Journal of International Money and Finance 85 (2018) 1–19

Contents lists available at ScienceDirect

Journal of International Money and Finance

journal homepage: www.elsevier.com/locate/jimf

Conditioning carry trades: Less risk, more return ☆

Arjen Mulder *, Ben Tims

Rotterdam School of Management, Erasmus University, Burg. Oudlaan 50, 3062PA Rotterdam, The Netherlands

ARTICLE INFO

Article history:
Available online 8 March 2018

2016 MSC:
D-16-00220-R2

Keywords:
Uncovered interest parity
Extreme sampling
Carry trade
Foreign exchange rates

ABSTRACT

Prior studies show that extreme interest rate differentials (IRDs) and high foreign exchange rate (FX) volatility have substantial explanatory power for the validity of UIP. We show that these contemporaneous drivers also have predictive power by implementing a conditional currency carry trade (CT) strategy that excludes regimes for which UIP is likely to hold. Conditioning high FX volatility only, or on both FX volatility and extreme IRDs outperforms the base-case unconditional CT strategy in virtually any of the settings analyzed. Conditioning on very large IRDs only shows mixed findings. Our strategy works best for smaller CT portfolios.

1. Introduction

In this paper we investigate currency carry trade (CT) strategies, that aim at exploiting exchange rate mispricing by taking short positions in low-interest rate currencies and long positions in high-interest rate ones (e.g. Koijen et al., 2018).[1] According to a standard international finance theory, the so-called 'uncovered interest rate parity' (UIP), these investments should yield a zero profit, because the expected return of the currency exchange rate should exactly offset the difference in the underlying interest rates. Empirically, however, these CT investments tend to yield a positive return over longer time horizons (Christiansen et al., 2011; Jordà and Taylor, 2012; Villanueva, 2007). As these gains of CT strategies are small but positive most of the time but are occasionally wiped away due to adverse movements in the exchange rate, *The Economist* has coined carry trades as 'picking up nickels in front of steamrollers' (The Economist, 2006).

Although according to the academic literature there is much evidence that UIP does not hold *in general* (e.g. Engel, 1996; Hodrick, 1987; Lewis, 1995; Sarno, 2005), which explains the profitability of CT strategies over longer windows, UIP might be valid in *specific* regimes. In this paper we focus on two such regimes that have a contemporaneous effect on the validity of UIP, namely when interest rate differentials (IRDs) are large and foreign exchange (FX) volatility is high. We show that these regimes also have predictive power such that exploiting them in a conditional currency carry trade strategy could improve its performance.

☆ We gratefully acknowledge Kees Koedijk (the editor), and an anonymous reviewer, Dion Bongaerts, Mathijs Cosemans, Mathijs van Dijk, Bilal Hafeez, Joop Huij, Hao Jiang, Abe de Jong, Richard Levich, Ronald Mahieu, Martin Martens, Spyros Mesomeris, Nick Roussanov, Elvira Sojli, Laurens Swinkels, Remco Zwinkels, and seminar participants at Erasmus University Rotterdam, Free University of Amsterdam, Deutsche Bank, and Robeco Asset Management for helpful comments on and suggestions for earlier versions of this paper.

* Corresponding author.
E-mail addresses: amulder@rsm.nl (A. Mulder), btims@rsm.nl (B. Tims).

[1] While there is widespread agreement that the currency carry trade (CT) market has rapidly grown over the past years, the exact number of billions of dollars traded per day in the CT market can hardly be determined with publicly available data (Curcuru et al., 2010; Galati et al., 2007).

https://doi.org/10.1016/j.jimonfin.2018.03.003
0261-5606/

We propose three conditional CT strategies. Our main conditional CT strategy, that we call the CONDFXIRD strategy, conditions on so-called 'extreme' IRDs during periods of high FX volatility by disregarding the corresponding investment opportunities. Using a sample of 25 countries over the period 1985-mid 2015 we show that this conditional CT strategy outperforms the conventional unconditional CT strategy (base-case) in terms of mean return, holding period return, Sharpe ratio, and skewness.[2] Our robustness checks include alternative extreme IRD cutoff rates, alternative definitions of extreme IRDs, various FX volatility thresholds, alternative definitions of FX volatility, VIX as an alternative regime switch indicator, the impact of the Lehman crisis, the number of long and short positions in the CT strategies, and the choice of currencies in the investment opportunity set. We show that our findings are robust to all these settings. We also show that the number of positions is important: our main conditional CT strategy works best for smaller CT portfolios, i.e. for portfolios with two long-two short positions or one long-one short positions.

