhibit holds no real value to investors, since such investors may implement the process as well using lagged information about each strategy's returns. Moreover, using an adequate bootstrap procedure, we show that any true timing ability that is initially detected is in fact the consequence of sampling variation, i.e. *luck* on the part of some multi-strategy managers.

Theoretically, we distinguish two components of the strategy timing process: *relative* strategy timing, which relates to comparing the strategies' returns and investing more in the best-performing ones, and *absolute* strategy timing, which relates to market timing since it is about switching to risk-free investments when all strategies yield negative excess returns overall. Our empirical results show that multi-strategy managers possess neither components of strategy timing ability, which is consistent with the existing literature.

Additionally, we conjecture that when multi-strategy managers react to public information to reallocate capital between strategies, they do so more aggressively if the strategies' returns are predicted to strongly diverge. On the opposite, managers may overlook small differences in their strategies' returns. A plausible explanation for this behaviour is simply that switching strategies is costly for managers, since a disinvestment process is required for some of the strategies in doing so.

Finally, we believe that our results may exhibit high external validity as our sample is in fact affected by significant *upward* data biases. Indeed, as we do not detect significant timing ability even among the best-performing funds, while their returns look as good as possible, it seems unlikely that different results would be obtained using other, more complete fund samples. Moreover, our analysis appears robust to alternative data filters and model specifications.

To conclude, this study thus raises serious doubts about multi-strategy managers being able to provide superior value to their investors by timing strategies. On the contrary, it rather appears that this supposed strategy timing ability is fundamentally a marketing argument, which allows these managers to better differentiate themselves from funds-of-hedge-funds. However, as multi-strategy hedge funds have greater flexibility in managing strategies than funds-of-funds, they are better able to react to changing market conditions, thereby explaining their documented outperformance of the latter. This investment flexibility thus ultimately appears as the one real advantage of multi-strategy funds over other hedge fund categories.

Appendix A. Strategy timing as a marketing argument.

We hereunder provide examples of mission statements coming from our multi-strategy funds which explicitly refer to the ability to time strategies. This is only meant to be an illustration, and we do not aim to generate the impression that all the multi-strategy funds in our sample declare timing their strategies. Actually, only a dozen of funds in our final sample explicitly do so.

- Fund 13: "The Fund invests in a weighted portfolio of investment strategies which have a low correlation with one another. This leads to improved risk adjusted returns through diversification. The weighting of each strategy is periodically adjusted according to its performance as well as the evolution in correlations between the strategies".
- Fund 14: "The Fund dynamically combines trade strategies that it believes will perform in either a weaker or stronger economic environment with strategies that will perform in unchanged markets".
- Fund 50: "The investment Manager expects that the Segregated Portfolio will allocate at any time approximately 50-70% of the assets under management of Segregated Portfolio to the Fixed income sub-strategy, with the balance of the assets under management being allocated [either] to cash, to the Foreign Exchange sub strategy, or to the Equities sub strategy in any proportion as determined by the Investment Manager in its sole discretion".
- Fund 106: "Investment Strategies include: Convertible Arbitrage, Credit Arbitrage, Capital Structure Arbitrage, Volatility/Options, RMBS and Special Situations among others. The Management selects investments for which it expects superior risk-adjusted returns by maintaining the flexibility to move capital among different investment strategies".

Appendix B. Analysis of the funds' mission statements.

To detect possible issues in the data as well as to get some insights into the multi-strategy funds' trading styles, we analyse the mission statements contained within the TASS database. Of course, although we employed a pre-defined set of rules to conduct this analysis, a subjectivity bias is likely to be present in the generated data. Again, the goal is to provide insights, and surely not definitive conclusions, into the multi-strategy hedge funds trading styles.

For each fund, we thoroughly read the mission statement and then make, if possible, conclusions on several points: (1) whether or not the fund mainly or only trades in one market; (2) if not, in which markets does it trade, or which strategies does it use; (3) whether or not the fund is focused on a geographic region; (4) whether or not the fund states it may use options or other derivative products; (5) whether or not the fund states that it only invests in other funds.

We also sort the funds by name to detect duplicates. Funds that share a common name with one or several other funds are marked as "duplicates" and we delete the copies in each group by following the methodology described in the next section.

For point (1), we aim to determine which funds are not likely to use multiple strategies or to trade in multiple markets, and which would thus not correspond to a classic multi-strategy fund. We hereunder provide two examples of funds that scored positive on this aspect.

- Fund 131: "The Fund aims to generate superior risk-adjusted returns by investing in the equities of growth companies both long and short."
- Fund 10: "The Fund aims to generate significant returns by investing in Digital Currencies, Blockchain, distributed ledgers, and related technologies."

For point (2), we aim to determine what the focus markets of the funds or the main strategies it might use are. We infer our conclusions from the words used in the mission statements. We provide two examples below.

- Fund 4: "The Fund aims to capture short-term momentum within a diverse group of commodity markets". This fund scored positive on the "Commodities / CTAs" and "Trend following" variables.
- Fund 35: "The Fund primarily aims to systematically capture statistical arbitrage and market neutral arbitrage opportunities in the equity and currency markets". This fund scored positive on the "Equity", "Statistical / quantitative arb." and "Currencies" variables.
- Fund 66: "The Fund seeks to consider all type of insurance perils and structures, with a preference for low frequency catastrophic events". This fund scored positive on the "Event driven / distressed" variable.

Point (3) seems relatively straightforward. A good example would be fund 46: "The Fund aims to achieve consistent absolute returns by investing primarily in equities listed in the Greater China (including Taiwan, Hong Kong and the PRC) and U.S."

In the same manner, point (4) and (5) are straightforward. We provide some examples of funds that were deemed to be funds-of-hedge-funds below.

- Fund 16: "The Fund aims to achieve medium-term capital growth by trading the Fund through investment in various instruments and equity or debt securities, including but not limited to investment in regulated or unregulated collective investment schemes or other pooled vehicle(s) managed by the Investment Manager or its affiliates. The Fund may also hold cash or equivalent instruments."
- Fund 1: "To pursue capital appreciation, regardless of market conditions, principally by investing in other collective investment schemes and investments funds."
- Fund 62: "The Fund aims to achieve capital appreciation through the allocation of assets to Underlying Funds which pursue Alternative Investment Strategies."

Appendix C. Deletion of duplicate funds in the data.

As previously stated, we identified 66 "duplicate" funds through the preliminary data analysis. These duplicates share a common name and are generally differentiated through letters or numbers. We deduct that these funds probably operate under the same investment management company but are diversified in some way. Some funds are also distinguished within their group through legal suffixes. We deduct these funds are clones created to leverage some legal advantages.

For each group of duplicates, we conduct a linear correlation analysis. Funds that exhibit correlation coefficients superior to 0.8 are considered real duplicates and are marked for deletion. Funds that exhibit lower correlation with each other are further analysed. Unless some special data feature makes it clear the funds are in fact real duplicates, we consider the funds sufficiently distinguish themselves from one another to be kept in the final sample. Special data features include:

- Low correlation coefficients because of a small number of months where returns differ significantly. When these months are omitted, the correlation coefficient goes above 0.99.
- Low correlation coefficients because the original fund manager created other funds at some point in time and backfilled the original fund's returns into the newer fund's history. As such, the return history is identical up until the splitting point.

Group of funds that are marked for deletion are treated as follows. If the length of return histories differs among the group, we only keep the fund which has the longest track record. If we cannot make this choice because some or all funds in the groups have the same inception date, we choose randomly the fund which will be kept in the sample by using a discrete pseudo-random number generator. Finally, out of the initial 66 funds that were marked for deletion, we keep 24 into the final sample.

Appendix D. Selection procedure for time-varying coefficients.

As explained in section 4, the decision regarding which coefficients in the FH model are allowed to vary over time is quite important. Based on previous insights, we allow the coefficients on the *MKT*, *SIZE*, and *CRDT* factors to do so. However, we seek to validate this choice by testing other possible specifications, which are detailed in the table below. Note that specification 1 is the one we originally proposed, and ultimately used. For each specification, we test the extended, relative and generalized strategy timing models separately. We also test both the TM and HM versions of our models to allow for a global comparison. We thus have 36 different specifications to compare in this test.

