Dipartimento di Economia Internazionale, delle Istituzioni e dello Sviluppo (DISEIS)

WORKING PAPER SERIES

WP 1705

December, 2017

Carry trade returns with Support Vector Machines

Emilio Colombo, Gianfranco Forte,

Roberto Rossignoli

DISEIS WP 1705

DISEIS
Dipartimento di Economia Internazionale, delle Istituzioni e dello Sviluppo
Università Cattolica del Sacro Cuore
Via Necchi 5
20123
Milano

Carry trade returns with Support Vector Machines

Emilio Colombo*

UNIVERSITÀ CATTOLICA DEL SACRO CUORE

Gianfranco Forte

UNIVERSITÀ MILANO-BICOCCA

Roberto Rossignoli†

MONEYFARM

December 2017

Abstract

This paper proposes a novel approach to directional forecasts for carry trade strategies based on Support Vector Machines (SVMs), a learning algorithm that delivers extremely promising results. Building on recent findings in the literature on carry trade, we condition the SVM on indicators of uncertainty and risk. We show that this provides a dramatic performance improvement in strategy, particularly during periods of financial distress such as the recent financial crises. Disentangling the measures of risk, we show that conditioning the SVM on measures of liquidity risk rather than on market volatility yields the best performance.

*Corresponding author: Università Cattolica del Sacro Cuore, Largo Gemelli 1, 20123 Milano, Italy, e-mail: emilio.colombo@unicatt.it

†We thank participants to Infiniti, Money, Macro and Finance, ICMAIF and IFABS conferences for helpful comments.

1 Introduction

Uncovered interest parity (UIP) is probably one of the simplest and most intuitive equilibrium conditions in financial markets. For risk neutral investors with rational expectations, the expected exchange rate change has to compensate for the interest differential that may arise between two currencies. Such an equilibrium condition is most likely to hold in the FX market because it is the closest approximation to the notion of market efficiency. Yet the empirical evidence fails to support the UIP. On the contrary, the evidence shows the opposite: high yielding currencies tend to appreciate rather than depreciate, as theory predicts. This is known as the "forward bias puzzle"[1] and has a natural implication: the possibility of realizing excess returns from carry trade, that is, the practice of investing in high yield currencies by going short on low yield ones.

Simple carry trade strategies deliver positive average excess returns for substantially long periods, coupled with Sharpe ratios significantly higher than those measured in other financial markets (such as the US stock market), spurring considerable attention from economists and practitioners alike. During the last decades, a large body of literature has investigated the reasons underlying the UIP puzzle and explanations for the excess of returns from carry trade.[2]

As an equilibrium condition, the carry trade does not require any model. As such, it is a rather naïve strategy and we know that more sophisticated investors would use any information that they may find useful, particularly if it is from some established model. Unfortunately, the literature is of little help in this respect. In fact, in a carry trade strategy, the only unknown is the exchange rate and improving the strategy necessarily implies making a correct guess about the future change in the exchange rate. However, since the seminal work by Meese and Rogoff (1983), researchers struggled to find a model capable of sufficient predictive power for the exchange rate. Recently, there has been some improvements; Della Corte, Sarno, and Tsiakas (2009) show that forward premia and stochastic volatility have predictive ability for exchange rates; Li, Tsiakas, and Wang (2014) reaffirm the out of sample predictive power of economic fundamentals; and Molodtsova and Papell (2009) stress the role of Taylor rule fundamentals. Nonetheless, as shown by the recent survey by Rossi (2013), we are far from having a good predictive model for exchange rates.

In explaining excess returns from carry trade, prior studies follow two main approaches.

[1]The name derives from the fact that the failure of UIP results in the forward rate being a biased predictor of future spot rates.

[2]See Engel (2014) for a recent survey.

On the one hand, the "traditional view" emphasizes the importance of fundamentals in providing useful information for exchange rate forecasting; in particular, Jordà and Taylor (2012) show that conditioning the carry trade strategy on the predictions of a simple fundamental equilibrium exchange rate model yields a better performance with a significant improvement in the Sharpe ratio. On the other hand, building on the fact that carry trade returns are negatively skewed, the "risk view" posits that they are essentially a compensation for currency crashes, reflecting a sort of "Peso problem." Indeed, Brunnermeier, Nagel, and Pedersen (2009), Burnside et al. (2011), and Farhi et al. (2015) show that crash risk has accounted for a high proportion of the carry trade risk premium in advanced countries over the last 20 years.[3] Menkhoff et al. (2012) find that excess returns to carry trades are a compensation for time-varying risk, and in particular, for global foreign exchange volatility risk. More recently, Cenedese, Sarno, and Tsiakas (2014) provide some theoretical underpinnings to the relationship between volatility and carry trade returns using an intertemporal capital asset pricing model and show that conditioning the carry trade strategy on FX risk measures results in a clear performance improvement, even accounting for transaction costs.

Our paper contributes to this line of research providing innovative contributions in several domains: first, we propose a novel approach to directional forecasts for carry trade strategies based on Support Vector Machines (SVM), a learning algorithm that provides extremely promising results and a wide range of possible applications in finance. To our knowledge, this is the first application of these tools to exchange rate modelling and carry trade. Second, we condition the SVM model on indicators of uncertainty and risk. We show that this provides a dramatic performance improvement in the strategy, particularly during periods of financial distress, such as the recent financial crises. This provides clear support for the view that considers excess returns from carry trade as compensation for risk. Third, we disentangle different measures of risk, showing that measures of liquidity risk rather than measures of market volatility yield the best performance.

The remainder of the paper is structured as follows: section 2 illustrates the methods and data, section 3 presents the results, and section 4 concludes.

[3] Using different techniques, Jurek (2014) downsizes the importance of crash risk. See also Smales and Kininmonth (2016), who stress the role of investorsâĂŹ fear.

2 Methods and data

There are two general approaches to constructing the carry trade. The traditional approach defines the carry strategy in terms of interest rate differentials, an approach followed by Brunnermeier, Nagel, and Pedersen (2009) and Jordà and Taylor (2012), among others. An alternative approach casts the carry trade in terms of forward currency contracts; see, for example, Burnside et al. (2011), Cenedese, Sarno, and Tsiakas (2014), and Bakshi and Panayotov (2013). Clearly, the two approaches are equivalent if the Covered Interest Parity holds, which is an assumption well supported by empirical evidence.

We follow the latter approach, which has the major advantage of allowing researchers to incorporate transaction costs simply, which are crucial in determining the real return on carry trade strategies.[4]

The return on the carry trade strategy can be briefly illustrated as follows:

$$z_{t+1} = \begin{cases} \left(\frac{F^b_{t,t+1}}{S^a_{t+1}}\right)^{\gamma_t} - 1 & \text{if } \gamma_t > 0 \\ \left(\frac{F^a_{t,t+1}}{S^b_{t+1}}\right)^{\gamma_t} - 1 & \text{if } \gamma_t < 0 \end{cases} \tag{1}$$

where[5]:

$$\gamma_t = \begin{cases} +1 & \text{if } F_{t,t+1} > S_t \\ -1 & \text{if } F_{t,t+1} < S_t \end{cases} \tag{2}$$

Clearly, given that we consider several currency pairs among the countries this study examines, there is an issue of identifying the “best” strategy. To avoid the problem of data snooping, we concentrate only on two portfolios: one is an equally-weighted portfolio consisting of placing a uniform bet of size $1/N$ where N is the number of currency pairs. The second is a dynamically rebalanced portfolio, in which each periodâĂŹs currencies are ranked by their interest rate differential, with the investment in only the first three currencies. Therefore we go long (with equal weights) on the top three interest rate differentials and short (with equal weights) on the bottom three interest rate differentials. Given that the US dollar is the reference currency against which returns are calculated whenever the US dollar is among, for example, the lowest-yielding currencies in a given period, we do not short any other currency against it and adjust the weights accordingly, in order to main-

[4]As Lyons (2001) shows, bid ask spreads quoted by standard data providers are approximately twice the size of inter-dealer spreads. Therefore, our transaction cost estimates are very conservative.