Our second and third conditional CT strategies test whether both conditions (i.e. extreme IRDs during high FX volatility) have to be met in order to outperform the base-case. In our second conditional CT strategy (that we call CONDFX) we condition on FX volatility only, which in our setup implies that in high FX volatility regimes we switch to a passive investment strategy. This strategy has the best performance in terms of average return, Sharpe ratio, holding period return, and skewness. Yet, this CONDFX trategy does not consistently outperform the base-case throughout all the robustness checks. In some settings this strategy outperforms the base-case on some performance indicators because of its dramatically lower standard deviation of returns (due to periods where no investments are made), which boosts its Sharpe ratio in spite of its low(er) returns. In our third conditional CT strategy (that we call CONDIRD) we condition on extreme IRDs only. That is, the highest IRDs are truncated from the investment opportunity set regardless of the FX volatility regime. This strategy shows mixed performance relative to the base-case across all settings. Fig. 1 summarizes our main findings.

Our motivation to condition on both extreme IRDs and FX volatility regimes is based on prior studies that have extensively researched the role of FX volatility and extreme IRDs as key drivers of currency risk premia. A key finding of the FX volatility literature is that in periods of high FX volatility the returns of the carry trade strategy is low or even negative whilst periods of low FX volatility deliver positive CT returns (Christiansen et al., 2011; Clarida et al., 2009; Lustig et al., 2011; Menkhoff et al., 2012). Furthermore, the so-called extreme sampling literature shows that UIP is more likely to hold for extreme IRD observations (cf. Huisman et al., 1998; Lothian and Wu, 2011) which seems to contradict the logic of conventional CT strategies because CT strategies try to obtain their risk-arbitraging profits exactly from these extreme IRD positions. Bansal and Dahlquist (2000) show that for developed markets' currencies, UIP is particularly violated if US interest rates exceed foreign rates. Lastly, Baillie and Chang (2011) combine the findings of both strands of literature (FX volatility and extreme IRDs) and show that UIP is likely to hold for positive extreme IRDs during high FX volatility periods.[3] Therefore we expect that these regimes do not positively contribute to the performance of a CT strategy. As a consequence, we expect that truncating these extreme IRDs during periods of high FX volatility from the investment opportunity set should enhance CT performance.

A possible explanation why UIP is likely to hold for extreme IRDs during high FX volatility periods is provided by the 'limits to speculation hypothesis'. Under this hypothesis transaction and opportunity costs might prevent traders from exploiting currency mispricing (Hochradl and Wagner, 2010; Lyons, 2001; Sarno, 2005). Furthermore, these costs are idiosyncratic (and thus heterogeneous across investors) and time-varying. If the interest rate differentials (IRDs) are small then only a small number of traders are able to exploit the mispricing. Yet, the larger the IRD the more likely that it overcomes the transaction and opportunity costs for other investors as well. Therefore, increases in IRDs will engage an increasing number of traders to risk-arbitrage away the mispricing (see, e.g. Baldwin, 1990) and consequently trade volumes as measured by order flows will increase, which in turn leads to higher FX volatility (Evans, 2010; Evans and Lyons, 2002a,b; Sager and Taylor, 2008). In sum, under the limits to speculation hypothesis, UIP is thus likely to hold for large IRDs during periods of high FX volatility, and these investment opportunities will thus not contribute to the performance of a CT strategy.[4]

Our findings have the following important implications. First, conditioning a CT strategy on regimes where UIP is likely to hold can outperform the base-case conventional strategy. These gains are substantial, while the strategy is simple to implement, and does not come at a real cost. Second, the individual performances of our conditional CT strategies vary. When considering average and holding period returns and Sharpe ratios, our CONDFX CT strategy performs best. The CONDFXIRD strategy (which conditions on both FX volatility and extreme interest rate differentials) has a slightly lower outperformance of the conventional CT strategy, but its outperformance is much more robust across all robustness tests. Though our results can be improved even more by means of an exact parametrization of our FX volatility indicator and 'extreme' IRD definition, the purpose of our paper is not to calibrate these indicators to obtain the best-possible outperformances (for one data set). Instead, we show that conditioning carry trades on volatility and extreme IRDs improve their performance.