Specification	**Time-varying coefficients**				
	MKT	SIZE	BOND	CRDT	EMRG
1	X	X		X	
2	X	X	X	X	
3	X			X	
4	X	X		X	X
5	X		X		X
6		X		X	X

We regress the monthly returns of the equally-weighted portfolio of funds on each corresponding set of factors and look at the R^2 and adjusted R^2 statistics. The results are presented in [table 4]. The specifications are labelled by the letter *G*, *R*, or *E*, depending on which variant of the model they correspond to – the generalized strategy timing model, the relative strategy timing model, or the extended model, respectively. They are then labelled by the specification number, as detailed in the table above. Panel A presents the goodness-of-fit of TM versions of the models, while panel B presents results for the HM counterparts of these. In each panel, specifications are sorted from the highest adjusted R^2 to the lowest.

All the specifications but one yield improvements, although of different magnitudes, over the adjusted R^2 of the basic FH eight-factor model. Overall, the TM model versions exhibit higher goodness-of-fit than the HM model versions. Also, the relative strategy timing specifications yield the smallest improvements in terms of adjusted R^2. The extended specifications yield better improvements, and the generalized strategy timing specifications exhibit the highest adjusted R^2 statistics. The only exceptions to this overall ranking are specifications 3 and 5, of which the extended, relative, and generalized versions all perform relatively badly. These results allow us to confirm our preference for specification 1, the TM version of which yield the third highest improvement in terms of adjusted R^2. Specifications 2 and 4 yield even higher improvements, however, they both require the use of two additional variables compared to specification 1, and one would generally prefer a parsimonious model. Furthermore, the increase in R^2 over specification 1 they exhibit is statistically insignificant[38]. We thus settle for specification 1, in particular its TM generalized version, which we use in all our empirical analyses. We keep using the HM counterpart of this model as a robustness test.

[38] We test this using the classic multiple linear restrictions F-test on the R^2.

References

Admati, A. R., Bhattacharya, S., Pfleiderer, P., and Ross, S. A. (1986), "On timing and selectivity", *The Journal of Finance*, 41(3), 715-730.

Agarwal, V., Fung, W. H., Loon, Y. C., and Naik, N. Y. (2011), "Risk and return in convertible arbitrage: Evidence from the convertible bond market", *Journal of Empirical Finance*, 18(2), 175-194.

Agarwal, V., and Kale, J. R. (2007), "On the relative performance of multi-strategy and funds of hedge funds", CFR Working Paper, No. 07-11.

Agarwal, V., Mullally, K., and Naik, N. Y. (2015), "Hedge funds: A survey of the academic literature", *Foundations and Trends in Finance*, forthcoming.

Agarwal, V., and Naik, N. Y. (2000), "Performance evaluation of hedge funds with option-based and buy-and-hold strategies".

Agarwal, V., and Naik, N. Y. (2004), "Risks and portfolio decisions involving hedge funds", *The Review of Financial Studies*, 17(1), 63-98.

Agarwal, V., Ruenzi, S., and Weigert, F. (2017), "Tail risk in hedge funds: A unique view from portfolio holdings", *Journal of Financial Economics*, 125(3), 610-636.

Baquero, G., Ter Horst, J., and Verbeek, M. (2005), "Survival, look-ahead bias, and persistence in hedge fund performance", *Journal of Financial and Quantitative analysis*, 40(3), 493-517.

Cai, L., & Liang, B. (2012), "On the Dynamics of Hedge Fund Strategies", *The Journal of Alternative Investments*, *14*(4), 51-68.

Cao, C., Chen, Y., Liang, B., and Lo, A. W. (2013), "Can hedge funds time market liquidity?", *Journal of Financial Economics*, 109(2), 493-516.

Chan, L. K., Chen, H. L., and Lakonishok, J. (2002), "On mutual fund investment styles", *The Review of Financial Studies*, 15(5), 1407-1437.

Chen, Y. (2006), "Timing ability in the focus market of hedge funds".

Chen, Y., and Liang, B. (2007), "Do market timing hedge funds time the market?", *Journal of Financial and Quantitative Analysis*, 42(4), 827-856.

Coggin, T. D., Fabozzi, F. J., and Rahman, S. (1993), "The investment performance of US equity pension fund managers: An empirical investigation", *The Journal of Finance*, 48(3), 1039-1055.

Durbin, J. (1970), "Testing for serial correlation in least-squares regression when some of the regressors are lagged dependent variables", *Econometrica*, 38(1), 410-421.

Dybvig, P. H., and Ross, S. A. (1985), "Differential information and performance measurement using a security market line", *The Journal of Finance*, 40(2), 383-399.

Efron, B. (1979), "Computers and the theory of statistics: thinking the unthinkable", *SIAM review*, 21(4), 460-480.

Fama, E. F., and French, K. R. (1992), "The cross-section of expected stock returns", *The Journal of Finance*, 47(2), 427-465.

Ferson, W. E., and Schadt, R. W. (1996), "Measuring fund strategy and performance in changing economic conditions", *The Journal of Finance*, 51(2), 425-461.

Fung, W., and Hsieh, D. A. (1997), "Empirical characteristics of dynamic trading strategies: The case of hedge funds", *The Review of Financial Studies*, 10(2), 275-302.

Fung, W., and Hsieh, D. A. (2000), "Performance characteristics of hedge funds and commodity funds: Natural vs. spurious biases", *Journal of Financial and Quantitative Analysis*, 35(3), 291-307.

Fung, W., and Hsieh, D. A. (2001), "The risk in hedge fund strategies: Theory and evidence from trend followers", *The Review of Financial Studies*, 14(2), 313-341.

Fung, W., and Hsieh, D. A. (2002), "Risk in fixed-income hedge fund styles", *Journal of Fixed Income*, 12(2), 6-27.

Fung, W., and Hsieh, D. A. (2004), "Hedge fund benchmarks: A risk-based approach", *Financial Analysts Journal*, 60(5), 65-80.

Fung, W., and Hsieh, D. A. (2005), "Extracting portable alphas from equity long/short hedge funds", in *The world of hedge funds: characteristics and analysis* (pp. 161-180).

Fung, W., Hsieh, D. A., Naik, N. Y., and Ramadorai, T. (2008), "Hedge funds: Performance, risk, and capital formation", *The Journal of Finance*, 63(4), 1777-1803.

Fung, H. G., Xu, X. E., and Yau, J. (2002), "Global hedge funds: risk, return, and market timing", *Financial Analysts Journal*, 58(6), 19-30.

Getmansky, M., Lo, A. W., and Makarov, I. (2004), "An econometric model of serial correlation and illiquidity in hedge fund returns", *Journal of Financial Economics*, 74(3), 529-609.

Goetzmann, W. N., Ingersoll, J., and Ivković, Z. (2000), "Monthly measurement of daily timers", *Journal of Financial and Quantitative Analysis*, 35(3), 257-290.

Griffin, J. M., and Xu, J. (2009), "How smart are the smart guys? A unique view from hedge fund stock holdings", *The Review of Financial Studies*, 22(7), 2531-2570.

Henriksson, R. D., and Merton, R. C. (1981), "On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills", *The Journal of Business*, 54(4), 513-533.

Jagannathan, R., and Korajczyk, R. A. (1986), "Assessing the market timing performance of managed portfolios", *The Journal of Business*, 59(2), 217-235.

Jensen, M. C. (1972), "Optimal utilization of market forecasts and the evaluation of investment performance".

Jiang, W. (2003), "A nonparametric test of market timing", *Journal of Empirical Finance*, 10(4), 399-425.

Mitchell, M., and Pulvino, T. (2001), "Characteristics of risk and return in risk arbitrage", *The Journal of Finance*, 56(6), 2135-2175.

Newey, W. K., and West, K. D. (1987), "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix", *Econometrica*, 55(3), 703-708.

Lehmann, B. N., and Modest, D. M. (1987), "Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons", *The Journal of Finance*, 42(2), 233-265.

Treynor, J., and Mazuy, K. (1966), "Can mutual funds outguess the market?", *Harvard Business Review*, 44(4), 131-136.

White, H. (1980), "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", *Econometrica*, 48(4), 817-838.

Table 1.

Summary statistics of the data.

Panel A contains summary statistics for fund returns and available fund characteristics. "All funds" statistics are computed on the entire return sample, without any differentiation being made between funds or periods. A total of 10236 monthly returns are available in the sample. The last column (ρ1) contains the average autocorrelation coefficient across all funds. It is obtained by running an AR(1) regression for each return series. "E.W. Portfolio" contains statistics for the equally-weighted portfolio of live funds. The last column contains the autocorrelation coefficient obtained from the AR(1) regression. These return statistics are not computed in excess of the risk-free rate. "Age" is the funds' age in months. "Min. Investment" is minimum initial investment in thousands U.S. dollars.