[5]In the formulas, subscripts a and b represent bid and ask, respectively.

tain a zero-cost strategy. We determine the direction of the forward (long or short USD) following equations 1 and 2 in the text (naïve Carry Trade) or using the SVM as explained below.

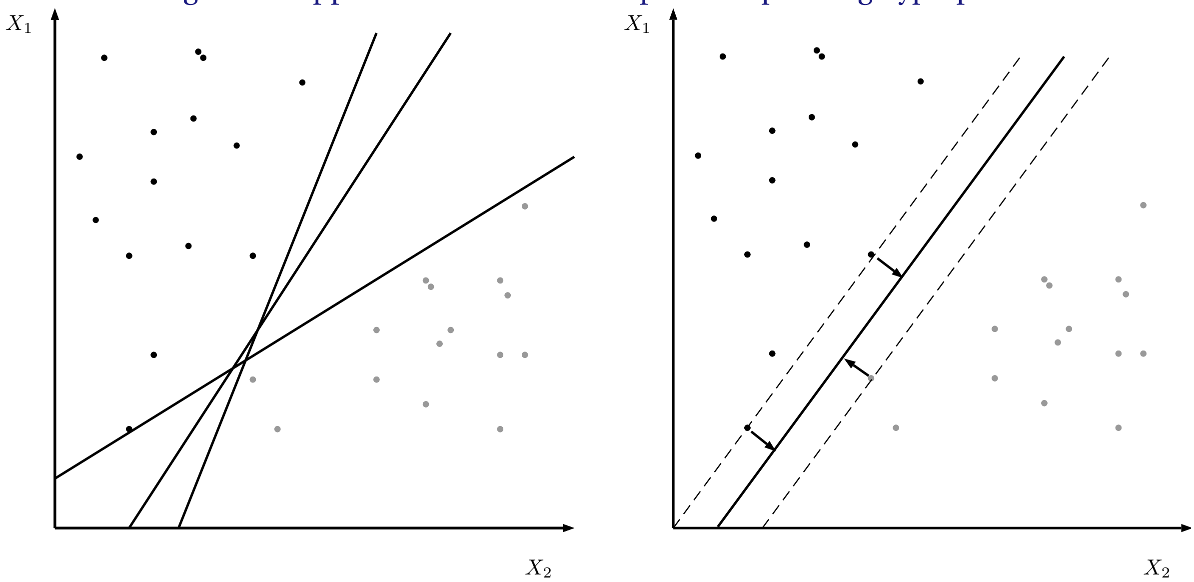
2.1 Carry trade and Support Vector Machines

As stated in the introduction, the standard carry trade strategy is rather simple and naïve, and could be ideally improved by adding information that helps predict future exchange rate changes. We do so by conditioning the long-short trade choices of a carry trade strategy on the prediction of a model. Instead of using the direction of the forecast of a standard regression model, we use an SVM in which the input variables are indicators of market uncertainty. Thus, conditional on observing indicators of market uncertainty at time *t*, the SVM yields a prediction about whether to continue the carry trade strategy or to invert it at time *t+1*. SVM is currently one of the most popular machine learning algorithms and has been applied in numerous fields successfully, and recently to financial market forecasts.[6] The SVM is a binary classification algorithm that classifies observations with certain features into two classes. This is particularly interesting for purpose carry trades, since what really matters for the investor or the trader is the ability to predict the direction of the trade, not its magnitude. In order to gain excess returns from the carry trade, one in fact has to guess the direction of the change in exchange rates correctly rather than accurately forecast the exchange rate itself. In this respect, binary classification systems may work better than standard econometric tools such as Probit or Logit regressions.

Intuitively, the Support Vector is an algorithm that constructs the hyperplane that maximizes the distance between two classes of observations (positive and negative returns, in our case). When the two classes are clearly separable, the SVM is analogous to a linear optimization problem that we can solve with standard tools (i.e. Lagrange multiplier). Figure 1 helps to refine the intuition. In the figure, we represent two classes of observations (black and grey dots), each of which is defined by two variables (X_1, X_2); there are infinite possible linear separating hyperplanes we can use as classifiers (the figure in the left panel shows three). SVMs select the optimal separating hyperplane in the following way: first, the perpendicular distance between any observation and the given hyperplane is computed; second, it is identified the smallest distance from the observation to the hyperplane, which defines the margin, that is, the maximal width of the slab parallel to the hyperplane that has no interior data points. The optimal separating hyperplane maximizes the margin. The

[6] See, for example, Huerta, Corbacho, and Elkan (2013) and Papadimitriou, Gogas, and Stathakis (2014).

Figure 1: Support Vector Machine: Optimal separating hyperplane



right panel of Figure 1 has three observations (two black and one grey) with minimum distances from the hyperplane. These observations are called support vectors, and they identify the margin (the distance between the two dashed lines). The intuition underlying the margin maximization is that a large margin on the training data is expected to deliver a good classification of the test data.

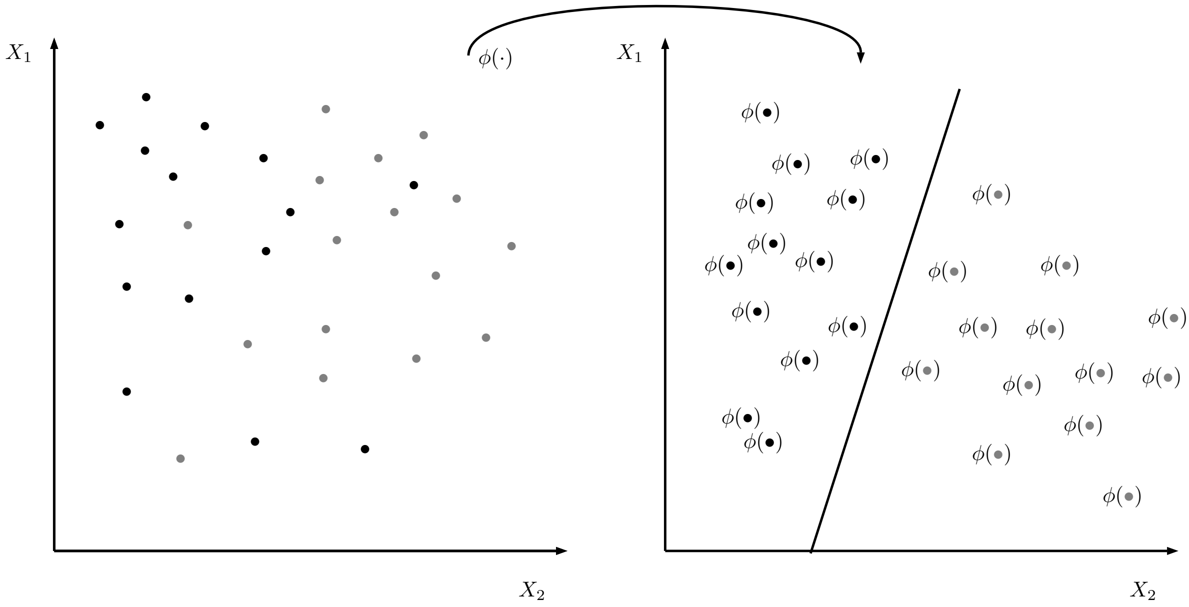
A classification based on the optimal separating hyperplane could be extremely efficient, but with a drawback: it could be too sensitive to individual observations, as Figure 1 illustrates. A change in the support vectors would imply a potentially large change in the position of the maximal margin hyperplane. SVMs are flexible instruments that solve this problem by introducing a form of soft margin classification; in other words, they allow the misclassification of few training observations in order to reach a better classification for most of the training observations.

