

This PDF is a selection from a published volume from the National Bureau of Economic Research

Volume Title: NBER International Seminar on Macroeconomics 2010

Volume Author/Editor: Richard Clarida and Francesco Giavazzi, organizers

Volume Publisher: University of Chicago Press

Volume ISBN: 978-0-226-10736-3 (cloth); 0-226-10738-8 (paper)

Volume URL: <http://www.nber.org/books/clar10-1>

Conference Date: June 18-19, 2010

Publication Date: September 2011

Chapter Title: Currency Carry Trades

Chapter Authors: Travis Berge, Òscar Jordà, Alan M. Taylor

Chapter URL: <http://www.nber.org/chapters/c12216>

Chapter pages in book: (p. 357 - 387)

---

## Currency Carry Trades

Travis Berge, *University of California, Davis*

Òscar Jordà, *University of California, Davis*

Alan M. Taylor, *University of California, Davis, and NBER*

### I. Introduction

One of the oldest and most frequently recurring questions in international finance concerns the efficiency of the foreign exchange market. Indeed it is one of the most durable and intriguing questions in the field of finance as a whole since the market for major currencies is one of the largest, most liquid, and most actively traded asset markets in existence. Thus, treated as a laboratory, this market more than any other may have the potential to reveal how close actual financial markets are to attaining their textbook idealized form: are asset returns essentially random or do they have systematically predictable elements?<sup>1</sup>

For several decades a long literature has sought to explore whether currency returns are forecastable, and the simple “carry trade” logic of trading based on the interest differential has been very widely studied. Here systematic *ex post* profits are widely observed, a phenomenon that is merely a manifestation of the long-studied forward discount puzzle (see, e.g., Frankel 1980; Fama 1984; Hodrick 1987; Froot and Thaler 1990; Bekaert and Hodrick 1993; Engel 1996). Notwithstanding this broadly accepted puzzle, a number of metrics have been used to evaluate the predictability and profitability of exchange rate forecasts, and the results have by no means created consensus. Researchers have asked whether such forecasting power delivers statistically significant fit relative to random walk and if the forecast can generate economically significant profits for a risk-neutral investor after transaction costs (Meese and Rogoff 1983; Kilian and Taylor 2003). Researchers have also sought to account for the possibility of time-varying risk premia—but they must then navigate between the inevitably circular reasoning that *ex post* risk premia could be found that in principle explain any *ex post* returns observed and the problem that observable so-called risk factors

are (apart from consumption growth) often atheoretic and ad hoc regressors vulnerable to a “ketchup” critique.<sup>2</sup>

Carry trades are now under scrutiny again. In conjunction with a dramatic rise in real world currency trading in the last decade, a recent wave of research on exchange rate forecastability has appeared in the last few years asking new questions and sharpening new tools. Much of this literature has continued to focus on strategies based on the naive carry signal, where investors go long high-yield currencies and short low-yield currencies (e.g., Burnside et al. 2006, 2008a, 2008b; Brunnermeier, Nagel, and Pedersen 2008). These pre-financial-crisis studies also often found attractive residual profits, with moderately impressive Sharpe ratios. However, the strategies often came with unattractive third moments, with high negative skew resulting from the occasional tendency of target currencies to crash, or conversely funding currencies to suddenly appreciate (e.g., the well-known Japanese yen events of 1998). While one could in principle truncate the downside risks by augmenting the strategy with put options (Burnside et al. 2008b; Jurek 2008; Clarida, Davis, and Pedersen 2009), these insurance mechanisms are not inexpensive and entail the further complication of making assumptions concerning liquidity and counterparty risk in derivative markets.

The working of the foreign exchange market during the financial crisis challenged some of these findings. Investors using pure carry strategies fared very poorly indeed. Moreover, at key moments derivative markets malfunctioned, and counterparty risks could no longer be ignored, raising questions about options-based insurance strategies. However, alternative strategies with attractive returns and crash protection have come to light. These could be described as “augmented carry” models, as they exploit additional conditioning information. In some recent research that is directly antecedent to the current paper, Jordà and Taylor (2009, 2011) developed new tools to study directional trading strategies based on a set of three signals: carry, momentum, and value (CMV). They applied the receiver operating characteristic (ROC) curve to evaluate the directional performance of these signals one at a time, and jointly. They also extended the ROC techniques by constructing a return-weighted ROC\* curve, with analogous properties, which could be used to evaluate the profitability of various signals when used for trading.

In this paper, we refine and extend the ROC techniques and apply them to a broader set of signals that includes information on the forward curve, and we examine the robustness of our results when confronted with explanations based on the so-called risk factors. We focus on methods which are based on the fact that ROC analysis is equivalent

to the analysis of a *correct classification* frontier (CC), a concept that we think has a more natural economic interpretation and that extends easily to the return-weighted case. And we are interested in exploring whether the Jordà and Taylor CMV signals still contain useful predictive information even when one includes forward-curve data following the insights of Clarida and Taylor (1997), Chen and Tsang (2009), Clarida et al. (2009), and Ang and Chen (2010).

Using our new tools we are able to show that the CMV signals and the forward curve signals each contain independent and valuable predictive information. We find very strong evidence of forecastability using the full set of signals, both in sample and out of sample. Our preferred model generates economically meaningful returns on a portfolio of G-10 currencies, with favorable Sharpe and skewness characteristics. From an efficiency standpoint, a risk-neutral investor would find these trades profitable even allowing for transaction costs. The returns are also uncorrelated with the microfounded consumption growth risk factor. And although explanations based on unobservable time-varying risk premia are not testable, we find no relationship between our returns and a long menu of so-called risk factors either, casting doubt on the potential objection that our currency trade profits could reflect beta rather than alpha.

## II. Statistical Design

Uncovered interest rate parity (UIP) in an ideal, risk-neutral, frictionless world is a condition that suggests that nominal excess returns to currency speculation, based on arbitraging differences in nominal interest rates across countries, are expected to be zero. Let  $x_{t+1}$  denote the ex post, monthly excess returns (in logarithms) given by

$$x_{t+1} = \Delta e_{t+1} + (i_t^* - i_t), \tag{1}$$

where  $e_{t+1}$  is the logarithm of the nominal exchange rate in U.S. dollars per foreign currency unit, and  $i_t^*$  and  $i_t$  denote the 1-month London interbank offered rates (LIBOR) abroad and in the United States, respectively. Therefore, standard arbitrage arguments suggest that  $E_t x_{t+1} = 0$ .

Excess returns can be easily expressed in real terms, offering a natural link to the purchasing power parity (PPP) condition. Let  $\pi_{t+1}$  denote the inflation rate calculated as the log-difference of the price level, that is,  $\pi_{t+1} = \Delta p_{t+1}$  (and similarly for  $\pi_{t+1}^*$ ). Adding and subtracting  $\pi_{t+1}^* - \pi_{t+1}$  to the

right-hand side of expression (1) and defining the real exchange rate as  $q_{t+1} = \bar{q} + e_{t+1} + (p_{t+1}^* - p_{t+1})$ , then

$$x_{t+1} = \Delta q_{t+1} + [(i_{t+1}^* - \pi_{t+1}^*) - (i_{t+1} - \pi_{t+1})]. \quad (2)$$

Since  $\Delta q_{t+1}$ ,  $i_t^*$ ,  $i_t$ ,  $\pi_{t+1}^*$ , and  $\pi_{t+1}$  are  $I(0)$  variables,  $q_{t+1}$  can be most naturally thought of as a cointegrating vector under PPP where  $\bar{q}$  is the fundamental equilibrium exchange rate (FEER) toward which  $q_{t+1}$  reverts in the long run.

A speculator trying to take advantage of condition (1), say, will be interested in constructing a forecast  $\Delta \hat{e}_{t+1} = E_t \Delta e_{t+1}$  since at time  $t$ ,  $(i_{t+1}^* - i_{t+1})$  is known. A linear forecast that articulates the underlying UIP and PPP conditions captured in expressions (1) and (2) is best formulated by considering the stochastic process for the  $I(0)$  random vector

$$\Delta y_{t+1} = \begin{bmatrix} \Delta e_{t+1} \\ \pi_{t+1}^* - \pi_{t+1} \\ i_{t+1}^* - i_{t+1} \end{bmatrix},$$

where  $q_t$  is a natural (and unique) cointegrating vector due to PPP. Therefore, it is appropriate to entertain that the stochastic process for  $\Delta y_{t+1}$  is a vector error correction model (VECM), where, for example, the first-order expression for  $\Delta e_{t+1}$  is easily seen to be

$$\Delta e_{t+1} = \beta_0 + \beta_e \Delta e_t + \beta_\pi (\pi_t^* - \pi_t) + \beta_i (i_t^* - i_t) + \gamma (q_t - \bar{q}) + u_{t+1}. \quad (3)$$

Expression (3) nests four popular approaches to currency trading: carry, value, and momentum signals used singly and a composite based on a mix of all three CMV signals. For example, the CMV approach underlies each of the three popular tradable ETFs created by Deutsche Bank, where in each case a nine-currency portfolio is sorted into equal-weight long-neutral-short thirds based on the relative strength of each of the three signals, and regularly rebalanced. In addition, Deutsche Bank offers a composite rebalancing portfolio split one-third between each of the CMV portfolios. Similar tradable indices and ETF products have since been launched by other financial institutions (e.g., Goldman Sachs's FX Currents, and Barclays Capital's VECTOR).