Panel B contains summary statistics for the factors used in the Fung and Hsieh (2004) 8-factors model. MKT is the total monthly return on the S&P 500. SIZE is the difference between the total monthly return of the Russel 2000 index and the S&P 500. BOND is the monthly change in the 10-year constant maturity Treasury bill rate. CRDT is the monthly change in the spread between Moody's Baa corporate bond index and the 10-year constant maturity Treasury bill rate. The PTFS factors are the primitive trend-following strategy factors from Fung and Hsieh (2001), respectively in the bond, currencies, and commodities markets. EMRG is the total monthly return on the MSCI Emerging Markets index.

	Mean	Std. Dev.	Min	Max	Skew.	Kurtosis	% Pos.	ρ_1
Panel A. Summary statistics for fund returns and characteristics.								
All funds	0.751[a]	4.432	-61.44	135.991	3.515	121.961	66.15	0.149[b]
E.W. Portfolio	0.980[c]	1.929	-6.4	9.251	0.307	5.305	71.43	0.262[d]
Age	113.733	66.757	24	287				
Min. Investment	1864.454	6058.862	2.5	50000				
Panel B. Summary statistics for the Fung and Hsieh (2004) model factors.								
MKT	0.850	4.167	-16.795	10.929	-0.690	4.313	66.20	
SIZE	0.019	3.220	-16.376	18.406	0.298	7.858	49.83	
BOND	-0.011	0.261	-1.080	0.950	-0.009	4.219	49.13	
CRDT	0.000	0.205	-0.790	1.530	1.340	14.925	48.43	
$PTFS_{BD}$	-1.844	15.097	-26.630	68.860	1.347	5.386	37.28	
$PTFS_{FX}$	-0.902	19.291	-31.810	90.270	1.371	5.607	39.02	
$PTFS_{COM}$	-0.542	14.079	-24.650	64.750	1.099	4.685	41.46	
EMRG	0.673	6.539	-29.168	17.156	-0.685	5.088	58.54	

[a] The mean monthly return goes up to 0.82% when we exclude the period spanning from September 2008 to December 2008 from the computation. It goes down to 0.67% when we use returns in excess of the risk-free rate, which we proxy by the 3-months T-bill rate.

[b] Across the funds sample, 48.9% of AR(1) autocorrelation coefficients are significant at the 5% level.

[c] The mean monthly return of the equally-weighted portfolio of funds goes down to 0.77% when we use returns in excess of the risk-free rate.

[d] Significant at the 1% level.

Table 2.

Results of the analysis of the funds' mission statements.

For each aspect, the percentage of answers that could be given with certainty is shown in the first column. The next two columns contain the percentages of positive and negative answers, relative to the whole sample (thus including funds for which no definite answer could be given).

"Single focus market" defines whether or not the fund explicitly states that it trades in one specific market. "Fund-of-funds" defines whether or not the fund explicitly states that it primarily or mainly invest in other funds (if it is not the case, it is considered a normal hedge fund). Rows 2 to 11 contain market and/or strategy indicators. They can relate to a focus market, a classic hedge fund strategy, or both. For example, "Equity" defines whether or not the fund states that it trades in the equity markets or employs an equity-focused strategy (long/short, dedicated short bias). "Geographical focus" defines whether or not the fund explicitly states that it trades with some focus on a geographical location. "Use of derivatives" defines whether or not the fund explicitly states that it may use options and other derivative products. "Leveraged" is not related to the analysis of the statements as it comes from Lipper TASS. The variable defines whether or not a fund uses leverage.

	Percentage Definite Answer	% Yes	% No
Single focus market	60.00	23.33	36.67
Fund-of-funds	97.78	6.67	86.67
Equity	43.33	43.33	
Fixed-income	33.33	33.33	
Convertible arb.	7.78	7.78	
Global macro	7.78	7.78	
Commodities / CTAs	16.67	16.67	
Trend following	4.44	4.44	
Statistical / Quantitative arb.	13.33	13.33	
ED / Distressed	4.44	4.44	
ED / Risk arbitrage	6.67	6.67	
Currencies	15.56	15.56	
Geographical focus	26.67	12.22	14.44
Use of derivatives	18.89	18.89	0.00
Leveraged	38.89	37.78	1.11

Table 3.

Regressions of the equally-weighted portfolio's returns on different sets of factors.

Returns are not computed in excess of the risk-free rate. The first column contains the intercept estimate. The last column contains the value of the R^2 statistic. Coefficients in bold are statistically significant at the 5% level, and *t*-statistics are shown in parentheses. The full sample period is used, from Feb. 1994 to Dec. 2017, totalling 287 monthly return observations.

Panel A variables: *CA* is the Convertible Arbitrage index, *EM* is the Emerging Markets index, *EMN* is the Equity Market Neutral index, *ED* is the Event Driven index, *FI* is the Fixed-Income index, *GM* is the Global Macro index, *LSE* is the Long/Short Equity index, and *MF* is the Managed Futures index. The Dedicated Short Bias index was dropped due to some part of the return series being incomplete for the period. **Panel B variables**: *EM* is the Emerging Markets index, *EH* is the Equity Hedge index, *ED* is the Event Driven index, *GM* is the Macro index, *RV* is the Relative Value index, *CA* is the Convertible Arbitrage sub-index. **Panel C variables** are the Fung and Hsieh (2004) factors as defined in the description of [table 1].

Panel A: Credit Suisse Hedge Fund Indices

alpha	*CA*	*EM*	*EMN*	*ED*	*FI*	*GM*	*LSE*	*MF*	R^2
0.0046	**0.3268**	0.0321	0.0049	0.0019	-0.0218	0.0326	**0.3272**	**0.1317**	0.549
(5.39)	(4.28)	(1.05)	(0.15)	(0.02)	(-0.25)	(0.83)	(6.75)	(5.12)	

Panel B: HFRI indices

alpha	*EM*	*EH*	*ED*	*GM*	*RV*	*CA*			R^2
0.0043	0.0104	**0.2743**	-0.0125	**0.3783**	-0.0218	**0.2484**			0.578
(4.66)	(0.30)	(4.09)	(-0.12)	(7.37)	(-0.13)	(3.02)			

Panel C: Fung & Hsieh (2004) 8-factors model

alpha	*MKT*	*SIZE*	*BOND*	*CRDT*	*PTFS_BD*	*PTFS_FX*	*PTFS_COM*	*EMRG*	R^2
0.0081	**0.1110**	**0.0783**	**-1.637**	**-1.6192**	0.0032	0.0087	0.0073	**0.1042**	0.405
(8.84)	(3.47)	(2.66)	(-3.97)	(-2.72)	(0.50)	(1.66)	(1.08)	(4.87)	

Table 4.

Comparison of the different strategy timing model specifications and their goodness-of-fit.

We regress the returns of the equally-weighted portfolio of funds on the sets of factors corresponding to the different model specifications, as detailed in [Appendix D]. Each model specification is labelled as *G*-, *R*-, or *E*-, which indicates whether they follow the generalized strategy timing specification, the relative strategy timing specification, or the extended specification, as explained in the methodology section. Each model specification is also labelled by a number between 1 and 6, which indicates which of the FH model factors are subject to strategy timing. The details of this terminology is found in [Appendix D].

For each model specification, we report the R^2 statistic and its adjusted counterpart, the number of degrees of freedom in the portfolio regression, and the results of a multiple linear restrictions F-test, with the FH 8-factors model taken as the restricted model. As such, this test measures if the R^2 improvement over the FH model for a given specification is statistically significant. *p-values* for this F-test are reported in the last column.

Panel A presents these statistics for the TM versions of the models, while panel B does so for the HM counterparts. In each panel, the specifications are sorted on their adjusted R^2 statistic, from highest to lowest.