The example in Figure 1 is rather simple; however, there are several cases where a linear separation is not possible, for instance, when the relationship between the predictors and the outcome is non-linear. This is indeed the case with exchange rates, where a growing body of literature stresses the presence of non-linearities in their dynamic adjustments (Sarno, Valente, and Leon, 2006). The left panel of Figure 2 shows one such example, where a linear classifier would perform rather poorly.

In principle, it would be possible to enlarge the feature space with non-linear functions of the predictors. However, this would result in a large number of features and an unmanageable computational cost.

Fortunately, a useful result from Cortes and Vapnik (1995) shows that it is possible to

Figure 2: Support Vector Machines: The kernel principle



project the dataset through a kernel function into a higher dimensional space (i.e., feature space), where the dataset is linearly separable (see the right panel of the figure).

More formally, let $\mathbf{x}$ be a p element vector of variables (in our case, the returns from carry trade and the conditioning factors such as uncertainty) and let n be the number of training observations. The SVM is a classification function (James et al., 2013):

$$f(x) = \sum_{i=1}^{n} \alpha_i y_i K(x, x_i) + \beta_0 \tag{3}$$

where $\mathbf{x}_i$ defines the vector of variables in the training set; y_i is a classifier that takes values of $+1$ or -1, formally assigning each observation in one of the two groups; β_0 is a constant that shifts the SVM output; α_i are parameters that are non-zero only for the support vectors and depend on the tuning parameter for the soft margin classification. Both β_0 and α_i are chosen optimally within the training period. Finally, we chose the kernel function K following the literature, selecting the Radial Kernel:

$$K(x, x_i) = \exp\left(-\gamma \sum^{p} (x - x_i)^2\right) \tag{4}$$

where γ is a parameter chosen to optimize the sample fit of the model.

In practice, when a test observation x is far from a training observation x_i, it implies that over the p dimensions, the Euclidean distance $(x - x_i)$ is large, so $\sum^{p}(x - x_i)^2$ is large and $K(x, x_i)$ is extremely small. Therefore, the training observation x_i plays no role in predicting the class for the test observation x. Using the kernel function is computationally very effective since the kernel has to be computed for each possible pair x, x_i, that is, $n(n-1)/2$.[7]

SVMs have several interesting features that make them potentially extremely useful for providing a directional forecast of the exchange rate. The two main advantages are the following.

- By construction, the SVM is optimized to discriminate around the decision margin, while it attaches no weight to data that are easily classifiable. This is the main difference from a regression-based approach: the latter weights all observations, and not just those close to the decision margin, making a regression-based approach less effective in binary classification problems. In principle, a logistic regression is also efficient as a classifier, but is limited by the fact that it estimates a *linear* decision boundary.

- SVM deals with non-linearity quite naturally through the kernel function without imposing a particular functional form, which could be valuable in cases such as our case, since it has no well-established theory that provides clearly testable implications.[8] This could help overcome a general problem with non-linear models that are known to perform well in-sample but fail in out-of-sample forecasting (Teräsvirta, 2006).[9]

As with any machine learning algorithm, the SVM needs a training period. We select the previous 5 years as the training period. We then apply the algorithm with a rolling window of 5 years. The rolling window approach is widespread in exchange rate modelling since the seminal work of Meese and Rogoff (1983), and has the benefit of reducing the problem of parameter instability over time (a known issue in this field), at the cost of not considering possible efficiency gains from an increasing sample size over time.

2.2 Measuring performance

In comparing the results of a SVM strategy with the standard carry trade, we use a number of measures. In a mean-variance setting, we provide, alongside with the average return, standard deviation, skewness, and the Sharpe ratio. It is well known that standard tests

[7] The technical appendix provides a more formal explanation of the maximization problem underlying the SVM and illustrates differences from and similarities to a logistic regression.

[8] Indeed, Jordà and Taylor (2012) show that their augmented carry trade strategy improves the most when they use a non-linear model.

[9] In theory, kernel functions can also be applied to logistic regressions, but so far, no application has been developed for this purpose.

comparing the Sharpe ratios of different investment strategies fail when the tails are heavier than the normal distribution or display time series correlation. We therefore use the Ledoit and Wolf (2008) test, which is robust to non-normality and serial correlation in returns. Given the skewness in carry trade returns, we compute additional measures common in finance and particularly suitable for non-normal returns: the Omega ratio, Sortino Index, and upside potential. The Omega ratio measures the probability-weighted ratio of gains versus losses for some threshold return target, and it employs all information contained within the distribution of returns, not assuming normality in the distribution of returns.

$$\Omega = \frac{\frac{1}{T}\sum_{t=1}^{T} i^{+}(r_t - R_{min})}{\frac{1}{T}\sum_{t=1}^{T} i^{-}(r_t - R_{min})} \tag{5}$$

where R_{min} is the minimum acceptable return (0 in our case), T is the total number of periods, i^{+} and i^{-} are two indicator functions constructed as follows: $i^{+} = 1$ if $r_t >= R_{min}$ and $i^{+} = 0$ if $r_t < R_{min}$; $i^{-} = 1$ if $r_t <= R_{min}$, and $i^{-} = 0$ if $r_t > R_{min}$

The Sortino Index is analogous to the Sharpe ratio with the difference that in computing the standard deviation of excess returns, it considers only negative returns.

$$SI = \frac{R - T}{DR}, \tag{6}$$

where R is the average return and $DR = \left(\frac{1}{T}\sum_{t=1}^{T} i^{-}(r_t - R_{min})^2\right)^{1/2}$ is the downside risk.

When return distributions are symmetrical and the target return is close to the median of the distribution, the Sortino and Sharpe ratios provide similar results. However, in the presence of skewness of returns, there can be substantial differences. Generally, the larger the Sortino index is, the lower is the risk of large losses.

We also include the upside potential, which measures the upside potential relative to the downside risk.

$$UP = \frac{\frac{1}{T}\sum_{t=1}^{T} i^{+}(r_t - R_{min})}{\sqrt{\frac{1}{T}\sum_{t=1}^{T} i^{-}(r_t - R_{min})^2}} \tag{7}$$

Finally, following Della Corte and Tsiakas (2012), Fleming, Kirby, and Ostdiek (2001), and Thornton and Valente (2012), we compute a utility-based measure of performance. In particular, we calculate the performance fee, that is, the maximum fee that an investor is willing to pay to switch from the standard carry trade strategy to the improved SVM strategy. We derive the fee (F) by equating the average utility of the portfolios based on the two investment strategies. Formally, F solves the following equation:[10]

$$\sum_{t=0}^{T-1}\left\{\left(R_{t+1}^{CTSVM}-F\right)-\frac{\lambda}{2(1+\lambda)}\left(R_{t+1}^{CTSVM}-F\right)^2\right\}=\sum_{t=0}^{T-1}\left\{R_{t+1}^{CT}-\frac{\lambda}{2(1+\lambda)}\left(R_{t+1}^{CT}\right)^2\right\} \tag{8}$$

Clearly, the measure above depends on the shape of the utility function, particularly the relative risk aversion parameter λ.[11]

2.3 Data

We extracted exchange rate data from Factset, which provides bid, ask, and mid quotes for spot and forward contracts on a daily basis. We conducted our analysis using both monthly and weekly forward rates and our sample period starts in October 1997 and ends in June 2017.[12] In terms of currency selection, in order to compare our results with those reported in most of the literature, we consider only the most liquid currencies, that is, the so called G10 currencies. Specifically, we selected the Australian Dollar (AUD), Canadian Dollar (CAD), Swiss Franc (CHF), Euro (EUR), UK Pound (GBP), Japanese Yen (JPY), Norwegian Krona (NOK), New Zealand Dollar (NZD), Swedish Krona (SEK), and US Dollar (USD). In computing the returns to carry trade, we must define a reference currency in terms of which they are calculated. As in most studies, we select the US Dollar, so we express each exchange rate as a unit of currency per USD.