This paper has the following contributions to the growing literature on currency carry trades. First, we show that conditioning on regimes where UIP is likely to hold always leads to outperformance in terms of average return, holding period returns, Sharpe ratio, and crash risk (i.e., skewness). Though the importance of extreme IRDs and FX volatility have been

[2] Our main analysis focuses on CT portfolios with 2 long and 2 short positions, that rebalance monthly. See Section 2 for the full specifications.

[3] Although Baillie and Chang (2011) show that UIP is likely to hold for *positive* extreme IRDs during high FX volatility periods, we follow the extreme sampling literature by using *absolute* extreme IRDs in our main analysis. As robustness check we will separately test the effect of positive and negative extreme IRDs on CT performance.

[4] Since transaction and opportunity costs vary across traders and across time, pinpointing the exact threshold of UIP regimes is not possible. Therefore observing higher FX volatility and higher IRDs can only imply an increased *likelihood* that UIP will hold.

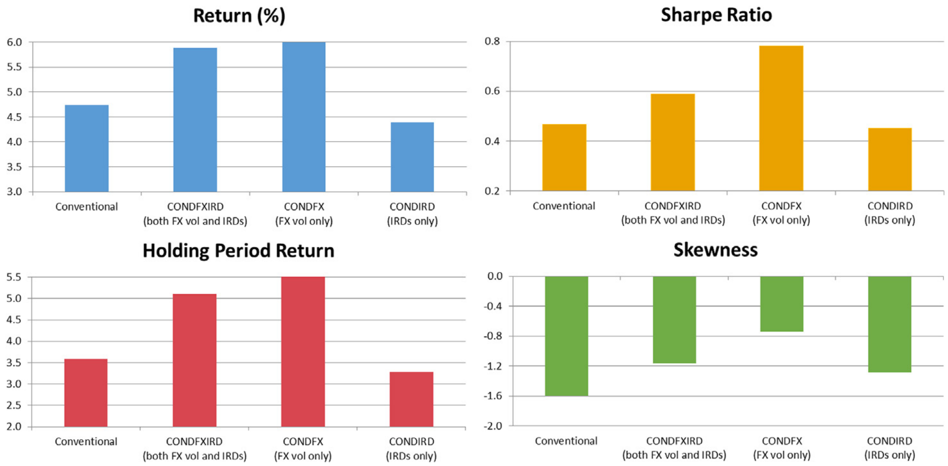

Fig. 1. Performance. Fig. 1 shows the performance measured by the annualized mean return (in %), annualized standard deviation (in %), holding period return (end-of-period accumulated wealth of $1 investment), and Sharpe ratio for the conventional CT strategy (base-case) and our three conditional CT strategies. The conventional strategy is an unconditional CT strategy based on 2 long positions in the currencies with the highest interest rates and 2 short positions in the currencies with the lowest interest rates. The second strategy, which conditions on extreme IRDs during high FX volatility regimes (CONDFXIRD), truncates the investment opportunity set on an ex ante basis by ignoring the 10% extreme IRDs during periods when FX volatility is higher than the 90%-th quantile of the FX volatility distribution. A 5 year rolling window is used to determine the IRD distribution and the FX volatility distribution. The third strategy conditions on FX volatility only (CONDFX), and switches to the conventional CT strategy during low FX volatility regimes and to a passive (non-investment) strategy during high FX volatility regimes. The fourth strategy conditions on the 10% extreme IRDs only (CONDIRD), thereby ignoring the FX volatility regime. All strategies rebalance every month. The monthly data set covers 25 currencies spanning the 1985 to mid-2015 period.

widely researched and recognized, it is by no means obvious that these contemporaneous effects also have predictive power on the validity of UIP.