Model	R^2	$\bar{R}^2$	*df*	F-test	*p-value*
FH 8-factors	0.417	0.400	278	-	-
Panel A. TM versions of the strategy timing model					
G4	0.482	0.451	270	4.277	0.0001
G2	0.479	0.448	270	4.051	0.0001
G1	0.474	0.447	272	4.949	0.0001
E6	0.465	0.443	275	8.219	0.0000
G6	0.470	0.443	272	4.601	0.0002
E4	0.465	0.441	274	6.176	0.0001
E1	0.460	0.438	275	7.309	0.0001
E2	0.460	0.436	274	5.487	0.0003
R2	0.451	0.427	274	4.261	0.0023
R6	0.448	0.426	275	5.156	0.0018
R4	0.448	0.424	274	3.902	0.0042
R1	0.444	0.422	275	4.603	0.0037
G5	0.441	0.413	272	2.017	0.0636
R5	0.430	0.407	275	2.158	0.0932
R3	0.426	0.405	276	2.270	0.1052
E3	0.424	0.403	276	1.855	0.1583
E5	0.425	0.402	275	1.269	0.2852
G3	0.426	0.401	274	1.127	0.3438
Panel B. HM versions of the strategy timing model.					
G2	0.463	0.431	270	2.908	0.0040
G1	0.458	0.431	272	3.502	0.0024
E6	0.449	0.427	275	5.395	0.0013
E4	0.450	0.425	274	4.101	0.0030
E1	0.445	0.423	275	4.679	0.0033
G6	0.449	0.421	272	2.680	0.0152
E2	0.445	0.421	274	3.505	0.0082
G4	0.452	0.420	270	2.215	0.0267
R2	0.440	0.416	274	2.897	0.0225
R1	0.434	0.411	275	2.820	0.0394
R6	0.430	0.407	275	2.136	0.0960
G3	0.430	0.405	274	1.640	0.1644
R4	0.430	0.405	274	1.567	0.1833
E5	0.426	0.403	275	1.563	0.1987
R3	0.423	0.402	276	1.442	0.2382
E3	0.423	0.402	276	1.432	0.2407
R5	0.423	0.400	275	1.080	0.3579
G5	0.427	0.397	272	0.797	0.5733

Table 5.

Summary statistics for the timing variables in the selected generalized strategy timing model.

Panel A presents summary statistics for timing variables in the TM version of the model. All statistics are multiplied by a factor 100, excepted of course for the skewness and kurtosis estimates. However, because of the way they are constructed, they are not expressed in percent, but rather in squared percentages. Each variable is named after the factor on which it is based, and is followed by *_ga* (for gamma) for the relative timing variables, or by *_de* (for delta) for the absolute timing variables. For example, the MKT_{ga} and MKT_{de} variables are defined as follows:

$$MKT_{ga} = MKT \times (MKT - \bar{f})$$
$$MKT_{de} = MKT \times \bar{f}$$
$$\bar{f} = (MKT + SIZE + CRDT)/3$$

Panel B presents the same summary statistics for timing variables in the HM version of the model. The terminology follows the same principle as in panel A. Note, however, that because they are constructed differently, the HM timing variables are expressed in percent. We thus report the summary statistics in percentages, excepted for the skewness and kurtosis estimates. Taking back the example above, the timing variables on *MKT* are now defined as:

$$MKT_{ga} = MKT \times I(MKT - \bar{f} > 0)$$
$$MKT_{de} = MKT \times I(\bar{f} > 0)$$

Where *I* is the usual indicator variable which takes the value 1 if the condition in parentheses is met, and 0 otherwise.

Variable	Mean	Std. Dev.	Min	Max	Skew.	Kurtosis	% Pos.
Panel A. TM version of the generalized strategy timing model.							
MKT_{ga}	0.116	0.200	-0.020	1.750	3.783	23.672	86
MKT_{de}	0.061	0.119	-0.207	1.084	4.620	34.173	84
$SIZE_{ga}$	0.066	0.209	-0.097	2.390	9.365	101.776	77
$SIZE_{de}$	0.038	0.086	-0.084	0.998	5.864	58.196	75
$CRDT_{ga}$	0.002	0.009	-0.015	0.122	9.430	124.480	66
$CRDT_{de}$	-0.002	0.007	-0.099	0.017	-8.245	104.305	35
Panel B. HM version of the generalized strategy timing model.							
MKT_{ga}	1.814	2.469	-1.863	10.928	1.209	3.835	54
MKT_{de}	1.657	2.510	-6.211	10.928	0.951	4.279	53
$SIZE_{ga}$	0.973	2.132	-4.958	18.406	2.757	18.734	34
$SIZE_{de}$	0.820	2.295	-5.596	18.406	2.138	15.001	43
$CRDT_{ga}$	0.031	0.148	-0.440	1.530	4.569	42.001	28
$CRDT_{de}$	-0.030	0.136	-0.790	0.520	-1.429	10.938	22

Table 6.

Correlation matrices of variables in the selected generalized strategy timing model. Coefficients with high absolute values (over 0.70) are in bold.

Panel A. Correlation coefficients for variables in the TM version of the generalized model.

	MKT-rf	MKT_{ga}	MKT_{de}	*SIZE*	$SIZE_{ga}$	$SIZE_{de}$	*BOND*	*CRDT*	$CRDT_{ga}$	$CRDT_{de}$	$PTFS_{BD}$	$PTFS_{FX}$	$PTFS_{COM}$	*EMRG*
MKT-rf	1.00													
MKT_{ga}	-0.17	1.00												
MKT_{de}	-0.25	**0.74**	1.00											
SIZE	0.08	-0.17	-0.05	1.00										
$SIZE_{ga}$	0.06	0.22	-0.22	0.06	1.00									
$SIZE_{de}$	-0.11	0.17	0.30	0.22	0.63	1.00								
BOND	0.18	-0.09	-0.03	0.20	-0.08	-0.03	1.00							
CRDT	-0.43	0.35	0.26	-0.22	0.10	0.10	-0.53	1.00						
$CRDT_{ga}$	-0.25	0.59	0.66	-0.08	-0.07	0.23	0.01	0.34	1.00					
$CRDT_{de}$	0.24	-0.57	-0.65	0.08	0.07	-0.22	0.00	-0.32	**-0.995**	1.00				
$PTFS_{BD}$	-0.25	0.17	0.19	-0.07	0.03	0.10	-0.24	0.25	0.11	-0.11	1.00			
$PTFS_{FX}$	-0.21	0.16	0.14	0.02	0.03	0.06	-0.15	0.28	0.18	-0.17	0.32	1.00		
$PTFS_{COM}$	-0.17	0.06	0.08	-0.05	-0.02	0.05	-0.10	0.17	0.15	-0.14	0.20	0.34	1.00	
EMRG	**0.73**	-0.24	-0.24	0.24	0.01	-0.08	0.20	-0.49	-0.24	0.24	-0.26	-0.19	-0.16	1.00

Panel B. Correlation coefficients for variables in the HM version of the generalized model.

	MKT-rf	MKT_{ga}	MKT_{de}	*SIZE*	$SIZE_{ga}$	$SIZE_{de}$	*BOND*	*CRDT*	$CRDT_{ga}$	$CRDT_{de}$	$PTFS_{BD}$	$PTFS_{FX}$	$PTFS_{COM}$	*EMRG*
MKT-rf	1.00													
MKT_{ga}	**0.80**	1.00												
MKT_{de}	**0.76**	**0.89**	1.00											
SIZE	0.08	-0.02	0.06	1.00										
$SIZE_{ga}$	0.07	-0.08	-0.10	**0.80**	1.00									
$SIZE_{de}$	0.00	-0.12	-0.15	**0.80**	**0.86**	1.00								
BOND	0.18	0.09	0.13	0.20	0.13	0.14	1.00							
CRDT	-0.43	-0.22	-0.27	-0.22	-0.17	-0.14	-0.53	1.00						
$CRDT_{ga}$	-0.41	-0.12	-0.14	-0.17	-0.13	-0.08	-0.36	**0.75**	1.00					
$CRDT_{de}$	-0.21	-0.20	-0.26	-0.14	-0.11	-0.13	-0.42	**0.70**	0.06	1.00				
$PTFS_{BD}$	-0.25	-0.11	-0.11	-0.07	-0.03	-0.04	-0.24	0.25	0.24	0.11	1.00			
$PTFS_{FX}$	-0.21	-0.08	-0.09	0.02	0.03	0.00	-0.15	0.28	0.29	0.13	0.32	1.00		
$PTFS_{COM}$	-0.17	-0.13	-0.12	-0.05	-0.02	-0.04	-0.10	0.17	0.20	0.03	0.20	0.34	1.00	
EMRG	**0.73**	0.52	0.56	0.24	0.20	0.11	0.20	-0.49	-0.42	-0.28	-0.26	-0.19	-0.16	1.00

Table 7.