In calculating the carry trade return, we build two kinds of strategies. The first is based on weekly rolling of one-week forward, so the forecast is for the one-week-ahead spot exchange rate. The second aims to forecast the one-month ahead movements in exchange rates using one-month forward.

We use different variables as measures of uncertainty in the SVM. First, we consider the well-known VIX index, that is, the measure of implied volatility in the S&P 500 index

[10] Following Della Corte, Sarno, and Tsiakas (2009), we can derive the equation below considering quadratic utility as good second-order approximation of investorsâĂŹ true utility functions. Thus $U(W_{t+1})=W_tR_{t+1}^p-\frac{1}{2}\alpha W_t^2(R_{t+1}^p)^2$, where W is the investors' wealth, α is the investorsâĂŹ absolute risk aversion, and R^p is the return on the investorsâĂŹ portfolios. Assuming constant αW_t, defining the coefficient of relative risk aversion $\lambda=\alpha W_t/(1-\alpha W_t)$ and defining W_0 as initial wealth, we can derive the average utility as a consistent estimate of expected utility:

$$\bar{U}(\cdot)=W_0\left(\sum_{t=0}^{T-1}R_{t+1}^p-\frac{\lambda}{2(1+\lambda)}(R_{t+1}^p)^2\right),$$

from which the equation in the text follows.

[11] In our calculations, we assumed a value of $\lambda=6$.

[12] More specifically, the sample period for the monthly forward contracts ranges from 10-1997 to 06-2017, while the sample for the weekly forward contract ranges from 10-1999 to 06-2017.

options. The VIX is probably the most popular measure of expected volatility in the stock market. Second, we consider the TED spread, the difference between the LIBOR and the US short-term government bond rate (T-bills); it is the most popular measure of credit risk. Third, we consider the Chicago Fed National Financial Conditions Index (NFCI), which measures U.S. financial conditions in money, debt, and equity markets, and the traditional and "shadow"? banking systems. Fourth, following Christiansen, Ranaldo, and Söderlind (2011), we construct a measure of the FX Implied Volatility using the implied volatility of the at-the-money 1-month options on the CAD, CHF, EUR, JPY, and GBP against the USD.[13] We obtained the VIX, TED, and NFCI indexes from the Federal Reserve Bank of St. Louis database and the data for the FX implied volatility measure from Bloomberg. The frequency and time availability are the same as in the corresponding carry trade strategy (weekly and monthly).

3 Results

The excess returns from carry trade are a direct consequence of the failure of uncovered interest parity. Indeed also on our data running standard UIP tests we would get a generalized rejection of it.[14]

3.1 Carry trade and uncertainty

In the previous sections we stressed before that one possible explanation for the excess returns of carry trade is that they are a compensation for "crash risk," where traders are exposed to larger losses as the exchange rate increasingly deviates from fundamentals. In fact, Brunnermeier, Nagel, and Pedersen (2009) and Clarida, Davis, and Pedersen (2009) document that carry trades are characterized by negative skewness. By linking funding and market liquidity, Brunnermeier and Pedersen (2009) provide a theoretical rationalization for this. On the one hand, tradersâĂŹ ability to provide market liquidity depends on their availability of funding; on the other hand, traders' funding depends on assets' market liquidity. In this context in a period of high volatility, liquidity can dry up, generating liquidity spirals that can cause large losses from the carry trade. More importantly, the relationship between volatility and market return is asymmetric, determining a negatively skewed distribution of asset returns. In periods of high volatility and market uncertainty, large negative

[13] We use the simple average across these measures; using the 1st principal component instead of the simple average would not change our results.

[14] Results are available upon request.

shocks generate losses and make traders' funding constraints binding, forcing them to sell assets, which in turn worsen volatility and the funding problem. On the contrary, funding constraints are not binding in the case of positive shocks, and the amplification mechanism would not occur.

In this setting, the relationship between carry trades returns and market uncertainty differs from that traditionally posited by the standard intertemporal capital asset pricing model - ICAPM - (Merton, 1973): $E_t[r_{t+1}] = \alpha + \beta Var_t[r_{t+1}]$.

According to the ICAPM, the condition above should hold for the full conditional distribution of asset returns. On the contrary, with a skewed distribution of returns, the coefficient β varies along the distribution of expected returns, possibly assuming different signs at its extremes.

Following Cenedese, Sarno, and Tsiakas (2014), we document this by estimating predictive regressions where the expected returns on a carry trade at different quantiles are regressed on measures of volatility. Our approach differs from that of Cenedese, Sarno, and Tsiakas (2014) in one important aspect: they use a measure of ex-post realized volatility in the exchange rate, while we use forward looking measures of market uncertainty. We believe that this is the most correct approach to frame the intertemporal capital asset relationship, which holds ex-ante in expected terms.

Specifically, we estimate the following regression:

$$Q_{r_{t+1}}(k|Unc_t) = \alpha(k) + \beta(k)Unc_t + \epsilon_{t+1}, \tag{9}$$

where $Q_{r_{t+1}}(k|\cdot)$ is the k-th quantile of the distribution of returns from carry trade and Unc are the following measures of uncertainty/risk: Ted spread, VIX, NFCI, and FX-implied volatility, as described in section 2.3.

Figure 3 reports values of standardised β coefficients from equation (9).[15] Following the argument by Brunnermeier and Pedersen (2009), we expect a negative relationship between market uncertainty and returns from carry trade at lower quantiles, and positive at higher quantiles of the return distribution. The figure fully confirms our prior for all measures of uncertainty we use, for both monthly and weekly data. In addition, Figure 4 reports the R2 of the regressions at different quantiles, revealing that at lower quantiles, not only does market uncertainty have a negative effect on carry trade returns, but its explanatory power is particularly high.

[15] All coefficients reported in the figure are highly significant. Results available upon request.

3.2 SVM strategies

If market uncertainty is so important in predicting large negative returns, we could use this information to improve carry trade strategies. We do so by implementing the SVM tool described in section 2.1; we therefore augment the carry trade strategies on the prediction of an SVM algorithm conditioned on the set of variables capturing market uncertainty. Our strategy is extremely simple: for any pair of currencies, given a direction of trade identified by the standard carry trade, we confirm it if the prediction of the SVM suggests it; otherwise, we invert it. We therefore use the SVM for a directional forecast, for which it is more efficient.

In order to compare our results with those in the literature, we implement the investment strategy for both monthly and weekly forward rates and compute two portfolios. The first is an equally weighted portfolio that places a uniform bet on every currency pair in the sample (Jordà and Taylor, 2012; Burnside, Eichenbaum, and Rebelo, 2008); the second is a dynamic portfolio that invests in the 3 currency pairs that display the largest forward-spot differential in each period (Bakshi and Panayotov, 2013). In every case, we compare the results of the SVM corrected model with the standard classic carry trade strategy.

Table 1 presents the results. A classic carry trade strategy conducted over the entire period 2002-2017[16] yields an average yearly return of 2.67% (considering transaction costs), in line with the results reported in the literature. We confirm that carry trade returns are also characterized by negative skewness, stressing the importance of considering measures that account for more than just the first two moments.

Transaction costs can make a big difference in terms of average returns;[17] in fact, they explain a large part of the difference between the strategies conducted in 1 week and 1 month forward contracts (the average return drops from 2.67 to -0.04%). This is because with weekly forward contracts, transaction frequency is higher, and hence so are the costs.