Second, to the best of our knowledge FX volatility and extreme IRDs have yet not been incorporated in the design of conditional CT strategies. Although prior studies have linked CT strategy performance to various determinants (Brunnermeier et al., 2009; Christiansen et al., 2011; Lustig and Verdelhan, 2007), very few propose *conditional* CT strategies that adjust conventional CT strategies in order to improve performance.[5] Our findings thus warn against the apparent attractiveness of extreme IRD investments during high FX volatility; truncating these opportunities from our investment set boosts the performance of our main conditional CT strategy. This is because of their poor reward-to-risk ratio.

Third, consistent outperformance is particularly obtained by jointly conditioning on extreme IRDs and high FX volatility. Conditioning on only one dimension will not always help to consistently outperform the base-case. Conditioning on extreme IRDs only (i.e., our CONDIRD strategy) therefore implies that we forego profitable investment opportunities during low volatility periods. Conditioning on FX volatility only (i.e., our CONDFX strategy) also implies that we sometimes forego profitable investment opportunities. This time, however, large but not extreme IRDs, that do add value to the portfolio, would also be ignored.

Combining the above contributions, our paper provides additional support for the limits to speculation hypothesis.

2. Conventional and conditional CT strategies

The standard practice of currency carry trade investments is to borrow money in one or multiple currencies with the lowest interest rates, and invest the proceeds in the same number of currencies with the highest interest rates. This CT strategy can only be profitable if UIP does not hold, so logically investors should not invest in those regimes for which UIP holds. According to the limits to speculation hypothesis UIP is likely to hold for extreme IRDs during high FX volatility periods, therefore we conjecture that a conditional CT strategy that exclude these regimes from the investment opportunity set will outperform the unconditional conventional CT strategy.

In this section we will first discuss how the conventional CT strategy (base-case) is implemented. Subsequently, we discuss how extreme IRDs and high FX volatility are determined before we turn to the implementation of our conditional CT strategies where we condition on both large IRDs and high FX volatility and on large IRDs and high FX volatility separately.

2.1. Conventional CT strategy

In this paper the conventional CT strategy is constructed as follows (cf. e.g. Brunnermeier et al., 2009; Clarida et al., 2009). An investor invests in the two currencies with the highest interest rates (the 2 long positions) and borrows in the two cur-

[5] An example of a paper using a conditional CT strategy is Jordà and Taylor (2012), who avoid crash risk by controlling for the Fundamental Equilibrium Exchange Rate (FEER).

rencies with the lowest interest rates (the 2 short positions).[6] The return of a single long position at time t in currency h vis-à-vis our benchmark currency, the US dollar, equals:

$$r_t^h = i_{t-1}^h - i_{t-1}^{US} - \Delta s_t^h \equiv IRD_{t-1}^h - \Delta s_t^h, \tag{1}$$

where i_t^h is the interest rate for currency h, i_t^{US} is the interest rate for the US dollar, IRD_t^h equals $i_t^h - i_t^{US}$, and s_t^h is the log spot exchange rate denoting the number of foreign currency units per US dollar. The return of the conventional [+2, −2] CT strategy is just an equally weighted average of the returns of 2 long and 2 short positions. We further assume that the CT investor rebalances every 1 month, cf., e.g., Baillie and Chang (2011), Burnside et al. (2011), Villanueva (2007).

2.2. Conditional CT strategies

2.2.1. Determination of extreme IRDs

One of the determinants of currency risk premia is the size of the IRD. One possible explanation for this is provided by the limits to speculation hypothesis. This hypothesis postulates that traders will only exploit the currency mispricing when the absolute IRD is high enough to cover the transaction and opportunity costs (Lyons, 2001; Sarno, 2005). As these costs are heterogeneous across investors and time-varying a larger absolute IRD will result in more traders entering the market to risk-arbitrage away the mispricing such that UIP is more likely to hold. Empirical support is provided by e.g. Huisman et al. (1998) and Lothian and Wu (2011). Therefore, the CT performance is expected to improve when conditioning on extreme absolute IRDs.

To determine whether an IRD is extreme we identify the 10% highest absolute IRDs across all currencies relative to the US dollar over the prior 5-year rolling window. Within the extreme sampling literature there is no clear choice for the size of the extreme tails where UIP should hold. For example, Huisman et al. (1998) analyze a bandwidth of 5 to 20%, while Lothian and Wu (2011) consider a 1 to 10% bandwidth. In the main analysis we will use the 10% cutoff value but as a robustness check we consider a range from 1 to 10%.