For our purposes, we will define four model-based strategies of this form for use in this paper. We shall assume that, at time  $t$ , the currency

trader determines which currency to go long with and which to short depending on

$$\hat{d}_{t+1} = \text{sign}(\hat{x}_{t+1}) \in \{-1, 1\}; \quad \hat{x}_{t+1} = \Delta \hat{e}_{t+1} + (i_t^* - i_t),$$

where  $\Delta \hat{e}_{t+1}$  is the one-period-ahead forecast of  $\Delta e_{t+1}$  and where  $\hat{x}_{t+1}$  is determined as follows for each of the four strategies considered.

**Carry:**  $\Delta \hat{e}_{t+1} = 0;$

**Momentum:**  $\Delta \hat{e}_{t+1} = \hat{\beta}_e \Delta \hat{e}_t;$

**Value:**  $\Delta \hat{e}_{t+1} = \hat{\gamma}(q_t - \bar{q});$

**VECM:**  $\Delta \hat{e}_{t+1} = \hat{\beta}_0 + \hat{\beta}_e \Delta e_t + \hat{\beta}_\pi (\pi_t^* - \pi_t) + \hat{\beta}_i (i_t^* - i_t) + \hat{\gamma}(q_t - \bar{q}).$

We make a few asides about these candidate strategies. First, the pure carry strategy as written is somewhat simplified. It could be recast in more general or less restrictive terms. For a pure directional bet using a single strategy, it need not be the case that the exchange rate has a zero expectation; it is enough for this strategy to work that the interest differential predict the direction of profitable trade; that is, that  $\Delta \hat{e}_{t+1} = \beta_i (i_t^* - i_t)$  with  $\beta_i > -1$ , such that predicted profit  $\Delta \hat{x}_{t+1} = (\beta_i + 1)(i_t^* - i_t)$  is still forecast to be positive on the carry-based directional bet. Second, one might also consider the inflation differential  $(\pi_t^* - \pi_t)$  as providing a possible fifth signal for traders, which might be referred to as the “monetary policy” or “Taylor rule” signal. The logic for this signal is that if central banks are inflation targeting and are using some kind of feedback rule, then “bad news” about inflation could be “good news” for the exchange rate. See, for example, Clarida and Waldman (2007), who find evidence for this in some cases where policy regimes shifted in the 1990s. Third, one could obviously envisage many other candidate strategies that we do not consider here for reasons of space. One obvious candidate would be to use measures of implied FX volatility as additional conditioning variables or as interacted variables with the above signals. Another set of signals could be built around measures that capture financial market distress, as in Melvin and Taylor (2009).

Focusing henceforth on the main four strategies that we have identified above, ex post returns realized by a trader engaged in any of these strategies are therefore

$$\hat{\mu}_{t+1} = \hat{d}_{t+1} x_{t+1}. \tag{4}$$

Note that the trader need not be particularly accurate in predicting  $\Delta e_{t+1}$  (which has been known to be a futile task at least since Meese and Rogoff [1983]), as long as  $\hat{d}_{t+1}$  correctly selects the direction of the carry trade. Recent work by Cheung, Chinn, and Garcia Pascual (2005) and Jordà and Taylor (2011) suggests that directional forecasts of exchange rate movements perform better than a coin toss, leaving the door open for us to evaluate the economic value of a carry trade investment.

The carry trade is a zero net-investment strategy. As such, fundamental models of consumption-based asset pricing in frictionless environments with rational agents would suggest that, if  $m_{t+1}$  denotes the stochastic discount factor (see, e.g., Cochrane 2001), then

$$E_t(m_{t+1}x_{t+1}) = 0.$$

Thus, in order to explain the observation that the carry trade enjoys long periods of persistently positive net returns, one has to examine how good a hedge against consumption-growth risk the carry trade is relative to other investments and risk factors. For example, Burnside et al. (2008a, 2008b) argue that the correlation of carry trade returns with conventional risk factors is insufficient to justify carry trade returns but that these could be reconciled with standard results by interpreting carry trade returns as compensation for large tail risk (dubbed “peso events” in their paper).

Recent theoretical work (e.g., Shleifer and Vishny 1997; Jeanne and Rose 2002; Bacchetta and van Wincoop 2006; Fisher 2006; Brunnermeier et al. 2008; Ilut 2008) try to explain what generates this tail risk using a combination of market microstructure mechanisms, such as models of noise traders, heterogeneous beliefs, rational inattention, liquidity constraints, herding, “behavioral effects,” and other factors that may serve to limit arbitrage.

Our empirical strategy follows a two-pronged approach that differs from what is usually done. In the first prong, we extend the four basic carry trade strategies outlined above with country-specific yield curve factors extracted using Nelson and Siegel’s (1987) approach. There are some obvious and intuitive reasons for doing this: first, because the Nelson and Siegel factors are natural predictors of relative cyclical positions between two countries and hence of relative UIP and PPP; and second, because yield curves are natural candidates as risk factors in many asset markets, and so it makes sense to examine their covariation with carry trade returns. More formally, as Clarida and Taylor (1997) have shown, in a model with persistent short-run deviations

from the risk-neutral efficient markets hypothesis, expectational errors can induce a nonzero correlation between information in the forward yield curve and the future path of the exchange rate. However, even if one chooses to be agnostic about the theoretical channel, and even if Nelson and Siegel yield curve factors do not necessarily serve to justify carry trade returns, it might still be the case that they could predict carry trade direction and help dilute tail risk. To make the link explicit to prior work, Jordà and Taylor (2011) found FEER to provide a superior hedge (in terms of Sharpe ratio) against this tail risk than the options-based hedge proposed in Burnside, Eichenbaum, and Rebelo (2008). Here we are interested in comparing the FEER hedge with a Nelson and Siegel hedge and a combination of the two.

The second prong examines how country-specific carry trade returns with respect to the United States covary with U.S. risk factors such as value-weighted excess returns in the U.S. stock market (CAPM); three Fama and French (1993) factors (excess returns to another measure of value-weighted U.S. stock market); the size premium; and the value premium; U.S. industrial production growth; the federal funds rate; the term premium (measured by the spread between 10-year Treasury bonds and 3-month Treasury bills); the liquidity premium (measured as the spread between the 3-month Eurodollar rate and the 3-month Treasury bill rate); the Pastor and Stambaugh (2003) liquidity measures; and four measures of market volatility: VIX, VXO, the change in VIX, and the change in VXO.<sup>3</sup> Covariation with any of this long list of conventional risk factors can help us understand why it appears that there are excess returns to be made with the carry trade. Before investigating these questions, we discuss some important methodological issues.

### III. Evaluating Realized Carry Trade Returns

Before we present the main results of our empirical strategy, two novel approaches to out-of-sample investment performance evaluation from a trader's perspective are discussed in this section (Jordà and Taylor 2009). Recall that ex post realized returns for a given carry trade strategy are given by expression (4), repeated here for convenience:

$$\hat{\mu}_{t+1} = \hat{d}_{t+1} x_{t+1},$$

where  $\hat{d}_{t+1} = \text{sign}(\hat{x}_{t+1})$ . Moreover, we are interested in investigating the decisions a typical trader makes given the information available



to him at time  $t$ , and for this reason we are less interested in model fit and more interested in out-of-sample evaluation. That is, our objective is to assess carry trade returns from an investor's point of view.

From this perspective, the natural loss functions required for out-of-sample predictive evaluation are determined by investor-performance measures rather than by the more habitual root mean-squared error (RMSE) metric. In addition and because we are interested in examining one-period ahead forecasts from a rolling sample of fixed length (where estimation uncertainty never disappears regardless of the sample size), it is appropriate to rely on conditional predictive ability tests à la Giacomini and White (2006), rather than unconditional predictive ability tests à la Diebold and Mariano (1995).

Note that the key ingredient in expression (4) is  $\hat{a}_{t+1} \in \{-1, 1\}$ , which is a binary variable. Here conventional methods for evaluating predicted probability outcomes are not useful. Instead, we are interested in evaluating the ability to predict directional outcomes for the purposes of maximizing return, and this distinction turns out to be quite important: such an evaluation requires tools that take into account differences in risk attitudes across different investors. In the next sections we explain each of these evaluation tools in detail.