Turning to the main innovation of the paper, the use of the SVM conditioned on uncertainty and risk, which improves the returns from carry trade spectacularly. Considering the equally weighted portfolio, the average return jumps from 2.67 to 9.68% with monthly data and from -0.04 to 11.17% with weekly data. Additionally, the Sharpe ratio improves significantly and the negative skewness disappears; Omega, Sortino, and upside potential all increase significantly. In particular, the strong increase in the Sortino Index suggests that the SVM-based strategy minimizes the risk of heavy losses.

[16] As we stressed in section 2, we use the period before 2002 in the sample to train the SVM algorithm.

[17] In our sample, accounting for transaction costs reduces the average annualized return by approximately 1%.

Finally, turning to utility-based performance measures, an investor would be willing to pay a fee of up to 9% to switch from the standard carry trade strategy to the SVM-based strategy for the equally weighted portfolio with monthly forward rates. The fact that the fee is close to or even higher than the returns on the SVM model itself indicates the attractiveness of this strategy: the investor is willing to pay a lot, not only to capture the higher return, but also as a reward for the lower risk and a lesser exposure to negative outcomes.

To better understand the origin of this performance improvement, Figure 5 splits the results between cases where the SVM and the standard carry trade model agree (thus implying the same strategy) and where they disagree (implying a different strategy). In the first case, the results are identical by construction. However, when the models disagree, the differences are strong. The standard carry trade model yields negative returns, while the SVM model does the opposite. The same applies to the skewness, negative in case of agreement, which turns strongly positive for the SVM in case of disagreement. It is evident that accounting for market uncertainty allows investors to hedge against negative draw-downs.

This is clearly shown in Figure 6, which plots the distribution of returns from the SVM model and the standard carry trade for periods where the latter are positive and negative. Using the SVM model allows traders to gain in negative periods rather than improve performance in positive ones.

Figures 8 - 9 show the cumulative returns from carry trade strategies over the sample period. While the risk-based SVM does not improve the performance prior to the financial crisis dramatically (where the standard carry trade is highly profitable), it makes a huge difference during and after the crisis.

3.3 Inspecting the mechanism

The results of the previous section show that using the SVM algorithm conditioned on measures of market uncertainty improves the returns from carry trade significantly and in particular, allows investors to hedge against large negative drawdowns. The extremely positive result of the SVM model, however, comes at a price. As with any other machine learning algorithm, the SVM is a bit of a black box since it does not provide the information economists are accustomed to, such as the standard output from a regression (i.e. coefficients, degree of significance, goodness of fit, etc.). In this section, we provide additional analysis that sheds more light on the factors that drive the results.

3.3.1 Splitting time periods

With a dynamic perspective, Figure 7 plots the returns from the standard carry trade and from the SVM model during the sample period. It is striking that during periods of financial turbulence, such as the global financial crisis, the SVM model can hedge against large drawdowns of carry trade considerably.

In order to analyse these issues more thoroughly we have split the time horizon in three periods: one (2003-2007) which is the period leading up to the financial crisis with rising interest rates and emerging market boom, the period of the great financial crisis itself (2008-2012) which includes also the European debt crises, and the more recent period (2013-2017) characterised by slow recovery with globally flat rates. The Table 2 shows that indeed, the major gains from the SVM model are obtained during and after the financial crisis, when, in the presence of low or negative returns from carry trade, correcting the strategy for uncertainty yields consistently higher returns coupled with substantially higher Sharpe and Omega ratios. This seems the period in which the learning algorithm uses information about market uncertainty more efficiently.

3.3.2 Disentangling uncertainty measures

As stated above, the SVM model uses information from four indicators of market uncertainty to improve the carry trade strategy. As a matter of fact, the SVM is more efficient in using disaggregated rather than aggregated information; thus, its performance worsens slightly if we collapse the four indicators by, for example, a principal component or factor analysis.

However, the measures we considered for market uncertainty capture somewhat different aspects and it would be interesting to understand what effect the SVM model is really capturing. To shed light on this issue, we split the measures of market uncertainty in two groups: one comprising the VIX and the FX volatility index, and another comprising the TED and the NFCI index. While the former captures market volatility, the latter captures uncertainty linked to liquidity issues, and as such, are more in line with the Brunnermeier, Nagel, and Pedersen (2009) argument. Table 2 shows that both groups provide useful information to the SVM model: using only one group of indicators instead of both reduces the average returns significantly. However, it seems that conditioning the SVM on liquidity measures yields better performance than conditioning on volatility measures only, particularly during the financial crises, where the differences in the Sharpe ratios and Sortino index are extremely relevant.

In addition, we performed predictive regressions where the realized returns of SVM carry trade are regressed on lagged uncertainty measures. This analysis is very stylized, as we cannot interpret the coefficients as proper factor loadings since the SVM can by construction use a non-linear approach, whereas here we are using a linear regression model. Nevertheless, despite its limitations, this approach can be fruitful in shedding light on the relative role of the different indicators of market uncertainty. Table 3 shows that in the whole sample both Forex volatility and NFCI are significant. Comparing the indicators of volatility and liquidity (2nd and 3rd column), the former seem to be less relevant than the latter, both in terms of significance and the explained variance of the returns. The analysis by time periods reveals that before the global financial crisis, indicators of market volatility are predictive of carry trade returns, while during the crisis, they become insignificant while market liquidity becomes significant (NFCI). The last column of the table we reports the results of predictive regressions of the *difference* between the returns from the SVM carry trade and from the standard carry trade on lagged uncertainty measures. The analysis shows that the outperformance of the SVM is explained mainly by liquidity variables confirming the results of Table 2.

4 Conclusions

This paper has contributed to the literature on carry trade by proposing a novel approach to provide directional forecasts for carry trade strategies. This approach is based on SVMs, a binary classification mechanism with potential for application in several fields in economic forecasting with extremely promising results. This is particularly interesting for carry trades since, where the investorsâĂŹ or tradersâĂŹ ability to predict the direction of the trade rather than its magnitude matters most. We condition the SVM model on indicators of uncertainty and risk. We show that this provides a dramatic performance improvement in the strategy, particularly during periods of financial distress, such as the recent financial crises. Finally, we show that the relative contribution of liquidity variables is more relevant than that of the volatility variables.

Albeit at its infancy, the use of machine learning algorithms can help researchers understand several aspects of international finance, where standard tools perform rather poorly. Further research on these issues is certainly needed.

References

Bakshi, Gurdip and George Panayotov. 2013. "Predictability of currency carry trades and asset pricing implications." *Journal of Financial Economics* 110 (1):139–163.

Brunnermeier, Markus K., Stefan Nagel, and Lasse H. Pedersen. 2009. "Carry Trades and Currency Crashes." In *NBER Macroeconomics Annual 2008, Volume 23*, NBER Chapters. National Bureau of Economic Research, Inc, 313–347.

Brunnermeier, Markus K. and Lasse Heje Pedersen. 2009. "Market Liquidity and Funding Liquidity." *Review of Financial Studies* 22 (6):2201–2238.

Burnside, Craig, Martin Eichenbaum, Isaac Kleshchelski, and Sergio Rebelo. 2011. "Do Peso Problems Explain the Returns to the Carry Trade?" *Review of Financial Studies* 24 (3):853–891.

Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo. 2008. "Carry Trade: The Gains of Diversification." *Journal of the European Economic Association* 6 (2-3):581–588.

Cenedese, Gino, Lucio Sarno, and Ilias Tsiakas. 2014. "Foreign exchange risk and the predictability of carry trade returns." *Journal of Banking & Finance* 42 (C):302–313.