2.2.2. Determination of high FX volatility regimes

Aside from the size of the IRDs, FX volatility is another important determinant of currency risk premia. One reason is again provided by the limits to speculation hypothesis. Following the same logic as discussed in the previous section, the limits to speculation hypothesis implies that larger absolute IRDs will result in more traders entering the market to risk-arbitrage away the mispricing such that FX volatility will increase due to higher trade volumes (Evans, 2010; Evans and Lyons, 2002a,b; Sager and Taylor, 2008) and UIP is more likely to hold.

This implied relation between the likelihood of UIP to be valid and the degree of FX volatility is also supported empirically. For example, Clarida et al. (2009) and Christiansen et al. (2011) show that CT strategies are only profitable when FX volatility is not too high. Therefore, we expect that conditioning the CT strategy on high FX volatility regimes would contribute to the CT performance.

To determine the high FX volatility regimes we perform two steps. First, based on the returns r_t of the conventional CT strategy (base-case) we calculate the realized volatility $FXVOL_t$ over the prior 5-year rolling window. Specifically, $FXVOL_t$ is an exponentially weighted moving average (EWMA) such that[7,8]

$$FXVOL_t = \frac{\sum_{i=0}^{\tau-1} \lambda^i \left(\Delta s_{t-i}^{\text{carry}} - \overline{\Delta s_t^{\text{carry}}} \right)^2}{\sum_{i=0}^{\tau-1} \lambda^i} \tag{2}$$

where s_t is the return on the currencies selected in the conventional CT strategy (excluding the IRDs), $\bar{s}_t$ is the average return of the currencies selected by the conventional CT strategy over the prior 5-year rolling window (for which τ is 60 months), and the exponential decay parameter λ is such that the weight of the midpoint of the rolling window equals 0.5.[9] See also Eq. (1). As our sample starts in January 1975 this implies that we have to sacrifice the first five years of our sample to calculate $FXVOL_t$.

Second, based on this measure we set the FX volatility regime to 'high' when $FXVOL_t$ exceeds the 90th percentile of the $FXVOL$ distribution determined over the prior 5-year rolling window, thereby again reducing the sample by five years.[10] De

[6] Next to this [+2, −2] strategy we will also investigate the performance of the [+1, −1], [+3, −3], [+4, −4], and [+5, −5] strategies as robustness check.

[7] Clarida et al. (2009): also apply this EWMA measure on realized CT returns. Furthermore, we are interested in the behavior of currencies employed in the CT strategy, not in FX volatility of all currencies. The latter might be driven by currencies that are not included in the CT strategy. To ensure our measure of volatility is not driven by any IRD effect, we specified Eq. (2) using Δs_{t-i} instead of r_{t-i}. When testing it using r_{t-i}, it appears that the measure of FX volatility hardly deviates with or without IRDs, which we do not find surprising given the low frequency of IRD adjustments for most currencies, and the very fast adjustment of FX rates to changes in IRDs.

[8] Note that due to our definition of CT return this FX volatility measure only uses data known to the investor at time t.

[9] In this case $\lambda^{60/2} = 0.5$ such that $\lambda \approx 0.977$.

[10] Clarida et al. (2009) and Menkhoff et al. (2012) use a 75% threshold, albeit that the volatility measure of the latter is different from ours. Menkhoff et al. (2012) focus on volatility innovations in their analysis. The term 'high' is quite arbitrary, and we therefore test the robustness of our findings by varying the threshold in Section 3.5.

facto, we thus sacrifice 10 years of our sample in total such that effectively our sample starts in January 1985. Fig. 2 visualises the construction of our volatility measures.