#### A. Trading-Based Predictive Ability Tests

Given a sample of size  $T$ , suppose we reserve the first  $R$  observations to produce a forecast for  $t = R + 1$  and then roll the sample by one observation. This generates  $P = T - (R + 1)$ , one-period-ahead forecasts obtained with rolling samples of size  $R$ . Accordingly, let  $\{L_{t+1}\}_{t=R}^{T-1}$  denote the sequence of loss functions associated with a given forecasting model. The Giacomini and White (2006) test statistic that evaluates the out-of-sample, conditional marginal ability between two models is

$$GW = \frac{\Delta \bar{L}}{\hat{\sigma}_L / \sqrt{P}} \rightarrow N(0, 1),$$

with

$$\Delta \bar{L} = \frac{1}{P} \sum_{t=R}^{T-1} (L_{t+1}^1 - L_{t+1}^0); \quad \hat{\sigma}_L^2 = \frac{1}{P} \sum_{t=R}^{T-1} (L_{t+1}^1 - L_{t+1}^0)^2,$$

where the superscripts 0 and 1 refer respectively to the null and alternative models under consideration. The loss functions that we consider for each model include the traditional MSE given by

$$\Delta L_{t+1} = (\hat{x}_{t+1}^1 - x_{t+1})^2 - (\hat{x}_{t+1}^0 - x_{t+1})^2$$

and the following three investment-performance measures:

**Return:**

$$\Delta L_{t+1} = \hat{\mu}_{t+1}^1 - \hat{\mu}_{t+1}^0,$$

**Sharpe Ratio:**

$$\Delta L_{t+1} = \frac{\hat{\mu}_{t+1}^1}{\hat{\sigma}_1} - \frac{\hat{\mu}_{t+1}^0}{\hat{\sigma}_0},$$

where  $\hat{\sigma}_i$  are calculated for each country individually over the predictive sample.

**Skewness:**

$$\Delta L_{t+1} = \left( \frac{\hat{\mu}_{t+1}^1}{\hat{\sigma}_1} \right)^3 - \left( \frac{\hat{\mu}_{t+1}^0}{\hat{\sigma}_0} \right)^3.$$

We remark that  $\hat{\sigma}_L^2$  is estimated with a cluster robust option to account for country-specific effects. Briefly, the *return* loss function compares the relative returns between a given carry trade strategy (carry, momentum, value, and VECM) relative to a coin-toss null model; the *Sharpe ratio* loss function examines the Sharpe ratio instead so as to down-weight carry trade returns by country-specific risk; and the *skewness* loss function examines a country-specific skewness proxy of carry trade returns that keeps the general format of the other investment-performance loss functions and puts more weight on returns that are positively skewed and hence avoids left tail risk.

*B. Trading-Based Classification Ability Tests*

Realized returns  $\hat{\mu}_{t+1}$  depend critically on  $\hat{d}_{t+1} \in \{-1, 1\}$ , a binary prediction on the profitable direction of the carry trade, that is,  $d_{t+1} = \text{sign}(x_{t+1})$ .

It would seem natural to construct  $\hat{d}_{t+1}$  using a prediction  $\hat{x}_{t+1}$  in a latent probability model. However, note that we do not require a probability prediction but an actual decision on the direction that the investor should take. In general such a prediction will be based on a single index that appropriately combines information up to time  $t$ , say  $\hat{\delta}_{t+1}$  of which a special case is  $\hat{\delta}_{t+1} = \hat{x}_{t+1}$ . Given  $\hat{\delta}_{t+1}$ , then we determine  $\hat{d}_{t+1} = \text{sign}(\hat{\delta}_{t+1} - c)$ , where  $c$  is a threshold whose value depends critically on an investor's preferences and attitudes toward risk as well as the distribution of returns. We will explain this issue in more detail momentarily. Notice that in this setup,  $\hat{\delta}_{t+1}$  need not be bounded between 0 and 1, as is customary in a logit or a probit model.

It is useful to define the following classification table associated to this binary prediction problem:

Outcome	Prediction	
	Negative	Positive
Negative	$TN(c) = P(\hat{\delta}_{t+1} < c   d_{t+1} = -1)$	$FP(c) = P(\hat{\delta}_{t+1} > c   d_{t+1} = -1)$
Positive	$FN(c) = P(\hat{\delta}_{t+1} < c   d_{t+1} = 1)$	$TP(c) = P(\hat{\delta}_{t+1} > c   d_{t+1} = 1)$

where  $TN$  and  $TP$  (true negative and true positive) refer to the correct classification rates of negatives and positives, respectively;  $FN$  and  $FP$  (false negative and false positive) refer to the incorrect classification rates of negatives and positives, respectively; and clearly  $TN(c) + FP(c) = 1$  and  $FN(c) + TP(c) = 1$ . Customarily,  $TP(c)$ , the true positive rate, is called *sensitivity* and  $TN(c)$ , the true negative rate, is called *specificity*.

The space of combinations of  $TP(c)$  and  $TN(c)$  for all possible values of  $c$  such that  $-\infty < c < \infty$  summarizes a sort of production possibilities frontier (to use the microeconomics nomenclature for the space for two goods, in our case the two correct classification rates that we contemplate) for a classifier  $\hat{\delta}_{t+1}$ , that is, the maximum  $TP(c)$  achievable for a given value of  $TN(c)$ . We will call this curve the *correct classification* (CC) frontier. In statistics and other scientific fields, it is more common to represent the curve associated with the combinations of  $TP(c)$  and  $FP(c)$ , called the receiver operating characteristics curve (ROC), but since  $FP(c) = 1 - TN(c)$ , the CC frontier is equivalent to the ROC curve if one reverses the horizontal axis. Because economists may prefer the manner in which the CC frontier is constructed, we maintain the different nomenclature to avoid confusion.

In  $(TN, TP)$  space, let the abscissa represent  $TN$  and the ordinate  $TP$ . Notice that for any classifier, as  $c \rightarrow -\infty$ , then  $TN(c) \rightarrow 1$  but  $TP(c) \rightarrow 0$  and vice versa, as  $c \rightarrow \infty$ ,  $TN(c) \rightarrow 0$  as  $TP(c) \rightarrow 1$ .

Let us first consider an *uninformative* classifier as a benchmark. Imagine we have a classifier that randomly calls  $P$  with probability  $p$  and  $N$  with probability  $n = 1 - p$ , like a biased coin. These calls are independent of the true state. The true positive rate is  $TP = p$  and the true negative rate is  $TN = n = 1 - p$ , and these satisfy  $TN + TP = 1$  by construction. Thus, an *uninformative* classifier that is no better than a coin toss is one with frontier given by  $TN(c) = 1 - TP(c)$  for all  $c$ , that is, the diagonal or 1-simplex that runs from  $(0, 1)$  to  $(1, 0)$  in  $(TN, TP)$  space. In contrast, a perfect classifier instead hugs the northeast corner of the  $[0, 1] \times [0, 1]$  square subset of  $(TN, TP)$  space.

An example of a typical CC frontier is shown in figure 1. We also show CC frontiers for the cases of a perfect classifier and for a coin-toss classifier.

How does an investor put to use a binary classifier of this type? Faced with a CC frontier of this kind he could choose to operate his trading strategy at any one of the points on the frontier by choosing a particular threshold parameter  $c$ . What, then, is the optimal choice of  $c$ ?

A risk-neutral investor will choose the value of  $c$  where the slope of the CC frontier is equal to his utility's marginal rate of substitution

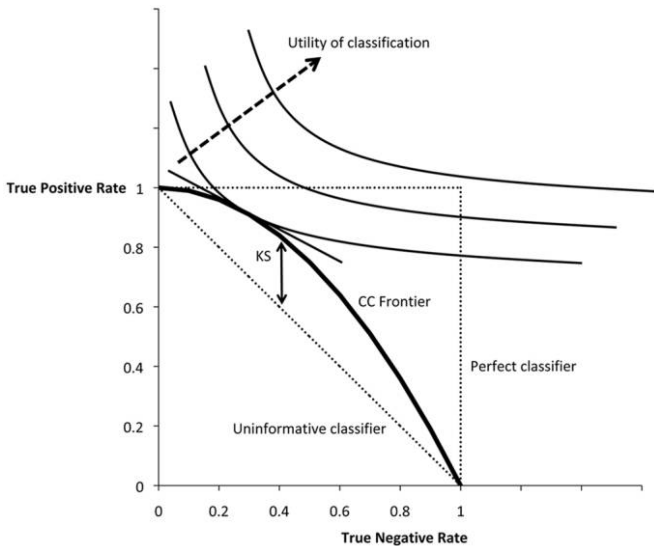


Fig. 1. The correct classification frontier

(MRS) between the utility of  $TP(c)$  and  $TN(c)$  outcomes. When  $x_{t+1}$  is symmetrically distributed, then this MRS is  $-1$ , and the distance between the CC frontier at this point and the coin-toss diagonal turns out to be the Kolmogorov-Smirnov (KS) statistic (see Jordà and Taylor [2011] for an explanation of this result).

Intuitively, the KS statistic is a nonparametric test based on the distance given by the average correct classification rates of the classifier under investigation and those for a coin-toss classifier. In the latter case, recall that  $TN(c) = 1 - TP(c)$  for all  $c$ , and hence the average is 0.5. The formula for the KS statistic is

$$KS = \max_c 2 \left| \left( \frac{TN(c) + TP(c)}{2} \right) - \frac{1}{2} \right| \rightarrow \sup_{\tau \in [0,1]} |B(\tau)|, \quad (5)$$

where  $B(\tau)$  is a Brownian bridge and  $KS \in [0, 1]$  (see Conover 1999). The KS statistic can also be interpreted as identifying the operating point that maximizes the Youden (1950) index for medical diagnostic testing, or Peirce's (1884) "science of the method."