Summary statistics for the STX factor and its linear variant.

The *STX* factor is constructed as the monthly variance of the returns of the *MKT*, *SIZE*, and *CRDT* factors. Its linear counterpart, labelled as the STX_L factor, is defined as the monthly average absolute deviation from the mean return between the same three variables. Mathematically, we write:

$$STX = \left((MKT - \bar{f})^2 + (SIZE - \bar{f})^2 + (CRDT - \bar{f})^2 \right) / 3$$
$$STX_L = \left(|MKT - \bar{f}| + |SIZE - \bar{f}| + |CRDT - \bar{f}| \right) / 3$$

Panel A presents the usual summary statistics for the *STX* factor and its linear variant, the STX_L factor. The values of the first four statistics are multiplied by a factor 100 so that the STX_L factor is expressed in percent. As the *STX* factor is the product of two percentages, it is expressed in squared percentages. Any direct comparison between the summary statistics of the two variables is thus meaningless, excepted for the skewness and kurtosis estimates. The last column shows the autocorrelation coefficient obtained from an AR(1) model.

Panel B presents estimated correlation coefficients between the two variants of the *STX* factor and the FH model variables. Values are expressed in decimal format.

Panel A. Summary statistics.

	Mean	Std. Dev.	Min.	Max.	Skew.	Kurtosis	ρ_1
STX	0.0612	0.1074	0.0000	1.1325	5.7848	48.4962	0.3732[a]
STX_L	1.8257	1.2726	0.0166	9.4302	2.0171	10.2364	0.3166[a]

Panel B. Correlations with the FH factors.

	MKT	*SIZE*	*BOND*	*CRDT*	$PTFS_{BD}$	$PTFS_{FX}$	$PTFS_{COM}$	*EMRG*
STX	-0.075	-0.069	-0.105	0.293	0.131	0.121	0.028	-0.149
STX_L	-0.064	-0.044	-0.081	0.231	0.147	0.138	0.037	-0.122

[a] Significant at the 1% confidence level.

Table 8.

Testing for the presence of heteroscedasticity and serial correlation in aggregate-level regression residuals.

The first column shows the different model specifications on which we conduct these test. The TM version of the generalized timing model is labelled as G_{TM}, and its HM counterpart as G_{HM}. For each specification, we regress the returns of the equally-weighted portfolio of funds on the corresponding set of factors. We then compute White's (1980) statistic and compute the *p-value* using the appropriate chi-square cumulative distribution function. Finally, we compute Durbin's (1970) alternative statistic and report the associated *p-value* using the appropriate Student distribution.

Model	White's statistic	*p-value*	Durbin's statistic	*p-value*
STX	69.57	0.0630	1.89	0.060
STX_L	76.98	0.0217	2.05	0.041
G_{TM}	112.74	0.1466	2.38	0.018
G_{HM}	102.90	0.3474	2.61	0.010

Table 9.

Aggregate-level regressions for the generalized strategy timing model.

We regress the returns of the equally-weighted portfolio of funds on the sets of factors corresponding to the two versions of the generalized strategy timing model. The full sample period is used, from Feb. 1994 to Dec. 2017, totalling 287 monthly return observations. Variables follow the same definition as in [table 5]. Panel A presents coefficient estimates for the TM version of the model, while panel B presents coefficient estimates for the HM version of the model. t-statistics are computed using the Newey and West (1987) HAC-robust standard errors, and are shown below their respective coefficients in parentheses. Coefficients in bold are statistically significant at the 5% confidence level. The last column presents the adjusted $\bar{R}^2$ for each regression.

alpha	*MKT*	MKT_{ga}	MKT_{de}	*SIZE*	$SIZE_{ga}$	$SIZE_{de}$	*BOND*	*CRDT*	$CRDT_{ga}$	$CRDT_{de}$	$PTFS_{BD}$	$PTFS_{FX}$	$PTFS_{COM}$	*EMRG*	$\bar{R}^2$
Panel A. TM version of the generalized strategy timing model.															
0.0038	**0.1011**	2.2171	-2.8020	0.0430	-1.5614	8.4855	**-1.5754**	**-1.8359**	-55.244	-36.336	-0.0004	**0.0094**	0.0078	**0.1130**	0.447
(3.29)	(2.90)	(0.31)	(-0.19)	(1.26)	(-0.22)	(0.58)	(-4.10)	(-2.99)	(-0.48)	(-0.26)	(-0.06)	(2.12)	(1.01)	(4.60)	
Panel B. HM version of the generalized strategy timing model.															
0.0012	0.0095	-0.0011	**0.1876**	**-0.1194**	0.0806	**0.2645**	**-1.7029**	-1.7312	-0.8324	0.4347	0.0003	**0.0089**	0.0086	**0.1120**	0.431
(0.64)	(0.20)	(-0.01)	(2.10)	(-2.13)	(0.85)	(2.76)	(-4.47)	(-0.51)	(-0.26)	(0.13)	(0.05)	(1.92)	(1.08)	(4.23)	

Table 10.

Aggregate-level regressions for the *STX* factor and its variants.

We regress the returns of the equally-weighted portfolio of funds on the FH model factors to which we add the *STX* factor. The full sample period is used, from Feb. 1994 to Dec. 2017, totalling 287 monthly return observations. For each regression, we present the coefficient estimates as well as their respective t-statistics in parentheses. All t-statistics are computed using Newey and West (1987) HAC-robust standard errors. We also present the adjusted R^2 of each regression. Each variable is labelled as usual, and coefficients in bold are statistically significant at the 5% confidence level.

Panel A presents regression results for the *STX* and STX_L factors using the original sample of funds. Panel B presents the regression results for the *STX* factor when alternative data filtering rules are employed to construct the sample of funds. The first alternative sample contains 81 funds and is constructed by increasing the minimal return history requirement from 24 months to 36 months. The second alternative sample is obtained by removing suspected funds-of-funds from the original sample (see section 3 for details).

Panel C (on the next page) present regressions results when alternative specifications of the *STX* and STX_L factors are used. In these regressions, the full sample of funds is employed. As a reminder, the original *STX* and STX_L factors are constructed by assuming that managers only try to time the *MKT*, *SIZE*, and *CRDT* strategies (factors). We thus test some variants of this assumption. Alternative specification 1 assumes that the *MKT*, *SIZE*, *CRDT* and *BOND* strategies are subject to strategy timing. Alternative specification 2 assumes that the *MKT*, *SIZE*, *CRDT* and *EMRG* strategies are subject to strategy timing. Alternative specification 3 assumes that the *MKT*, *SIZE*, *CRDT*, *BOND* and *EMRG* strategies are subject to strategy timing.

alpha	*STX*	STX_L	*MKT*	*SIZE*	*BOND*	*CRDT*	$PTFS_{BD}$	$PTFS_{FX}$	$PTFS_{COM}$	*EMRG*	$\bar{R}^2$
Panel A. Base regressions for the STX and STX_L factors.											
0.0046	**2.8147**		**0.0911**	**0.0802**	**-1.7636**	**-2.1971**	0.0013	0.0078	0.0083	**0.1119**	0.421
(4.71)	(3.18)		(2.66)	(1.99)	(-4.71)	(-3.52)	(0.22)	(1.67)	(1.00)	(4.39)	
0.0023		**0.2204**	**0.0931**	**0.0794**	**-1.7442**	**-2.0481**	0.0010	0.0074	0.0081	**0.1111**	0.418
(1.35)		(2.38)	(2.87)	(1.99)	(-4.69)	(-3.37)	(0.17)	(1.62)	(0.98)	(4.34)	
Panel B. Regressions for the STX factor using alternative fund samples.											
Requiring 36 months of minimum return history for each fund (81 funds in the sample).											
0.0045	**2.8704**		**0.0923**	**0.0828**	**-1.7493**	**-2.2146**	0.0015	0.0080	0.0085	**0.1111**	0.422
(4.55)	(3.30)		(2.69)	(2.07)	(-4.67)	(-3.56)	(0.24)	(1.70)	(1.02)	(4.35)	
Excluding self-described funds-of-funds from the sample (84 funds in the sample).											
0.0048	**2.8001**		**0.0887**	**0.0799**	**-1.7869**	**-2.1624**	0.0013	0.0078	0.0086	**0.1141**	0.419
(4.85)	(3.17)		(2.57)	(1.99)	(-4.80)	(-3.48)	(0.21)	(1.65)	(1.03)	(4.43)	

Table 10 (continued).