2.2.3. Construction of conditional CT strategies

Our main conditional CT strategy basically ignores investment opportunities for which UIP is likely to hold, namely for large IRDs during high FX volatility, and follows the conventional CT strategy otherwise. In other words, for the current investment set, i.e. the currencies currently available, the investor first determines the FX volatility regime. If FX volatility, $FXVOL_t$, is high (i.e. higher than the 90th percentile of the *FXVOL* distribution over the prior 5-year rolling window) then the investor should determine which of the current IRDs are extreme. If a current IRD is extreme, i.e. this observation belongs to the 10% highest absolute IRDs over the prior 5-year rolling window, then the investor should exclude the underlying currency from his investment set before implementing an otherwise conventional CT strategy. If FX volatility is (relatively) low, then the investor should follow the base-case strategy. In sum, the investor should apply the following procedure to implement our main conditional CT strategy:

$$\mathrm{CONDFXIRD_t} = \begin{cases} \text{Conv. CT strategy excl. extreme IRDs} & \text{if } FXVOL_t \text{ is high} \\ \text{Conv. CT strategy} & \text{if } FXVOL_t \text{ is low} \end{cases} \tag{3}$$

As discussed earlier we expect that this conditional CT strategy outperforms the base-case.

We also implement two other conditional CT strategies that condition on high FX volatility and large IRDs separately. The former strategy follows the conventional CT strategy when FX volatility is low while it follows a passive strategy (i.e. no investments) when FX volatility is high:

$$\mathrm{CONDFX}_t = \begin{cases} \text{No investments/positions} & \text{if } FXVOL_t \text{ is high} \\ \text{Conv. CT strategy} & \text{if } FXVOL_t \text{ is low} \end{cases} \tag{4}$$

The base-case strategy will thus outperform the CONDFX strategy if the highest non-extreme IRDs during high FX volatility regimes are still profitable on average. In these cases these IRDs contribute to the base-case's performance, otherwise the CONDFX strategy would outperform the base-case.

The second alternative conditional CT strategy excludes the extreme IRDs from the investment set before following an otherwise conventional CT strategy regardless of the FX volatility regime:

$$\mathrm{CONDIRD}_t = \text{Conv. CT strategy excl. extreme IRDs} \tag{5}$$

When FX volatility is high, extreme IRDs may not contribute to the CT performance since UIP will likely hold. Therefore, removing these observations from the investment opportunity set should enhance the CT performance. On the other hand, when FX volatility is not too high we expect that not all currency mispricing is exploited due to the heterogeneity of investors' transaction and opportunity costs. As a consequence UIP does not hold and currency carry trades are still profitable. This implies that removing extreme IRDs in such regimes could hurt CT performance. Overall, depending on the size of the expected excess gains or losses relative to the base-case during high or low FX volatility regimes, the CONDIRD strategy could either out- or underperform the base-case.

3. Analysis

3.1. Data

To analyze the performance of the conventional and our three conditional CT strategies we use a sample of 25 countries that effectively covers the 1985-mid 2015 period.[11]

We use monthly spot exchange rates relative to the US dollar and 1-month interest rates (cf. e.g., Barroso and Santa-Clara, 2015; Hochradl and Wagner, 2010; Menkhoff et al., 2012). To construct this end-of-month data set we use daily spot exchange rates and interest rates, both taken from Datastream, similar to Burnside et al. (2008). For the exchange rates we use indirect quotes, i.e. exchange rates are expressed as units of foreign currency per US dollar. We use log spot rates in our analysis cf. prior studies. The exchange rate and/or interest rate were not available for the entire sample period for some countries resulting in an unbalanced panel. When a country becomes member of the euro area its exchange rate and interest rate are replaced by respectively the euro-dollar exchange rate and euro area interest rate.[12]

When available, we retrieved data from January 1975 onwards but as we had to sacrifice 10 years of data to identify the FX volatility regime our sample effectively starts in January 1985.[13]

[11] These countries are Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hong Kong, Iceland, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Portugal, Singapore, South Korea, Spain, Sweden, Switzerland, and the United Kingdom.

[12] See http://www.ecb.int/euro/intro/html/map.en.html for an overview of the euro area countries and their accession dates.

[13] To calculate FX volatility at time t we use data of the previous 5 years and we sacrifice another 5 years to determine the FX volatility regime by comparing current FX volatility to its empirical distribution. See Section 2.2.2 for further details.