An estimate of KS can be obtained nonparametrically by realizing that

$$\widehat{TP}(c) = \frac{\sum_{d_{t+1}=1} I(\hat{\delta}_{t+1} > c)}{\sum_{d_{t+1}=1} 1}; \quad \widehat{TN}(c) = \frac{\sum_{d_{t+1}=-1} I(\hat{\delta}_{t+1} \leq c)}{\sum_{d_{t+1}=-1} 1}.$$

However, the above choice of operating point is clearly very special. In general, an investor's preferences are unknown and the distribution of returns may not be symmetric, and so it is customary and sensible to construct a different test statistic based on the area under the entire CC frontier, rather than its height at any point. This area we shall denote with the acronym AUC for *area under the curve*. This area is equivalent to the area under the ROC curve. The AUC for a coin-toss classifier is clearly 0.5 (the area under the simplex), whereas  $AUC = 1$  for a perfect classifier. In most practical scenarios, the AUC will fall between these two extreme values.

The AUC turns out to have a convenient interpretation as a Wilcoxon-Mann-Whitney rank-sum statistic, and inference is simplified by the fact that its distribution is well approximated by the Gaussian distribution in large samples (see Hsieh and Turnbull 1996). This distribution is centered at 0.5 and hence provides an easy way to test a classifier's mettle against

the null of no-classification ability given by the coin toss, a natural null in an investment problem (for a discussion about the properties of AUC, see Jordà and Taylor [2011]). Let  $N_p = \sum_{d_{t+1}=1} 1$  and, similarly,  $N_N = \sum_{d_{t+1}=-1} 1$  and define two subsidiary random variables as follows:  $\{\phi_i\}_{i=1}^{N_p} = \{\hat{\delta}_{t+1} | d_{t+1} = 1\}$  and  $\{\eta_j\}_{j=1}^{N_N} = \{\hat{\delta}_{t+1} | d_{t+1} = -1\}$ , then the AUC can be calculated as

$$AUC = \frac{1}{N_p N_N} \sum_{j=1}^{N_N} \sum_{i=1}^{N_p} I(\phi_i > \eta_j),$$

where  $I(\cdot)$  is an indicator function that takes the value of one when the condition is true, zero when false. To find asymptotic formulae for the variance the reader is referred to Obuchowski (1994), which also includes a discussion on bootstrap procedures.

There is one final refinement we need to apply these binary classification methods to a broad set of economics and finance problems. Thus far we have weighted all correct (+1) and incorrect (-1) calls the same, but in reality, traders face nonuniform and stochastic returns from directional bets. More precisely, the AUC assesses classification ability but does not take into account the returns associated with each trade. Thus, using the simple AUC metric above, a model that correctly classifies many trades worth pennies but misses a few trades worth dollars will turn out to be a poor investment strategy even if its classification ability appears to be very good. Conversely, by correctly picking the direction in a few trades with large returns, even while missing many possibly small-return trades, an investment strategy will be very attractive, even if it is poor in the strict classification sense.

To address this issue, Jordà and Taylor (2009, 2011) introduce variants of the KS and AUC statistics to evaluate return-weighted classification ability, and these are called, respectively,  $KS^*$  and  $AUC^*$ . Both of these statistics turn out to have the same asymptotic distributions as their unweighted counterparts, so inference is still straightforward. To proceed, consider the calculation of the return-weighted versions of  $TP^*(c)$  and  $TN^*(c)$  denoted with a star superscript, both of these being key ingredients in a calculation of  $KS^*$  and  $AUC^*$ . Define the maximum profits attainable in each trading direction as

$$B \max = \sum_{d=1} x_{t+1}; \quad C \max = \sum_{d=-1} |x_{t+1}|;$$

and hence define

$$TP^*(c) = \frac{\sum_{\hat{d}(c)=1|d=1} x_{t+1}}{B \max}; \quad TN^*(c) = \frac{\sum_{\hat{d}(c)=-1|d=-1} |x_{t+1}|}{C \max}.$$

Each of these expressions represents the ability of the trading rule to extract trading gains, since each represents the ratio of actual profits to maximum potential profits in each type of bet.

It is now readily apparent how to compute  $KS^*$  using these expressions along with the definition in (5). Similarly, define weights

$$w_i = \frac{x_i}{B \max}; \quad w_j = \frac{|x_j|}{C \max},$$

then

$$AUC^* = \sum_{i=1}^{N_N} w_i \sum_{j=1}^{N_P} w_j I(\phi_i > \eta_j).$$

For a more detailed explanation of the properties of these statistics the reader is referred to Jordà and Taylor (2009, 2011). These techniques form the basis of the empirical analysis that follows.

#### IV. A Trading Laboratory for the Carry Trade

We now begin our empirical analysis by examining the first of the two prongs outlined in Section II, that is, we investigate whether Nelson and Siegel factors improve the ability of carry trade portfolios to generate positive returns with less risk and fewer “peso events.” The data used for this part of the analysis consist of a panel of nine countries (Australia, Canada, Germany, Japan, Norway, New Zealand, Sweden, Switzerland, and the United Kingdom) relative to the United States, with the sample period being monthly observations between January 1986 and December 2008 (i.e., the G-10 currency set).

The observed variables include end-of-month nominal exchange rates expressed in foreign currency units per U.S. dollar; government debt yields of the following maturities: 3, 6, 12, 24, 36, 60, 84, and 120 months (when available); the 1-month London interbank offered rates (LIBOR);

and the consumer price index. Exchange rate data and consumer price indices are obtained from the IMF's International Finance Statistics database, LIBOR data are from the British Banker's Association, and yields of government debt were obtained from Global Financial Data.

A. Four Benchmark Carry Trade Strategies

Section II discussed four carry trade strategies labeled as *carry*, *momentum*, *value*, and *VECM*, which we now investigate. The carry strategy is a naive trading model in which a trader borrows in the low interest currency and invests in the high yield currency, that is, it is entirely based on the sign of  $(i_t^* - i_t)$  (in our simplified model version it assigns  $\Delta \hat{e}_{t+1} = 0$ ). For this reason, this strategy is a natural benchmark against which to compare the other strategies that we examine.

Table 1 reports the panel-based estimates of our four models over the entire sample of data from January 1986 to December 2008 (top panel), where we omit fixed effects estimates for brevity. Table 2 then examines the out-of-sample performance over the sample January 2004–December

**Table 1**  
Four Benchmark Carry Trade Strategies: In-Sample Estimates, January 1986–December 2008

	Dependent Variable: $\Delta e_{t+1}$			
	Carry Trade Strategy			
	Carry	Momentum	Value	VECM
$\Delta e_t$	...	.114** (.031)	...	.128** (.032)
$i_t^* - i_t$	...	...	...	.864** (.358)
$\pi_t^* - \pi_t$	...	...	...	.230* (.122)
$q_t - \bar{q}$	...	...	-.014** (.003)	-.023** (.003)
$R^2$	...	.013	.004	.025
Currencies	9	9	9	9
Periods	276	276	276	276
Total observations	2,484	2,466	2,466	2,455

Note: Panel estimates with country fixed effects (not reported). Heteroskedasticity-robust standard errors reported in parentheses. Slight differences in the total number of observations are due to differences in data lags.

\*Indicates significance at the 90% confidence level.

\*\*Indicates significance at the 95% confidence level.



**Table 2**  
Four Benchmark Carry Trade Strategies: Out-of-Sample Performance, January 2004–December 2008

Realized Returns to an Equally Weighted Portfolio	Carry Trade Strategy			
	Carry	Momentum	Value	VECM
Mean (monthly)	-.0024	.0027	-.0022	.0025
SD	.018	.018	.015	.018
Skewness	-2.97	1.59	-.69	1.44
Coefficient of variation	-7.35	6.74	-6.77	7.35
Average annual return (%)	2.9	3.3	-2.6	3.0
Sharpe ratio (annual)	-.47	.51	-.51	.47
Giacomini and White (2006) <i>p</i> -Values <sup>a</sup>				
Loss function:				
MSE	...	.00	.31	.40
Return	...	.05	.93	.01
Sharpe ratio	...	.05	.92	.01
Skewness	...	.15	.86	.00
Directional Performance <sup>b</sup>				
KS	.07 [.52]	.08 [.35]	.07 [.55]	.13 [.02]
KS*	.02 [.83]	.16 [.00]	.02 [.89]	.17 [.00]
AUC	.52 (.025)	.53 (.025)	.51 (.025)	.57** (.025)
AUC*	.44** (.025)	.60** (.024)	.46 (.025)	.59** (.024)
Observations	540	540	540	540

<sup>a</sup>Null model is the naive carry model. To correct for cross-sectional dependence, a cluster-robust covariance correction is used.