Aggregate-level regressions for the *STX* factor and its variants.

Panel C present regressions results when alternative specifications of the *STX* and STX_L factors are used. In these regressions, the full sample of funds is employed. As a reminder, the original *STX* and STX_L factors are constructed by assuming that managers only try to time the *MKT*, *SIZE*, and *CRDT* strategies (factors). We thus test some variants of this assumption. Alternative specification 1 assumes that the *MKT*, *SIZE*, *CRDT* and *BOND* strategies are subject to strategy timing. Alternative specification 2 assumes that the *MKT*, *SIZE*, *CRDT* and *EMRG* strategies are subject to strategy timing. Alternative specification 3 assumes that the *MKT*, *SIZE*, *CRDT*, *BOND* and *EMRG* strategies are subject to strategy timing.

alpha	*STX*	STX_L	*MKT*	*SIZE*	*BOND*	*CRDT*	$PTFS_{BD}$	$PTFS_{FX}$	$PTFS_{COM}$	*EMRG*	$\bar{R}^2$
Panel C. Regressions using alternative specifications of the STX factor.											
Alternative specification 1: timed strategies also include BOND.											
0.0044	**3.7733**		**0.0919**	**0.0783**	**-1.7809**	**-2.2442**	0.0011	0.0077	0.0082	**0.1122**	0.424
(4.42)	(3.48)		(2.72)	(2.01)	(-4.72)	(-3.54)	(0.18)	(1.66)	(1.00)	(4.42)	
0.0016		**0.2916**	**0.0941**	**0.0766**	**-1.7488**	**-2.0676**	0.0005	0.0073	0.0082	**0.1111**	0.422
(0.86)		(2.55)	(2.93)	(1.98)	(-4.68)	(-3.37)	(0.09)	(1.60)	(1.00)	(4.34)	
Alternative specification 2: timed strategies also include EMRG.											
0.0043	**1.9179**		**0.0995**	**0.0831**	**-1.8659**	**-2.1678**	0.0009	0.0071	0.0063	**0.1111**	0.417
(4.09)	(2.27)		(3.02)	(1.99)	(-4.79)	(-3.45)	(0.14)	(1.56)	(0.77)	(4.65)	
0.0021		**0.1758**	**0.0989**	**0.0833**	**-1.8357**	**-2.0557**	0.0011	0.0069	0.0065	**0.1077**	0.416
(1.16)		(2.17)	(3.05)	(1.99)	(-4.86)	(-3.41)	(0.18)	(1.48)	(0.79)	(4.43)	
Alternative specification 3: timed strategies also include BOND and EMRG.											
0.0043	**2.0635**		**0.1008**	**0.0821**	**-1.8698**	**-2.1683**	0.0008	0.0072	0.0062	**0.1109**	0.417
(4.12)	(2.26)		(3.07)	(1.98)	(-4.75)	(-3.42)	(0.13)	(1.59)	(0.76)	(4.65)	
0.0021		**0.1893**	**0.1009**	**0.0823**	**-1.8299**	**-2.0530**	0.0009	0.0071	0.0065	**0.1069**	0.416
(1.15)		(2.24)	(3.11)	(1.98)	(-4.81)	(-3.35)	(0.14)	(1.53)	(0.78)	(4.38)	

Table 11.

Controlling for public information and market timing ability in aggregate-level tests of the *STX* model.

We regress the returns of the equally-weighted portfolio of funds on the FH model factors to which we add the *STX* factor, and control for reaction to public information, for the presence of market timing abilities, and for both jointly. The full sample period is used, from Feb. 1994 to Dec. 2017, totalling 287 monthly return observations. For each regression, we present the coefficient estimates as well as their respective t-statistics in parentheses. All t-statistics are computed using Newey and West (1987) HAC-robust standard errors. We also present the adjusted R^2 of each regression. Coefficients in bold are statistically significant at the 5% confidence level.

The original *STX* factor is labelled as usual. In panel A, we control for publicly available information by estimating the following model:

$$r_{t+1} = \alpha + \gamma STX_{t+1}^{r} + \varphi STX_t + \sum_{k=1}^{K} \beta_k f_{k,t+1} + \varepsilon_{t+1}$$

Where STX_{t+1}^{r} corresponds to the residuals of an AR(1) model estimated on the *STX* factor, and represents the unpredictable component of the latter. We label this variable as STX^r in the table below. Meanwhile, the lagged *STX* factor is labelled as STX_{lag}.

In panel B, we control for the possible presence of market timing abilities amongst managers by adding the squared market return as an explanatory variable to the usual *STX* model, which is labelled MKT^2 in the table below.

In panel C, we control for both possibilities jointly, by adding the squared market return to the model defined above.

alpha	*STX*	STX^r	STX_{lag}	*MKT*	MKT^2	*SIZE*	*BOND*	*CRDT*	$PTFS_{BD}$	$PTFS_{FX}$	$PTFS_{COM}$	*EMRG*	$\bar{R}^2$
Panel A. Controlling for publicly available information.													
0.0042		**2.1572**	**2.8041**	**0.0844**		**0.0881**	**-1.5929**	**-2.1831**	0.0044	0.0079	0.0080	**0.1142**	0.430
(4.20)		(2.15)	(3.75)	(2.43)		(2.04)	(-4.50)	(-3.43)	(0.84)	(1.78)	(0.98)	(4.38)	
Panel B. Controlling for market timing ability.													
0.0048	**3.1787**			**0.0896**	-0.2187	**0.0787**	**-1.7320**	**-2.1498**	0.0016	0.0079	0.0083	**0.1118**	0.419
(4.53)	(2.85)			(2.50)	(-0.54)	(2.03)	(-4.43)	(-3.36)	(0.26)	(1.67)	(1.00)	(4.37)	
Panel C. Controlling for publicly available information and market timing ability jointly.													
0.0045		**2.5567**	**2.9814**	**0.0829**	-0.2326	**0.0864**	**-1.5577**	**-2.1322**	0.0048	0.0080	0.0079	**0.1140**	0.428
(4.14)		(2.10)	(3.55)	(2.29)	(-0.59)	(2.08)	(-4.24)	(-3.29)	(0.90)	(1.79)	(0.97)	(4.36)	

Table 12.

Cross-sectional distribution of *t*-statistics for timing coefficients in the generalized strategy timing model.

We estimate both versions of the generalized strategy timing model for each fund in our sample, and then store the *t*-statistics on each timing coefficient. We theoretically use the full sample period for regressions, but in practice the return history of each fund varies. The table reports the percentages of funds which, for each timing coefficient separately, exhibit *t*-statistics which are above or below the indicated cut-off values. The timing variables are labelled as usual. All the *t*-statistics are computed using Newey and West (1987) HAC-robust standard errors.

Coefficient	Percentage of funds					
	$t < -2.326$	$t < -1.960$	$t < -1.645$	$t > 1.645$	$t > 1.960$	$t > 2.326$
Panel A. TM version of the generalized strategy timing model.						
MKT_{ga}	3.33	8.89	13.33	7.78	4.44	4.44
MKT_{de}	4.44	5.56	7.78	13.33	7.78	3.33
$SIZE_{ga}$	4.44	5.56	7.78	11.11	7.78	1.11
$SIZE_{de}$	2.22	6.67	11.11	7.78	5.56	4.44
$CRDT_{ga}$	3.33	7.78	11.11	15.56	13.33	5.56
$CRDT_{de}$	4.44	6.67	11.11	14.44	10.00	4.44
Panel B. HM version of the generalized strategy timing model.						
MKT_{ga}	0.00	2.22	4.44	2.22	1.11	1.11
MKT_{de}	2.22	4.44	5.56	6.67	5.56	2.22
$SIZE_{ga}$	0.00	5.56	8.89	8.89	4.44	1.11
$SIZE_{de}$	0.00	3.33	6.67	10.00	7.78	2.22
$CRDT_{ga}$	7.78	10.00	15.56	7.78	4.44	2.22
$CRDT_{de}$	7.78	10.00	16.67	11.11	8.89	4.44

Table 13.

Bootstrap analysis for the generalized strategy timing model.

For each fund in our sample, we estimate the TM version of the generalized strategy timing model as defined in equation (20), and then run the bootstrap procedure on the cross-section of *t*-statistics. All the *t*-statistics are computed using Newey and West (1987) HAC-robust standard errors. The number of bootstrap iterations is 5,000.