Christiansen, Charlotte, Angelo Ranaldo, and Paul Söderlind. 2011. "The Time-Varying Systematic Risk of Carry Trade Strategies." *Journal of Financial and Quantitative Analysis* 46 (04):1107–1125.

Clarida, Richard, Josh Davis, and Niels Pedersen. 2009. "Currency carry trade regimes: Beyond the Fama regression." *Journal of International Money and Finance* 28 (8):1375–1389.

Cortes, Corinna and Vladimir Vapnik. 1995. "Support-Vector Networks." *Machine Learning* 20 (3):273–297.

Della Corte, Pasquale, Lucio Sarno, and Ilias Tsiakas. 2009. "An Economic Evaluation of Empirical Exchange Rate Models." *Review of Financial Studies* 22 (9):3491–3530.

Della Corte, Pasquale and Ilias Tsiakas. 2012. "Statistical and Economic Methods for Evaluating Exchange Rate Predictability." In *Handbook of Exchange Rates*, edited by J. James, L. Sarno, and I.W. Marsh. Hoboken, NJ: Wiley.

Engel, Charles. 2014. "Exchange Rates and Interest Parity." In *Handbook of International Economics*, vol. 4, edited by Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff. Elsevier, North Holland, 435–522.

Farhi, Emmanuel, Samuel Paul Fraiberger, Xavier Gabaix, Romain Ranciere, and Adrien Verdelhan. 2015. "Crash Risk in Currency Markets." Mimeo, Harvard University.

Fleming, Jeff, Chris Kirby, and Barbara Ostdiek. 2001. "The Economic Value of Volatility Timing." *The Journal of Finance* 56 (1):329–352.

Huerta, Ramon, Fernando Corbacho, and Charles Elkan. 2013. "Nonlinear support vector machines can systematically identify stocks with high and low future returns." *Algorithmic Finance* 2:45–58.

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. *An Introduction to Statistical Learning*. Springer Verlag.

Jordà, Òscar and Alan M. Taylor. 2012. "The carry trade and fundamentals: Nothing to fear but FEER itself." *Journal of International Economics* 88 (1):74–90.

Jurek, Jakub W. 2014. "Crash-neutral currency carry trades." *Journal of Financial Economics* 113 (3):325–347.

Ledoit, Oliver and Michael Wolf. 2008. "Robust performance hypothesis testing with the Sharpe ratio." *Journal of Empirical Finance* 15 (5):850 – 859.

Li, Jiahan, Ilias Tsiakas, and Wei Wang. 2014. "Predicting Exchange Rates Out of Sample: Can Economic Fundamentals Beat the Random Walk?" *Journal of Financial Econometrics* 13 (2):293–341.

Lyons, Richard. 2001. *The Microstructure Approach to Exchange Rates*. MIT Press.

Meese, Richard A. and Kenneth Rogoff. 1983. "Empirical exchange rate models of the seventies : Do they fit out of sample?" *Journal of International Economics* 14 (1-2):3–24.

Menkhoff, Lukas, Lucio Sarno, Maik Schmeling, and Andreas Schrimpf. 2012. "Carry Trades and Global Foreign Exchange Volatility." *Journal of Finance* 67 (2):681–718.

Merton, Robert C. 1973. "An Intertemporal Capital Asset Pricing Model." *Econometrica* 41 (5):867–87.

Molodtsova, Tanya and David H. Papell. 2009. “Out-of-sample exchange rate predictability with Taylor rule fundamentals.” *Journal of International Economics* 77 (2):167–180.

Papadimitriou, Theophilos, Periklis Gogas, and Efthimios Stathakis. 2014. “Forecasting energy markets using support vector machines.” *Energy Economics* 44 (0):135 – 142.

Rossi, Barbara. 2013. “Exchange Rate Predictability.” *Journal of Economic Literature* 51 (4):1063–1119.

Sarno, Lucio, Giorgio Valente, and Hyginus Leon. 2006. “Nonlinearity in Deviations from Uncovered Interest Parity: An Explanation of the Forward Bias Puzzle.” *Review of Finance* 10 (3):443–482.

Smales, Lee A. and Jardee N. Kininmonth. 2016. “FX Market Returns and Their Relationship to Investor Fear.” *International Review of Finance* .

Teräsvirta, Timo. 2006. “Forecasting economic variables with nonlinear models.” In *Handbook of Economic Forecasting*, vol. 1, edited by C.W.J. Granger G. Elliott and A. Timmermann. Elsevier, 413 – 457.

Thornton, Daniel L. and Giorgio Valente. 2012. “Out-of-Sample Predictions of Bond Excess Returns and Forward Rates: An Asset Allocation Perspective.” *Review of Financial Studies* .

Table 1: Alternative carry trade strategies

	Mean	StDev	Skew	Sharpe	Fees	Omega	Sortino	Upside
Monthly								
Equal weights								
Carry trade	2.67	4.81	-0.737	0.39		1.52	0.24	0.70
SVM	9.68	5.07	1.064	1.75***	9.23	4.96	1.51	1.89
Dynamic								
Carry trade	2.35	4.85	-0.974	0.32		1.46	0.20	0.63
SVM	9.31	6.10	0.042	1.39***	9.08	3.35	0.87	1.25
Weekly								
Equal weights								
Carry trade	-0.04	5.53	-1.237	-0.15		1.00	0.00	0.46
SVM	11.17	5.56	1.008	1.87***	14.63	2.24	0.54	0.97
Dynamic								
Carry trade	-0.71	5.84	-1.878	-0.25		0.95	-0.02	0.41
SVM	11.25	6.84	0.724	1.53***	15.44	1.94	0.40	0.82

Equal weights = equally weighted portfolio
Dynamic = carry trade on first 3 currencies rebalanced every period
All values are computed on annual basis and include transaction costs.
Significance levels for the Sharpe ratio refer to the Ledoit and Wolf (2008) test.

Table 2: Carry trade strategies, disentangling between uncertainty measures

	Mean	StDev	Skew	Sharpe	Omega	Sortino	Upside
All Sample							
Carry Trade	-0.04	5.53	-1.24	-0.15	1.00	0.00	0.46
SVM	11.17	5.56	1.01	1.87***	2.24	0.54	0.97
SVM Liquity	8.17	5.68	0.78	1.31***	1.77	0.34	0.79
SVM Volatility	6.16	5.33	0.07	1.01***	1.60	0.26	0.68
2003-2007							
Carry Trade	1.28	4.78	-0.33	0.03	1.10	0.05	0.57
SVM	13.03	5.13	-0.22	2.32***	2.60	0.65	1.05
SVM Liquity	11.99	5.50	-0.12	1.97***	2.23	0.54	0.97
SVM Volatility	7.74	4.94	-0.63	1.34***	1.82	0.33	0.74
2008-2012							
Carry Trade	-1.70	7.99	-1.41	-0.23	0.91	-0.04	0.39
SVM	16.62	7.47	1.52	2.21***	2.58	0.69	1.13
SVM Liquity	9.45	7.51	1.39	1.24***	1.71	0.32	0.76
SVM Volatility	8.56	7.19	0.31	1.17***	1.67	0.27	0.69
2013-2017							
Carry Trade	-2.71	3.77	-0.16	-0.73	0.78	-0.13	0.45
SVM	11.74	4.23	0.42	2.77***	2.95	0.81	1.23
SVM Liquity	8.50	4.27	0.20	1.98***	2.12	0.51	0.96
SVM Volatility	6.32	4.10	0.03	1.53***	1.86	0.37	0.79

Values refer to weekly returns on equally weighted portfolios
SVM liquidity= carry trade with SVM conditioned on measures of liquidity (TED and NFCI)
SVM volatility= carry trade with SVM conditioned on measures of volatility (VIX and FX Vol.)
All values are computed on annual basis and include transaction costs.
Significance levels for the Sharpe ratio refer to the Ledoit and Wolf (2008) test.