<sup>b</sup>KS and KS\* *p*-values are shown in brackets. Standard errors for AUC statistics are reported in parentheses.

\*\*Indicates significance at the 95% confidence level.

2008 using fixed-window rolling samples starting with January 1986 to December 2003. We remark that for the value strategy, FEER is also calculated using the appropriate rolling sample rather than relying on the entire sample to avoid any look-ahead advantage.

In-sample estimates appear to give credence to the well-worn practice of momentum trading since the coefficient on lagged changes of exchange rates is positive and significant. Similarly, estimates for the value strategy also confirm the wisdom that currencies eventually return to their long-run fundamental value, although the speed of reversion is relatively slow (about 1.4% per month). The more sophisticated VECM

strategy encapsulates these two observations, although the results on the speed of reversion to long-run FEER are to be interpreted differently, here with a long-run equilibrium speed of adjustment of about 2.3%.

These results are reassuring to economists, but a currency speculator prefers an assessment based on out-of-sample performance in a realistic trading setting, and this is done in table 2. Note that the results of this exercise include the period of financial turbulence that started in the fall of 2007. We deliberately chose to include this in our window since it is the occurrence of "peso events" or flight to safety during such crash episodes that helps to ensure that we are not stacking the deck in the direction of favorable trading profits by only looking at tranquil times.

In contrast to other recent work on the carry trade, we find that when the crash period is included in the sample the results of this evaluation provide quite sobering reading. Realized profits are on average negative over the out-of-sample evaluation period for the carry and value strategies. Momentum and VECM strategies enjoy low but positive returns of around 3% annually with a Sharpe ratio of about 0.5 but with relatively low and positive skew (the skewness is negative for the carry and value portfolios).

Giacomini and White (2006) statistics are reported in the middle panel of table 2. Note that the carry strategy is our null model against which all the others are compared. With a traditional MSE loss function, only the momentum strategy appears to perform significantly better (in the statistical sense) than the naive carry. However, when we switch to investment performance loss functions then it becomes clear that the VECM strategy dominates all others, although the momentum strategy is successful in generating significant improvements in terms of realized returns and Sharpe ratio, but not in terms of skewness.

The bottom panel of table 2 examines the out-of-sample performance of these trading strategies using the classification statistics that we discussed in Section III. The unweighted KS and AUC statistics suggest that the marginal classification abilities of the carry, momentum, and value signals are not significantly different than that of a coin toss, although this null is firmly rejected for the VECM strategy.

However, as we argued previously, what is important is to classify the trades with high returns/losses correctly. Here the differences are clear: carry and value attain AUC\* values below the 0.5 coin-toss benchmark (indicating that it is best to do the opposite of what the strategy proposes!), whereas the momentum and VECM strategies have a rela-

tively high AUC\* value (near 0.6) and are statistically significant. A similar picture arises if one uses KS and KS\* statistics instead.

Where do we stand at this juncture? On the one hand, the naive carry strategy appears to have been blown out of the water by the turbulent period at the end of our sample. In some sense this is reassuring, as it suggests that the persistent carry trade profits observed prior to the recent downturn are compensation for (tail) risk. On the other hand, the momentum and VECM strategies appear to have weathered the storm rather well, and so we may ask whether an even more sophisticated trader could obtain higher returns.

The next step in our analysis therefore relies on a piece of information that we have so far neglected to use—that contained in the relative term structure of government yields between countries. The next section begins by constructing Nelson and Siegel term structure factors and then extends our four trading models to examine this proposition.

### B. Relative Nelson and Siegel Yield Factors

In principle, there are many ways to measure the term structure, for example, from the parsimonious method of taking differences between 10-year and 3-month government debt, using a vector of forward rates (Clarida and Taylor 1997; Clarida et al. 2009; Ang and Chen 2010) or fitting nonparametric or spline-based curves. We choose to impose a parametric form on the yield curve that is concise and simple to implement yet flexible enough to capture the relevant shapes of yield curves. In particular, we estimate factors of the yield curve following Nelson and Siegel (1987). Because our interest lies in movements in foreign yield curves relative to that in the United States, we follow Chen and Tsang (2009) and estimate the following relative yield curve as

$$i_t^{*m} - i_t^m = L_t + S_t \left( \frac{1 - e^{-\lambda m}}{\lambda m} \right) + C_t \left( \frac{1 - e^{-\lambda m}}{\lambda m} - e^{-\lambda m} \right), \quad (6)$$

where  $i_t^{*m}$  is the return of a foreign government bond with maturity  $m$ , and  $i_t^m$  is the return of a U.S. government bond of the same maturity. The parameter  $\lambda$  controls the speed of exponential decay and here is set to 0.0609, as recommended by Nelson and Siegel (1987).

The Nelson and Siegel setup is straightforward to estimate— $L_t$ ,  $S_t$ , and  $C_t$  are estimated for each country-period pair with standard regression techniques. Summary statistics for the estimates of these factors are

**Table 3**  
Summary Statistics for Relative Nelson and Siegel Factor Estimates (Means)

Country	Level	Slope	Curve
AUS	1.98 (1.55)	1.49 (1.91)	-.48 (1.97)
CAN	-.62 (.76)	.84 (1.65)	-.53 (1.64)
CHE	-1.89 (.43)	.77 (2.35)	-.21 (1.81)
DEU	-.47 (.80)	.36 (2.09)	-.44 (3.03)
JPN	-2.83 (.78)	.58 (1.88)	-1.96 (2.48)
NOR	1.20 (1.62)	1.73 (2.48)	-.57 (2.34)
NZD	1.93 (1.92)	2.71 (2.27)	.40 (2.46)
SWE	1.20 (1.63)	.96 (2.51)	-.11 (2.53)
UK	1.09 (1.14)	1.86 (1.74)	-1.40 (2.52)

Note: Refer to the text for an explanation of how the Nelson and Siegel factors are estimated.

reported in table 3. An additional benefit of the Nelson and Siegel yield curves is that the three factors have intuitive interpretations. The level factor,  $L_t$ , has a constant impact across the entire yield curve and is closely associated with the general direction of profitable carry as foreseen at all horizons, while factor loadings for slope,  $S_t$ , and curvature,  $C_t$ , vary across the maturity spectrum of the yield curve and give an indication on future movements in naive carry. The slope factor has a loading of 1 at maturity  $m = 0$  that decreases monotonically to zero as the maturity increases. Consequently, movements at the short end of the yield curve are mostly reflected by this factor, so that, for example, conditional on a long-term yield, a higher slope factor indicates a flatter relative yield curve. The curvature factor captures movements in the middle of the yield curve—its loading is zero at maturities of zero and very long maturities, with the maximum loading in the middle of the spectrum. We should note that, to a significant extent, the short-weighted slope factor  $S$  moves in tandem with the traditional carry signal, which may explain why of all the NS factors, this one seems to add little in the way of enhancement to the trading models in what follows: in general we find that it is mainly the  $L$  and  $C$  factors that contribute incremental value as signals.

As descriptions of the term structure, Nelson and Siegel models are known to fit the data well with very high  $R^2$  values. This turns out to be the case also for the relative Nelson and Siegel factors that we calculate with average  $R^2$  values above 0.90 for each country–U.S. pair. Over the sample, the estimates on the level factor in table 3 essentially represent the average interest rate differential in the short term, for example, the almost the 3% difference between Japan and the United States that was the source of considerable interest in the carry in Japan. Table 3 also reveals the considerable variation in the slope and curvature factors across countries. It is to be expected that such variation is particularly useful to construct clever carry trade strategies. Table 4 reports regression estimates for the same forex trading strategies analyzed in table 1, only now augmented with the level, slope, and curvature Nelson and

**Table 4**

Four Carry Trade Models with Nelson and Siegel Factors: In-Sample Estimates, January 1986–December 2008

	Dependent Variable: $\Delta e_{t+1}$			
	Carry Trade Strategy			
	Carry+	Momentum+	Value+	VECM+
$\Delta e_t$	...	.10** (.03)	...	.11** (.03)
$i_t^* - i_t$	-3.00** (1.2)	...	...	-3.24** (1.11)
$\pi_t^* - \pi_t$	...	...	...	.20 (.14)
$q_t - \bar{q}$	...	...	-.018** (.003)	-.023** (.003)
Level	-.003** (.001)	-.001** (.000)	-.002** (.000)	-.004** (.001)
Slope	-.002* (.001)	.000 (.000)	.000 (.001)	-.002** (.001)
Curve	-.002** (.000)	-.002** (.000)	-.002** (.000)	-.002** (.000)
$R^2$	.018	.027	.024	.039
Currencies	9	9	9	9
Periods	276	276	276	276
Total observations	2,416	2,408	2,408	2,397

Note: Panel estimates with country fixed effects (not reported). Heteroskedasticity-robust standard errors reported in parentheses. Slight differences in the total number of observations are due to differences in data lags.

\*Indicates significance at the 90% confidence level.

\*\*Indicates significance at the 95% confidence level.