In the table, and for each variable, we report lower and upper quantiles of the cross-sectional distribution of *t*-statistics in the first row, and the corresponding empirical *p-values* obtained from the bootstrap analysis in the second row. Empirical *p-values* are reported in decimal format.

	Lower *t*-statistics quantiles				Upper *t*-statistics quantiles			
Coefficient	1%	3%	5%	10%	10%	5%	3%	1%
MKT_{ga}	-2.67	-2.58	-2.44	-1.81	1.47	1.90	2.71	3.22
p-values	*0.96*	*0.71*	*0.46*	*0.49*	*0.77*	*0.86*	*0.56*	*0.80*
MKT_{de}	-3.43	-2.84	-2.00	-1.44	1.82	2.32	2.59	2.71
p-values	*0.73*	*0.46*	*0.79*	*0.80*	*0.50*	*0.59*	*0.71*	*0.95*
$SIZE_{ga}$	-3.36	-2.85	-1.97	-1.50	1.72	2.27	2.48	2.66
p-values	*0.75*	*0.45*	*0.80*	*0.72*	*0.64*	*0.66*	*0.80*	*0.96*
$SIZE_{de}$	-2.68	-2.55	-2.16	-1.67	1.55	2.00	2.84	3.24
p-values	*0.96*	*0.74*	*0.75*	*0.71*	*0.65*	*0.77*	*0.46*	*0.79*
$CRDT_{ga}$	-3.29	-2.66	-2.27	-1.89	2.02	2.72	3.80	4.52
p-values	*0.84*	*0.70*	*0.66*	*0.37*	*0.29*	*0.32*	*0.14*	*0.47*
$CRDT_{de}$	-3.35	-2.65	-2.17	-1.79	1.99	2.54	3.60	4.85
p-values	*0.81*	*0.69*	*0.73*	*0.48*	*0.30*	*0.44*	*0.18*	*0.38*

Table 14.

Cross-sectional distribution of *t*-statistics for the *STX* factor.

We estimate the FH model along with the *STX* factor for each fund in our sample, and then store the *t*-statistics of the strategy timing coefficient. We theoretically use the full sample period for regressions, but in practice the return history of each fund varies. The table reports the percentages of funds which exhibit *t*-statistics on the *STX* factor which are above or below the indicated cut-off values. The timing variables are labelled as usual. All the *t*-statistics are computed using Newey and West (1987) HAC-robust standard errors.

Panel A reports these results for the full sample, and for both the *STX* and the STX_L factors. Panel B reports the results for the *STX* factor only, when alternative data filtering rules are employed to construct the fund sample. The first alternative sample contains 81 funds and is constructed by increasing the minimal return history requirement from 24 months to 36 months. The second alternative sample is obtained by removing suspected funds-of-funds from the original sample (see section 3 for details).

Coefficient	Percentage of funds					
	$t < -2.326$	$t < -1.960$	$t < -1.645$	$t > 1.645$	$t > 1.960$	$t > 2.326$
Panel A. Cross-section of t-statistics in the full sample.						
STX	8.89	16.67	22.22	15.56	12.22	5.56
STX_L	3.33	10.00	11.11	14.44	7.78	4.44
Panel B. Cross-section of t-statistics in alternative samples.						
Requiring 36 months of minimum return history for each fund (81 funds in the sample).						
STX	7.41	16.05	20.99	16.05	12.35	6.17
Excluding self-described funds-of-funds from the sample (84 funds in the sample).						
STX	5.95	14.29	20.24	15.48	11.90	5.95

Table 15.

Bootstrap analysis for the *STX* factor.

For each fund in our sample, we estimate the FH model along with the *STX* factor, and then run the bootstrap procedure on the cross-section of *t*-statistics. All the *t*-statistics are computed using Newey and West (1987) HAC-robust standard errors. The number of bootstrap iterations is 5,000.

In the table, and for both the *STX* and STX_L factors, we report lower and upper quantiles of the cross-sectional distribution of *t*-statistics in the first row, and the corresponding empirical *p-values* obtained from the bootstrap analysis in the second row. Empirical *p-values* are reported in decimal format.

	Lower *t*-statistics quantiles				Upper *t*-statistics quantiles			
Coefficient	1%	3%	5%	10%	10%	5%	3%	1%
STX	-5.10	-3.35	-3.01	-2.52	2.24	2.85	3.19	4.04
p-values	*0.04*	*0.05*	*0.01*	*0.00*	*0.01*	*0.05*	*0.11*	*0.28*
STX_L	-3.58	-2.58	-2.42	-1.91	1.87	2.29	2.75	3.61
p-values	*0.22*	*0.25*	*0.09*	*0.04*	*0.06*	*0.18*	*0.17*	*0.28*

Table 16.

Bootstrap analysis for the *STX* factor while controlling for publicly available information.

For each fund in our sample, we estimate the FH model along with the *STX* factor while controlling for public information, as in equation (21):

$$r_{t+1} = \alpha + \gamma STX^{r}_{t+1} + \varphi STX_t + \sum_{k=1}^{K} \beta_k f_{k,t+1} + \varepsilon_{t+1}$$

Where STX^r is the unpredictable component of the *STX* factor, and measures true strategy timing ability, while the lagged *STX* variable, labelled STX_{lag} in the table below, represents public information. We then run the bootstrap procedure on the cross-section of *t*-statistics. All the *t*-statistics are computed using Newey and West (1987) HAC-robust standard errors. The number of bootstrap iterations is 5,000.

In the table, and for both the STX^r and STX_{lag} variables, we report lower and upper quantiles of the cross-sectional distribution of *t*-statistics in the first row, and the corresponding empirical *p-values* obtained from the bootstrap analysis in the second row. Empirical *p-values* are reported in decimal format.

	Lower *t*-statistics quantiles				Upper *t*-statistics quantiles			
Coefficient	1%	3%	5%	10%	10%	5%	3%	1%
STX^r	-4.77	-3.50	-2.91	-2.64	1.91	2.43	2.79	3.09
p-values	*0.05*	*0.02*	*0.02*	*0.00*	*0.13*	*0.22*	*0.29*	*0.67*
STX_{lag}	-3.79	-2.79	-1.96	-1.32	2.37	3.13	4.17	5.91
p-values	*0.27*	*0.25*	*0.66*	*0.86*	*0.00*	*0.00*	*0.00*	*0.01*

Table 17.

Bootstrap analysis for the *STX* factor while controlling for publicly available information and market timing jointly.

For each fund in our sample, we estimate the FH model along with the *STX* factor while simultaneously controlling for public information and possible market timing abilities, as follows:

$$r_{t+1} = \alpha + \gamma STX^{r}_{t+1} + \varphi STX_t + \gamma_{MT} MKT^{2}_{t+1} \sum_{k=1}^{K} \beta_k f_{k,t+1} + \varepsilon_{t+1}$$

Where STX^r is the unpredictable component of the *STX* factor, and measures true strategy timing ability, while the lagged *STX* variable, labelled STX_{lag} in the table below, represents public information. MKT^2 is the squared market return, and as such γ_{MT} is the market timing coefficient of the original TM model. We then run the bootstrap procedure on the cross-section of *t*-statistics. All the *t*-statistics are computed using Newey and West (1987) HAC-robust standard errors. The number of bootstrap iterations is 5,000.

In the table, and for the STX^r, STX_{lag}, *and* MKT^2 variables, we report lower and upper quantiles of the cross-sectional distribution of *t*-statistics in the first row, and the corresponding empirical *p-values* obtained from the bootstrap analysis in the second row. Empirical *p-values* are reported in decimal format.

	Lower *t*-statistics quantiles				Upper *t*-statistics quantiles			
Coefficient	1%	3%	5%	10%	10%	5%	3%	1%
STX^r	-3.69	-2.67	-2.22	-1.67	1.09	2.23	2.48	3.66
p-values	*0.33*	*0.29*	*0.31*	*0.28*	*0.98*	*0.26*	*0.39*	*0.26*
STX_{lag}	-2.93	-2.00	-1.82	-1.56	2.06	2.75	3.22	5.42
p-values	*0.68*	*0.87*	*0.77*	*0.46*	*0.02*	*0.02*	*0.03*	*0.02*
MKT^2	-3.92	-2.56	-2.31	-1.55	1.65	1.97	2.70	3.79
p-values	*0.14*	*0.28*	*0.15*	*0.43*	*0.36*	*0.61*	*0.25*	*0.24*

Table 18.