Table 3: Determinants of SVM carry trade returns

	All sample	All s.	All s.	2003-2007	2008-2012	2013-2017	All s.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Forex vol	-0.021	-0.006		-0.028**	-0.063	0.003	-0.038
	(0.016)	(0.011)		(0.012)	(0.046)	(0.014)	(0.025)
VIX	-0.001	0.007		0.005	0.004	-0.008*	-0.006
	(0.004)	(0.005)		(0.004)	(0.008)	(0.005)	(0.005)
TED	0.042		0.086**	-0.086	-0.007	0.136	0.177*
	(0.045)		(0.043)	(0.099)	(0.068)	(0.184)	(0.104)
NFCI	0.185**		0.074***	0.194	0.344**	0.175	0.268**
	(0.079)		(0.028)	(0.163)	(0.164)	(0.320)	(0.127)
R2	0.047	0.016	0.040	0.017	0.098	0.014	0.060
N	742	750	742	257	255	230	742

The table reports the coefficients of regressing returns of the SVM carry trade strategy on lagged indicators of market uncertainty.
Dependent variable: columns 1-6 returns of the SVM carry trade, column 7 the difference between the returns of the SVM and the standard carry trade.
SVM carry trade returns refer to weekly rates on equally weighted portfolios.
The reported standard errors are based on NeweyâĂŞWest approach with optimal lag selection.
* p<0.10, ** p<0.05, *** p<0.01

Figure 3: Standardized beta coefficients from quantile regressions on carry trade returns

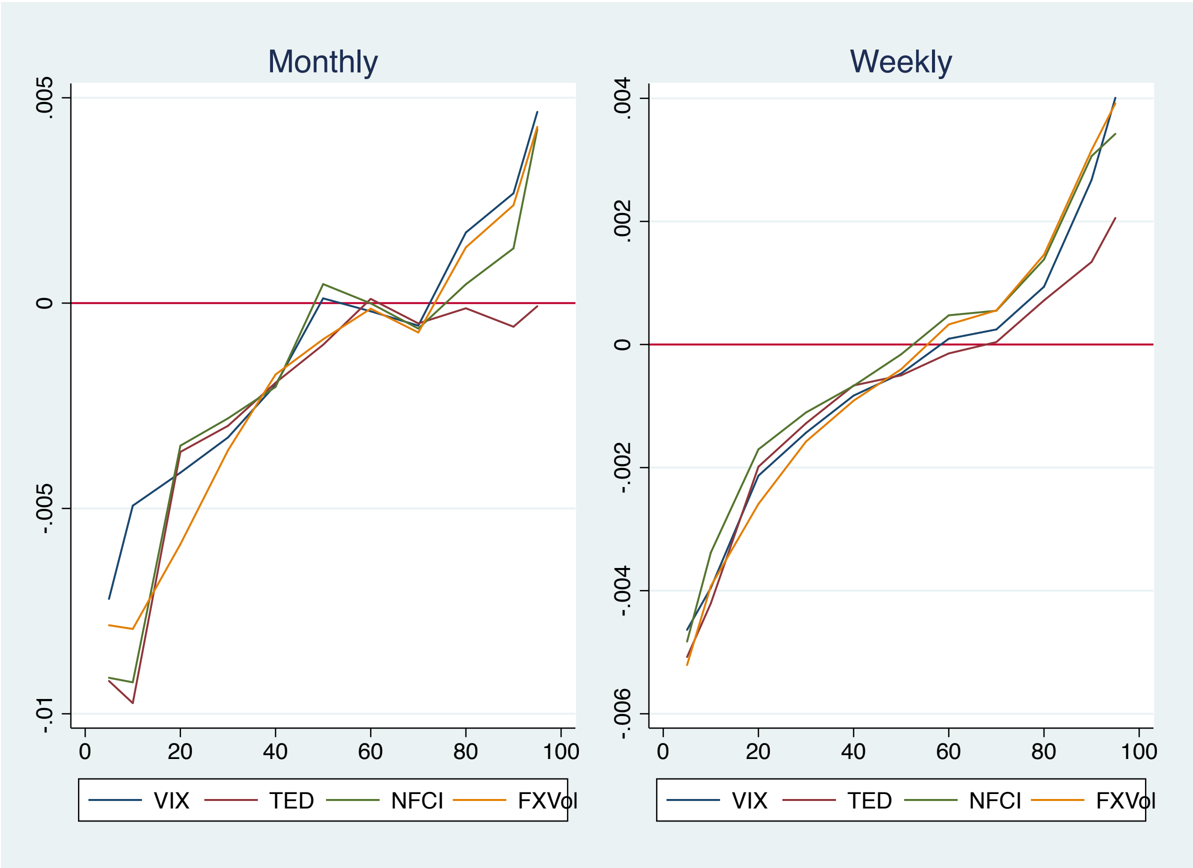


Figure 4: R2 from predictive regressions of the carry trade returns on various measures of uncertainty

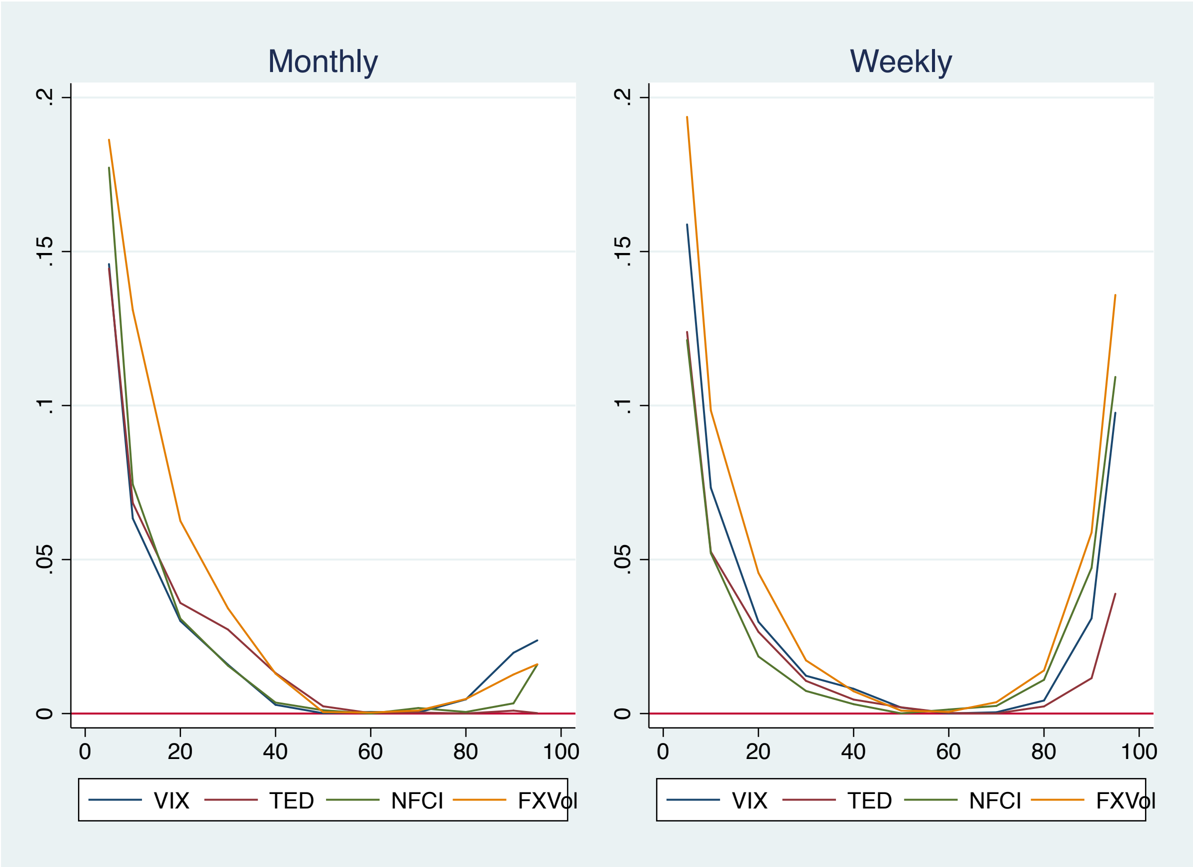


Figure 5: Agreement and disagreement between SVM and standard carry trade strategy

Figure 6: Distribution of returns of the carry trade and SVM model; equally weighted portfolio, monthly forward rates.