Siegel relative factors and thus denoted with a suffix “+” to differentiate them from the four benchmark strategies in table 1.

Broadly speaking, we can see that compared to the four benchmark strategies, the fit almost doubles in all cases and that level and curvature factors are significant for the four strategies (and for VECM+, all three factors enter significantly). Out-of-sample performance improves across the board as well. For example, the top panel of table 5 shows that now all four carry strategies have strictly positive returns ranging from 2.1% to 3.5% annually, with annualized Sharpe ratios around 0.6–0.7. Giacomini and White (2006) statistics reported in the middle

**Table 5**  
Four Carry Trade Models with Nelson and Siegel Factors: Out-of-Sample Performance, January 2004–December 2008

Realized Returns to an Equally Weighted Portfolio	Carry Trade Strategy			
	Carry+	Momentum+	Value+	VECM+
Mean (monthly)	.0029	.0029	.0017	.0026
SD	.014	.017	.014	.013
Skewness	-.73	.49	-.95	-.53
Coefficient of variation	4.84	5.96	8.29	5.03
Average annual return (%)	3.5	3.6	2.0	3.1
Sharpe ratio (annual)	.72	.58	.42	.69
Giacomini and White (2006) <i>p</i> -Values <sup>a</sup>				
Loss function:				
MSE	.71	.30	.44	.13
Return	.01	.02	.00	.00
Sharpe ratio	.01	.02	.00	.00
Skewness	.01	.22	.00	.00
Directional Performance <sup>b</sup>				
KS	.14 [.01]	.12 [.05]	.14 [.01]	.15 [.01]
KS*	.16 [.00]	.16 [.00]	.15 [.00]	.20 [.00]
AUC	.56** (.025)	.58** (.025)	.56** (.025)	.59** (.025)
AUC*	.55** (.025)	.60** (.024)	.55** (.025)	.60** (.024)
Observations	540	540	540	540

<sup>a</sup>The null model is the naive carry model. To correct for cross-sectional dependence, a cluster-robust covariance correction is used.

<sup>b</sup>KS and KS\* *p*-values are shown in brackets. Standard errors for AUC statistics are reported in parentheses.

\*\*Indicates significance at the 95% confidence level.

panel of table 5 suggest that although none of the four Nelson and Siegel augmented strategies would be significantly better than a naive carry trade by the MSE metric, the same is not true when the loss function is selected on the basis of investment performance of realized returns. Moreover, the Nelson and Siegel factors now make all four strategies significantly better over all, regardless of the loss function (returns, Sharpe, or skewness).

A natural explanation of these results is that to secure good returns all that is needed is a good directional forecast, and the bottom panel of table 5 confirms this observation. The KS and KS\* statistics improve uniformly for all strategies; a similar picture arises from the AUC and AUC\* statistics. Overall, looking across these metrics, the VECM+ still stands out as a preferred strategy although the differences across models have narrowed. One intuition for this finding is that the shape of the forward curve may capture some of the same composite information embedded in the CMV factors; for example, future yields may be informative concerning the forward-looking path of an exchange rate toward its FEER value. Our results are consistent with Ang and Chen (2010), who find that there is significant information in the yield curve to predict excess foreign exchange returns. Clarida et al. (2009) find that this correlation is robust to inclusion of the recent financial crash.

## V. Do Risk Factors Explain Carry Trade Profits?

The carry trade is a zero-net investment strategy. If traders had perfect foresight, then we would expect the carry trade to have zero-net returns (abstracting from transaction costs). Traders do not have perfect foresight, and markets have frictions so that average, non-zero-net returns are not necessarily surprising—they could be justified as compensation for risk in the sense that the carry trade provides an inadequate hedge against so-called risk factors that proxy for the stochastic discount factor in a typical asset-pricing model.

The question that we explore in this section is whether the carry trade strategies that we have analyzed are correlated with any of a long list of so-called risk factors that have been extensively analyzed elsewhere in the literature (e.g., Burnside et al. 2008a, 2008b). If they are, then we are interested in determining whether the “risk-adjusted” returns are still positive. This so-called alpha is calculated as the intercept from a regression of realized return on the demeaned risk factor variable(s). If alpha remains positive, and/or if the slope coefficients on the risk factors are

not significant, then it makes it difficult to explain what could justify the sort of returns that we have reported here.

The risk factors that we explore are all from the perspective of the United States and include the excess return to the value-weighted U.S. stock market (CAPM); the three Fama and French (1993) factors, namely, excess return to value-weighted U.S. stock market, the size premium, and the value premium; U.S. industrial production growth; the federal funds rate; the term premium measured as the spread between the 10-year Treasury Bond and the 3-month Treasury Bill; the Pastor and Stambaugh (2003) two liquidity measures; and four measures of market volatility, that is, the Chicago Board Options Exchange Volatility indexes VIX and VXO as well as their differences. An appendix to this paper contains a detailed description of how each risk factor was constructed.

Formally, we proceeded as follows. Using out-of-sample realized returns from the four benchmark portfolios and the Nelson and Siegel augmented versions, we then constructed equally weighted (across countries) portfolio returns, one for each strategy (for a total of eight cases). We then regressed these portfolio returns against each of the factors listed above, one at a time (except for the three Fama and French factors, which are entered jointly).

Specifically, portfolio returns are constructed as

$$\hat{\mu}_{t+1}^{EW} = \frac{1}{J} \sum_{j=1}^J \hat{\mu}_{t+1}^j,$$

where  $j \in \{\text{Australia, Canada, Germany, Japan, Norway, New Zealand, Sweden, Switzerland, United Kingdom}\}$  so that  $J = 9$ . Next, we regress

$$\hat{\mu}_{t+1}^{EW} = \alpha_k + \beta_k f_{k,t+1} + u_{k,t+1} \tag{7}$$

for  $k = 1, \dots, K$ , where  $f_{k,t+1}$  denotes the  $k$ th risk factor out of the  $K$  risk factors listed above and where the sample begins in January 2004 and ends in December 2008, as in tables 2 and 5. Therefore, the regressions in (7) are essentially those reported in tables 3 and 4 in Burnside, Eichenbaum, and Rebelo (2008), except that we focus on out-of-sample realized returns rather than on in-sample results and we use monthly rather than quarterly data.

We again remind the reader that our sample includes the turbulent period that begins in early 2007 and ends in our sample in December 2008. This period saw a major crash of G-10 carry trades with significant



appreciation of funding currencies such as the JPY and CHF and major declines in high-yield targets such as AUD, NZD, and GBP. There was also a sharp concentration of risk as evinced by the high observed correlation of risk factors during this period, with observers noting the unusual and unprecedented comovement of risk assets driven by daily risk-on/risk-off shifts in market sentiment (e.g., the almost overnight emergence of a strong correlation between JPYUSD and SPY right after the “Shanghai Surprise” event in 2007; Authers 2010).

The regression estimates are reported in tables 6–9. An estimate for  $\alpha_k$  corresponds to the risk-adjusted return to that particular carry trade

**Table 6**

Risk Factor Regressions: Out-of-Sample Risk-Adjusted Returns of Equally Weighted Portfolios Using the Carry Strategy, January 2004–December 2008

Factor	Carry				Carry+			
	Intercept		Betas		Intercept		Betas	
CAPM	-.0017 (.0018)	.0029** (.0009)			.0031 (.0020)	.0007* (.0004)		
Fama and French	-.0015 (.0017)	.0032** (.0010)	-.0014 (.0010)	.0014 (.0010)	.0033* (.0018)	.0012** (.0004)	-.0017* (.0009)	-.0012 (.0009)
IP growth	-.0022 (.0027)	.3097 (.3175)			.0032* (.0018)	.5068** (.2335)		
Federal funds rate	-.0016 (.0020)	.0016 (.0020)			.0036* (.0020)	.0014 (.0011)		
Term premium	-.0028 (.0030)	-.0026 (.0028)			.0026 (.0021)	-.0019 (.0014)		
Liquidity premium	-.0015 (.0017)	-.0146** (.0043)			.0031 (.0020)	-.0032 (.0021)		
Pastor and Stambaugh liquidity measures:								
Level	-.0022 (.0025)	.0747* (.0433)			.0030 (.0020)	.0211 (.0166)		
Innovation	-.0024 (.0027)	.0482 (.0356)			.0029 (.0020)	.0247 (.0289)		
Market volatility:								
VIX	-.0031 (.0019)	-.0013** (.0004)			.0027 (.0020)	-.0003 (.0002)		
VXO	-.0036* (.0019)	-.0013** (.0004)			.0026 (.0021)	-.0002 (.0002)		
$\Delta$ VIX	-.0022 (.0022)	-.0021** (.0009)			.0029 (.0020)	-.0007** (.0003)		
$\Delta$ VXO	-.0022 (.0022)	-.0018** (.0008)			.0030 (.0020)	-.0006* (.0003)		

Note: See text for details on risk factors and risk factor regressions. Standard errors robust to autocorrelation and heteroskedasticity are in parentheses.

\*Indicates significance at the 90% confidence level.