Persistence analysis for reaction to public strategy timing signals.

Starting from March 2003, we estimate for each fund the FH model along with the *STX* factor while simultaneously controlling for public information and possible market timing abilities, as follows:

$$r_{t+1} = \alpha + \gamma STX_{t+1}^{r} + \varphi STX_t + \gamma_{MT} MKT_{t+1}^{2} \sum_{k=1}^{K} \beta_k f_{k,t+1} + \varepsilon_{t+1}$$

The model is estimated on the past 36 months of return history of each available fund. We then rank the funds based on their ability to react to publicly available information, that is, on the magnitude of the estimated φ coefficient. Following this ranking, the funds are assigned to one of ten equally-weighted decile portfolios, ranging from portfolio 1 – the best "timers" – to portfolio 10 – the bottom "timers". These portfolios are held for a period of either 3, 6, 9, or 12 months, and are then rebalanced. This process is repeated until December 2017, and we thus obtain a 177-months-long return series for each decile portfolio. We then estimate the original FH model for each portfolio and report the alphas in the table, along with their respective t-statistics in parentheses. These t-statistics are computed using Newey and West (1987) HAC-robust standard errors as usual. Coefficients in bold denote statistical significance at the 5% level.

Portfolio	3 months	6 months	9 months	12 months
Portfolio 1 (top "timers")	0.518	0.523	0.529	0.487
	(1.85)	(1.94)	(1.97)	(1.66)
Portfolio 2	0.331	0.238	0.305	0.412
	(1.58)	(1.15)	(1.23)	(1.75)
Portfolio 3	0.208	**0.407**	0.217	0.270
	(1.05)	(2.29)	(1.05)	(1.26)
Portfolio 4	**0.286**	0.053	0.177	0.117
	(3.03)	(0.34)	(1.06)	(0.81)
Portfolio 5	**0.700**	**0.862**	**0.569**	**0.724**
	(3.18)	(4.13)	(3.36)	(3.85)
Portfolio 6	0.320	0.166	**0.433**	0.247
	(1.97)	(0.99)	(2.23)	(1.43)
Portfolio 7	0.208	0.264	0.203	0.305
	(1.08)	(1.39)	(1.08)	(1.89)
Portfolio 8	0.261	0.265	0.163	0.168
	(1.68)	(1.89)	(1.04)	(0.97)
Portfolio 9	0.065	0.008	-0.069	0.112
	(0.26)	(0.03)	(-0.31)	(0.45)
Portfolio 10 (bottom "timers")	**0.865**	**0.604**	**1.072**	0.206
	(2.69)	(2.53)	(3.29)	(0.58)

Figure 1.

Aggregate performance of the sample of funds over time.

We simulate the value of the equally-weighted portfolio of funds on the entire sample period, assuming that $100 were invested in the portfolio in January 1994. The portfolio reaches $1,559.24 in value in December 2017.

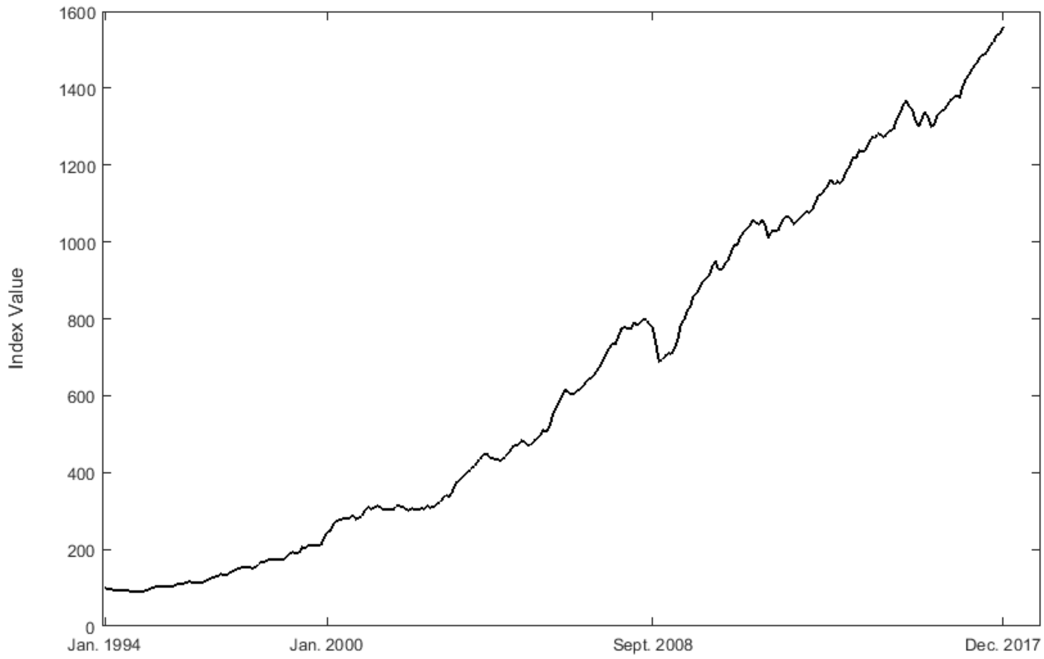

Figure 2.

Multi-strategy hedge funds' returns during the start of the subprime crisis.

The figure plots the monthly returns of the funds in our sample from January 2008 to December 2010. The lines in grey represent monthly returns of individual funds. The solid black line corresponds to the returns of the equally-weighted portfolio of funds. The dotted line indicates the value zero. Returns are not computed in excess of the risk-free rate.

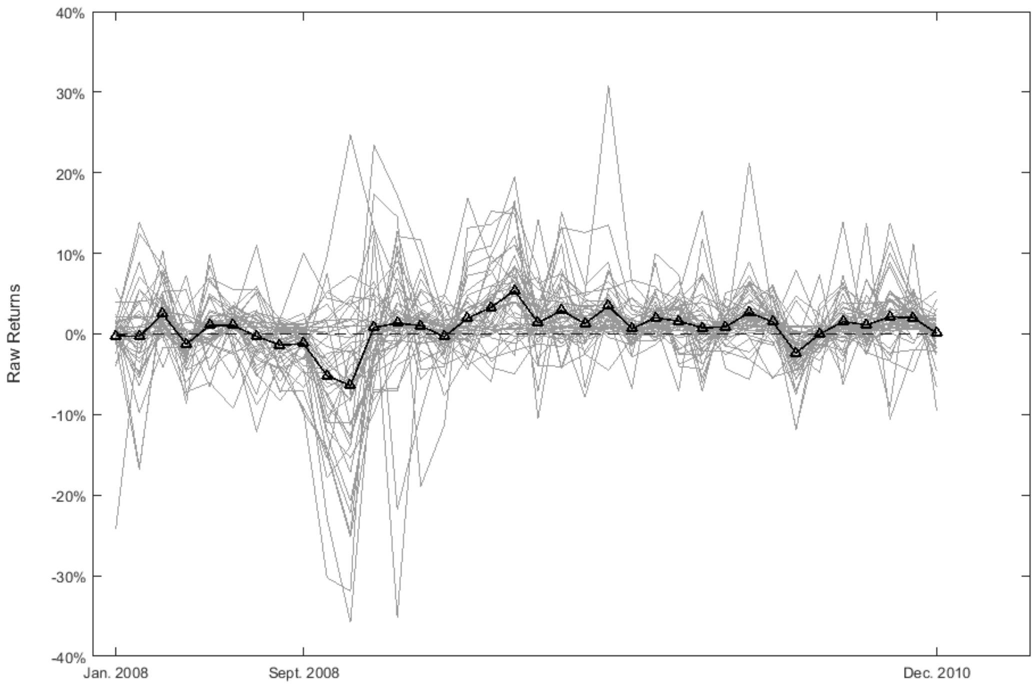

Figure 3.

Kernel density estimates for the empirical cross-section of *STX t*-statistics quantiles.

After generating the empirical distribution of selected quantiles of the *t*-statistics cross-section for the *STX* factor, we use Gaussian kernels to estimate the probability density function of each quantile. For each of these, we report the estimated density function (as the grey area) and the observed quantile value in the original sample (on the dotted line).

As reported in [table 15], the lower quantiles are all significant at the 5% confidence level, while it is obvious that the upper 1% quantile, in particular, has low statistical significance. Note, also, that none of the empirical distributions seem to follow a Student distribution.