Figure 7: Returns from carry trade strategies; equally weighted portfolio, monthly forward rates.

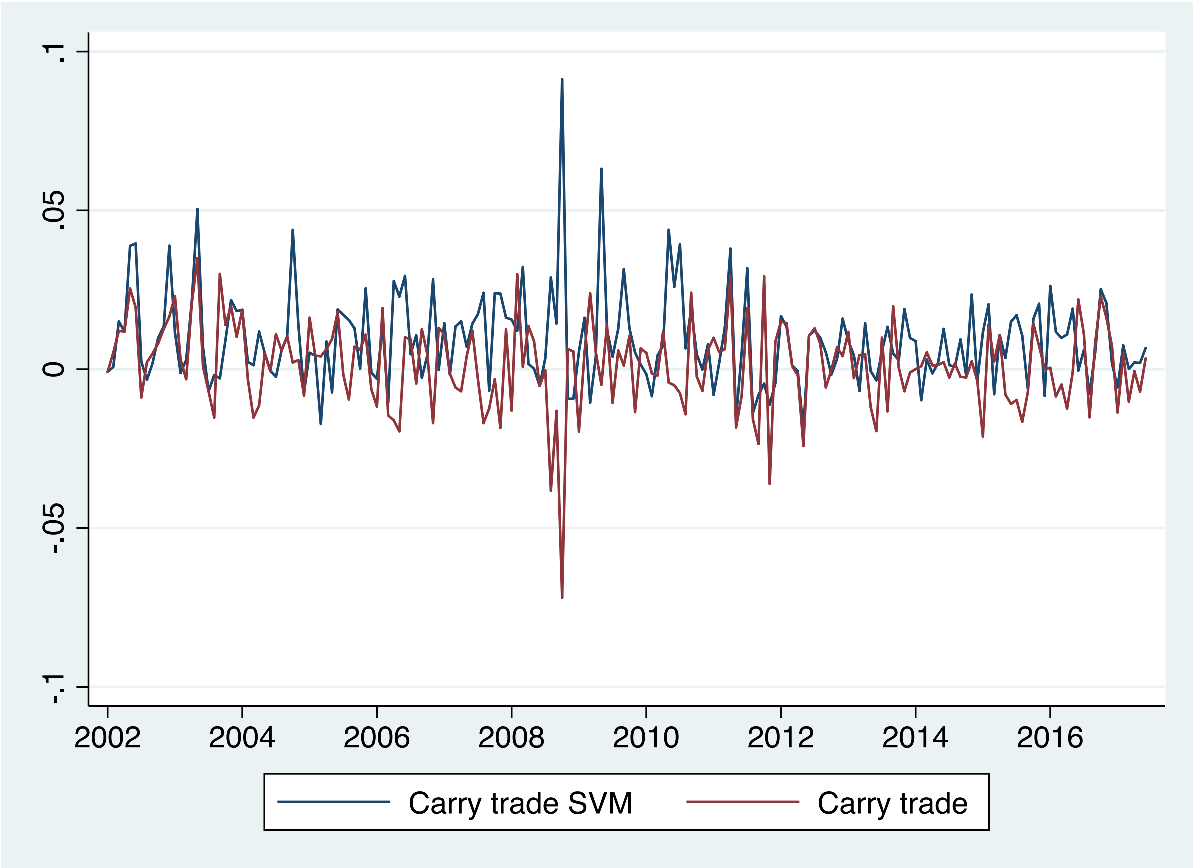


Figure 8: Cumulative returns from carry trade strategies, monthly rates

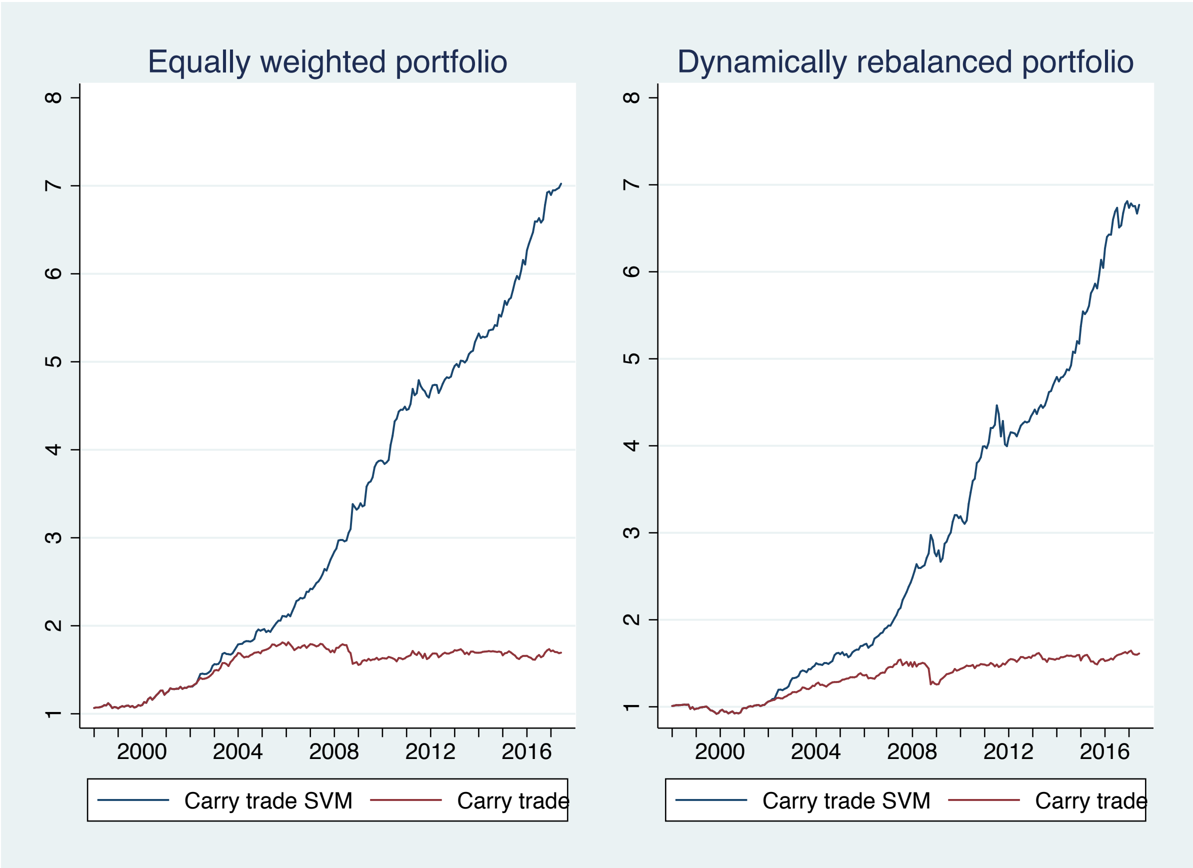


Figure 9: Cumulative returns from carry trade strategies, weekly rates