\*\*Indicates significance at the 95% confidence level.

**Table 7**  
 Risk Factor Regressions: Out-of-Sample Risk-Adjusted Returns of Equally Weighted Portfolios Using the Momentum Strategy, January 2004–December 2008

Factor	Momentum				Momentum+			
	Intercept		Betas		Intercept		Betas	
CAPM	.0023 (.0024)	-.0016 (.0012)			.0026 (.0023)	-.0011 (.0010)		
Fama and French	.0020 (.0022)	-.0020 (.0012)	.0018 (.0013)	-.0013 (.0012)	.0026 (.0022)	-.0010 (.0009)	.0000 (.0013)	-.0021* (.0013)
IP growth	.0027 (.0027)	-.0428 (.3239)			.0033 (.0024)	.5855** (.1843)		
Federal funds rate	.0013 (.0020)	-.0024 (.0019)			.0028 (.0020)	-.0002 (.0017)		
Term premium	.0031 (.0028)	.0033 (.0025)			.0030 (.0026)	.0004 (.0023)		
Liquidity premium	.0021 (.0022)	.0091* (.0052)			.0027 (.0022)	.0004 (.0053)		
Pastor and Stambaugh liquidity measures:								
Level	.0027 (.0026)	-.0026 (.0453)			.0030 (.0023)	.0157 (.0400)		
Innovation	.0027 (.0027)	.0124 (.0452)			.0030 (.0024)	.0406 (.0316)		
Market volatility:								
VIX	.0031 (.0024)	.0009** (.0004)			.0031 (.0024)	.0004 (.0004)		
VXO	.0035 (.0024)	.0008** (.0004)			.0033 (.0025)	.0004 (.0004)		
$\Delta$ VIX	.0025 (.0023)	.0016** (.0006)			.0028 (.0023)	.0011** (.0005)		
$\Delta$ VXO	.0025 (.0023)	.0015** (.0006)			.0028 (.0023)	.0010* (.0005)		

Note: See text for details on risk factors and risk factor regressions. Standard errors robust to autocorrelation and heteroskedasticity are in parentheses.

\*Indicates significance at the 90% confidence level.

\*\*Indicates significance at the 95% confidence level.

strategy relative to the risk factor considered in that regression. Therefore, we are looking for cases in which estimates of  $\beta_k$  enter significantly, in which case we need to check if  $\alpha_k \rightarrow 0$ , thus suggesting that excess carry trade returns can be explained as compensation for the risk described by the risk factor considered.

The results of this exercise can be broadly summarized as follows. The four benchmark portfolios appear to be consistently correlated with measures of market volatility (VIX, VXO, and their first differences). This result is striking. But for the simple carry strategy it is con-

**Table 8**  
Risk Factor Regressions: Out-of-Sample Risk-Adjusted Returns of Equally Weighted Portfolios Using the Value Strategy, January 2004–December 2008

Factor	Value		Value+	
	Intercept	Betas	Intercept	Betas
CAPM	-.0018 (.0020)	.0014** (.0005)	.0018 (.0019)	.0006 (.0004)
Fama and French	-.0017 (.0020)	.0015** (.0006)	-.0003 (.0010)	.0008 (.0010)
IP growth	-.0019 (.0021)	.3755* (.2355)	.0020 (.0018)	.5231** (.2029)
Federal funds rate	-.0020 (.0021)	.0002 (.0016)	.0024 (.0019)	.0013 (.0010)
Term premium	-.0022 (.0024)	-.0004 (.0020)	.0014 (.0020)	-.0018 (.0013)
Liquidity premium	-.0016 (.0018)	-.0087** (.0019)	.0019 (.0019)	-.0036** (.0018)
Pastor and Stambaugh liquidity measures:				
Level	-.0020 (.0021)	.0455** (.0228)	.0017 (.0020)	.0120 (.0156)
Innovation	-.0021 (.0022)	.0269 (.0237)	.0017 (.0020)	.0159 (.0290)
Market volatility:				
VIX	-.0026** (.0018)	-.0008** (.0001)	.0015 (.0019)	-.0003** (.0001)
VXO	-.0028** (.0018)	-.0007** (.0001)	.0013 (.0019)	-.0003** (.0001)
ΔVIX	-.0021 (.0022)	-.0008 (.0006)	.0017 (.0019)	-.0007** (.0003)
ΔVXO	-.0021 (.0022)	-.0006 (.0006)	.0017 (.0019)	-.0007** (.0003)

Note: See text for details on risk factors and risk factor regressions. Standard errors robust to autocorrelation and heteroskedasticity are in parentheses.

\*Indicates significance at the 90% confidence level.

\*\*Indicates significance at the 95% confidence level.

sistent with Brunnermeier et al. (2008), who argue that during volatile periods, traders prefer to liquidate carry trade positions to generate a cash cushion against domestic market instability. Even so, this result warrants a few caveats.

First, Jordà and Taylor (2009) find that such risk-factor correlations are not present in the pre-crash sample period actually studied by Brunnermeier et al. (2008) once one employs trading strategies that include the CVM factors; thus, it may be that the extreme events of the 2008 crash period indicate that it is only in very extreme “crisis” events

**Table 9**  
 Risk Factor Regressions: Out-of-Sample Risk-Adjusted Returns of Equally Weighted Portfolios Using the VECM Strategy, January 2004–December 2008

Factor	VECM				VECM+			
	Intercept		Betas		Intercept		Betas	
CAPM	.0020 (.0022)	-.0019 (.0012)			.0025 (.0017)	-.0005 (.0004)		
Fama and French	.0016 (.0021)	-.0025** (.0011)	.0027** (.0011)	-.0017 (.0013)	.0024 (.0017)	-.0004 (.0003)	-.0003 (.0009)	-.0015 (.0010)
IP growth	.0024 (.0027)	-.1011 (.2417)			.0026 (.0017)	.0011 (.1504)		
Federal funds rate	.0030 (.0021)	-.0039** (.0019)			.0026 (.0016)	.0000 (.0012)		
Term premium	.0031 (.0028)	.0044 (.0027)			.0025 (.0018)	-.0003 (.0016)		
Liquidity premium	.0019 (.0022)	.0098* (.0053)			.0025 (.0017)	.0021 (.0022)		
Pastor and Stambaugh liquidity measures:								
Level	.0023 (.0026)	-.0310 (.0392)			.0026 (.0017)	.0180 (.0186)		
Innovation	.0024 (.0027)	-.0194 (.0377)			.0026 (.0017)	.0179 (.0269)		
Market volatility:								
VIX	.0030 (.0022)	.0010** (.0004)			.0027 (.0017)	.0003 (.0002)		
VXO	.0033 (.0023)	.0009** (.0003)			.0028 (.0017)	.0002 (.0002)		
ΔVIX	.0023 (.0025)	.0012 (.0010)			.0026 (.0017)	.0001 (.0002)		
ΔVXO	.0024 (.0025)	.0010 (.0009)			.0026 (.0017)	.0002 (.0002)		

Note: See text for details on risk factors and risk factor regressions. Standard errors robust to autocorrelation and heteroskedasticity are in parentheses.

\*Indicates significance at the 90% confidence level.

\*\*Indicates significance at the 95% confidence level.

that such risk factors play a major role and can overwhelm even the more sophisticated hedging techniques built in to augmented carry models.

Second, in tables 6–9, the correlation with the volatility measures serves only to eliminate alpha for the carry and value strategies; for momentum and VECM it serves to enhance alpha; thus, the returns for these strategies still cannot be explained away. Moreover, the risk-factor explanation for returns completely evaporates for the momentum+ and VECM+ strategies (and to a large extent for the carry+ and value+ strategies as well). Out of the former two, momentum+ appears to have only some mild covariation with industrial production growth, whereas

VECM+ appears to be uncorrelated with all the risk factors that we consider and in fact, the risk-adjusted returns shown in table 9 (right panel alphas) are relatively constant and equal to the raw returns seen in table 5 (3% per annum with a Sharpe ratio of 0.69). Finally, although the alphas in table 9 have wide standard errors, the point estimate is consistent with our no-risk-factor model, and the sample size is small, and we consider that our Giacomini and White and Correct Classification statistics (AUC and KS) provide the compelling out-of-sample evidence of unweighted and return-weighted predictive ability against the null.

Thus our VECM+ (a CMV vector model augmented by information from the forward curve) appears to stand tallest among all the augmented carry models we have considered. It significantly beat a coin toss as a directional forecast, its return-weighted performance also delivers statistically significant profits, its profits avoid negative skew, and the profits have no correlation with a standard set of risk factors.

## VI. Conclusion

This paper has presented new data and methods to explore an ongoing debate about the currency carry trade. We find that many widely used carry trade strategies failed during the recent financial crisis, but those augmented by additional hedging signals have fared better. Building on prior work that identified the carry, momentum, and value (CMV) signals as jointly important (Jordà and Taylor 2009), in this paper we also find a complementary role for information drawn from the forward yield curves (Clarida and Taylor 1997; Ang and Chen 2009; Clarida et al. 2009). When this full set of signals is employed, the resulting portfolios of trades are profitable, exhibit attractive Sharpe and skewness properties, and cannot be rationalized away using any standard risk factors.

## Appendix

The list of sources we used to construct risk factors is as follows:

*Fama and French factors:* Kenneth French data library ([http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html))

*Industrial production growth:* Federal Reserve Board release G.17

*Federal funds rate:* Federal Reserve Board, table H.15

*Term premium:* Federal Reserve Board, table H.15; 10-year T-Bill (constant maturity) less 3-month T-Bill (secondary market rate)

*Liquidity premium*: Federal Reserve Board, table H.15; 3-month euro-dollar rate less 3-month T-Bill

*Pastor and Stambaugh*: Pastor Web site ([http://faculty.chicagobooth.edu/lubos.pastor/research/liq\\_data\\_1962\\_2008.txt](http://faculty.chicagobooth.edu/lubos.pastor/research/liq_data_1962_2008.txt))

VIX: Yahoo! Finance

VXO: Yahoo! Finance

## Endnotes

Paper presented at the NBER International Seminar on Macroeconomics, Amsterdam, June 2010. Taylor has been supported by the Center for the Evolution of the Global Economy at University of California, Davis, and Jordà by DGICYT grant (SEJ2007-63098-econ); part of this work was completed while Taylor was a Houblon-Norman/George Fellow at the Bank of England; all of this research support is gratefully acknowledged. We thank Craig Burnside for sharing his data on risk factors with us. We acknowledge helpful comments from our discussants Robert Cumby and Mark Taylor, and from participants at the ISOM meeting, and we thank De Nederlandsche Bank for their efficient organization and warm hospitality. All errors are ours.

1. For a full survey of foreign exchange market efficiency, see Sarno and Taylor (2002, chap. 2).

2. The critique is due to Summers (1985), who downplayed the usefulness of approaches which show that the prices of all risky assets move up and down in unison—like the prices of different sized bottles of tomato ketchup—without offering either theoretical insights or empirical evidence as to the fundamental shocks behind such comovements.

3. We thank Craig Burnside for sharing the quarterly version for most of these data. We constructed (from sources described in the appendix) monthly frequency data, updated for the sample that we examine.

## References

- Ang, Andrew, and Joseph S. Chen. 2010. "Yield Curve Predictors of Foreign Exchange Returns." Photocopy, University of California, Davis.
- Authers, John. 2010. "Book Extract: The Fearful Rise of Markets." *Financial Times*, May 21.
- Bacchetta, Philip, and Eric van Wincoop. 2006. "Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?" *American Economic Review* 96, no. 3:552–76.
- Bekaert, Geert, and Robert J. Hodrick. 1993. "On Biases in the Measurement of Foreign Exchange Risk Premiums." *Journal of International Money and Finance* 12, no. 2:115–38.
- Brunnermeier, Markus K., Stefan Nagel, and Lasse H. Pedersen. 2008. "Carry Trades and Currency Crashes." Working Paper no. 14473, NBER, Cambridge, MA.
- Burnside, A. Craig, Martin Eichenbaum, Isaac Kleshchelski, and Sergio Rebelo. 2006. "The Returns to Currency Speculation." Working Paper no. 12489, NBER, Cambridge, MA.
- . 2008a. "Do Peso Problems Explain the Returns to the Carry Trade?" CEPR Discussion Paper no. 6873, Centre for Economic Policy Research, London.

- . 2008b. "Do Peso Problems Explain the Returns to the Carry Trade?" Working Paper no. 14054, NBER, Cambridge, MA.
- Burnside, A. Craig, Martin Eichenbaum, and Sergio Rebelo. 2008. "Carry Trade: The Gains of Diversification." *Journal of the European Economic Association* 6, nos. 2–3:581–88.
- Chen, Yu-chin, and Kwok Ping Tsang. 2009. "What Does the Yield Curve Tell Us about Exchange Rate Predictability?" Working Paper UWEC 2009-04, University of Washington, Department of Economics.
- Cheung, Yin-Wong, Menzie D. Chinn, and Antonio Garcia Pascual. 2005. "Empirical Exchange Rate Models of the Nineties: Are Any Fit to Survive?" *Journal of International Money and Finance* 24, no. 7:1150–75.
- Clarida, Richard H., Josh Davis, and Niels Pedersen. 2009. "Currency Carry Trade Regimes: Beyond the Fama Regression." *Journal of International Money and Finance* 28, no. 8:1375–89.
- Clarida, Richard H., and Mark P. Taylor. 1997. "The Term Structure of Forward Exchange Premiums and the Forecastability of Spot Exchange Rates: Correcting the Errors." *Review of Economics and Statistics* 79, no. 3:353–61.
- Clarida, Richard H., and Daniel Waldman. 2007. "Is Bad News about Inflation Good News for the Exchange Rate?" Working Paper no. 13010, NBER, Cambridge, MA.
- Cochrane, John H. 2001. *Asset Pricing*. Princeton, NJ: Princeton University Press.
- Conover, W. J. 1999. *Practical Nonparametric Statistics*. 3rd ed. New York: Wiley.
- Diebold, Francis X., and Roberto S. Mariano. 1995. "Comparing Predictive Accuracy." *Journal of Business and Economic Statistics* 13, no. 3:253–63.
- Engel, Charles. 1996. "The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence." *Journal of Empirical Finance* 3:123–91.
- Fama, Eugene F. 1984. "Forward and Spot Exchange Rates." *Journal of Monetary Economics* 14, no. 3:319–38.
- Fama, Eugene F., and Kenneth R. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33:3–56.
- Fisher, Eric O'N. 2006. "The Forward Premium in a Model with Heterogeneous Prior Beliefs." *Journal of International Money and Finance* 25, no. 1:48–70.
- Frankel, Jeffrey A. 1980. "Tests of Rational Expectations in the Forward Exchange Market." *Southern Economic Journal* 46, no. 4:1083–1101.
- Froot, Kenneth A., and Richard H. Thaler. 1990. "Foreign Exchange." *Journal of Economic Perspectives* 4, no. 3:179–92.
- Giacomini, Raffaella, and Halbert White. 2006. "Tests of Conditional Predictive Ability." *Econometrica* 74, no. 6:1545–78.
- Hodrick, Robert J. 1987. *The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Markets*. Chur: Harwood.
- Hsieh, Fushing, and Bruce W. Turnbull. 1996. "Nonparametric and Semiparametric Estimation of the Receiver Operating Characteristics Curve." *Annals of Statistics* 24:25–40.
- Ilut, Cosmin. 2008. "Ambiguity Aversion: Implications for the Uncovered Interest Rate Parity Puzzle." Photocopy, Northwestern University.
- Jeanne, Olivier, and Andrew K. Rose. 2002. "Noise Trading and Exchange Rate Regimes." *Quarterly Journal of Economics* 117, no. 2:537–69.
- Jordà, Óscar, and Alan M. Taylor. 2009. "The Carry Trade and Fundamentals: Nothing to Fear but FEER Itself." Working Paper no. 15518, NBER, Cambridge, MA.
- . 2011. "Performance Evaluation for Zero Net-Investment Strategies." Working Paper no. 17150, NBER, Cambridge, MA.

- Jurek, Jakub W. 2008. "Crash-Neutral Currency Carry Trades." Photocopy, Princeton University.
- Kilian, Lutz, and Mark P. Taylor. 2003. "Why Is It so Difficult to Beat the Random Walk Forecast of Exchange Rates?" *Journal of International Economics* 60, no. 1:85–107.
- Meese, Richard A., and Kenneth Rogoff. 1983. "Empirical Exchange Rate Models of the Seventies." *Journal of International Economics* 14, nos. 1–2:3–24.
- Melvin, Michael, and Mark P. Taylor. 2009. "The Crisis in the Foreign Exchange Market." *Journal of International Money and Finance* 28, no. 8:1317–30.
- Nelson, Charles R., and Andrew F. Siegel. 1987. "Parsimonious Modeling of Yield Curves." *Journal of Business* 60, no. 4:473–89.
- Obuchowski, Nancy A. 1994. "Computing Sample Size for Receiver Operating Characteristic Curve Studies." *Investigative Radiology* 29, no. 2:238–43.
- Pastor, Lubos, and Robert F. Stambaugh. 2003. "Liquidity Risk and Expected Stock Returns." *Journal of Political Economy* 111, no. 3:642–85.
- Peirce, Charles S. 1884. "The Numerical Measure of the Success of Predictions." *Science* 4:453–54.
- Sarno, Lucio, and Mark P. Taylor. 2002. *The Economics of Exchange Rates*. Cambridge: Cambridge University Press.
- Shleifer, Andrei, and Robert W. Vishny. 1997. "A Survey of Corporate Governance." *Journal of Finance* 52, no. 2:737–83.
- Summers, Lawrence H. 1985. "On Economics and Finance." *Journal of Finance* 40, no. 3:633–35.
- Youden, W. J. 1950. "Index for Rating Diagnostic Tests." *Cancer* 3:32–35.