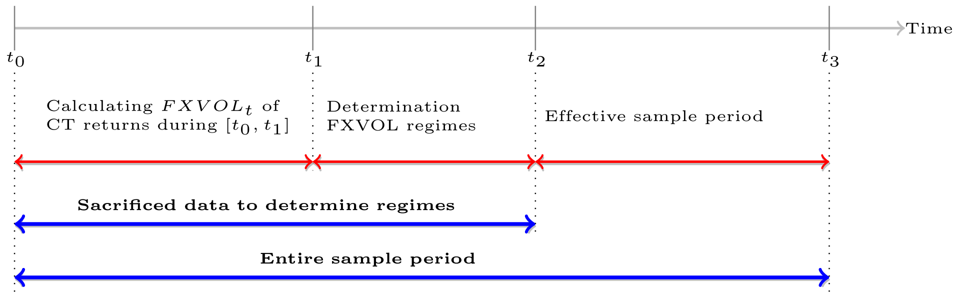

Fig. 2. Measuring volatility over time. Figure shows how we have constructed our volatility measures. In the above timeline, the period $[t_0, t_1]$ is used to calculate $FXVOL_t$, as specified in Eq. (2). We repeat this for 60 times over new intervals (i.e., a rolling window that shifts towards t_2) and then during the $[t_1, t_2]$ period, we determine the $FXVOL_t$ regime ('high' or 'low') by taking the 90th percentile of the past 5 year window of *FXVOL* observations. This reduces the active analysis period to $[t_2, t_3]$, during which we can infer whether an $FXVOL_t$ sampled between t_2 and t_3 is classified as 'high' or 'low'.

Table 1 shows some summary statistics of the currencies employed. For most currencies the average exchange rate change (5th column) does not match the corresponding average interest rate differential (6th column). Furthermore, they are often of opposite sign. Both indicate that, on average, UIP does not hold which is in line with the general consensus in this literature (e.g., Engel, 1996; Hodrick, 1987; Lewis, 1995; Sarno, 2005). This explains the average profitability of a carry trade strategy with a long position in the corresponding currency (see Eq. (1)) for most currencies, see the 7th column of Table 1.[14]

The right-hand side of Table 1 shows how often a currency has been deployed in our CT strategies. We observe that Australia, Greece, Iceland, Italy, New Zealand, and the United Kingdom have the highest number of long positions, while Japan, Singapore, and Switzerland have the highest number of short positions. Some of these currencies have also been identified in other studies as popular target currencies, e.g. New Zealand and United Kingdom, and funding currencies, e.g. Japan, Switzerland (see Christiansen et al., 2011; Galati et al., 2007; Kohler, 2010).

3.2. Results

We start off with determining the risk and return characteristics of a conventional (base-case) CT strategy, based on 2 long positions in the currencies with the highest interest rates and 2 short positions in the currencies with the lowest interest rates, that rebalances monthly. Table 2 reports various performance indicators. The base-case yields an average annual return of 4.74%, a holding period return of \$3.59, a Sharpe ratio of 0.47, and a standard deviation of 10.14%. Conform prior research, skewness is negative and kurtosis is high. Next we implement our conditional CT strategies (see Section 2 for details).

Our main conditional CT strategy (CONDFXIRD), that tries to avoid investing in extreme IRDs during high FX volatility periods (see Section 2 for more details), is able to outperform the conventional CT strategy as the average annualized returns are higher (5.88% instead of 4.74%), the holding period return increases (5.11 instead of 3.59), Sharpe ratio is higher (0.59 instead of 0.47), and skewness is lower (-1.16 instead of −1.60); see Table 2. Therefore, by disregarding investment opportunities for which UIP is likely to hold our main conditional CT strategy outperforms the base-case. Next, we analyze the separate effects of conditioning on IRDs only and on FX volatility regimes only.

Our second conditional strategy, CONDFX, solely conditions on FX volatility regimes. That is, when FX volatility is relatively high over the recent 5 years, it moves into a passive strategy (i.e., no investment). In the current setting, this strategy has the largest outperformance relative to the base-case: the average annualized return is higest (6.00% instead of 4.74%), and so is the holding period return (5.64 instead of 3.59), Sharpe ratio (0.78 instead of 0.47), and skewness (−0.74 instead of −1.60). Furthermore, note that the standard deviation has decreased dramatically relative to the base-case (7.67% instead of 10.14%), which is due to the no-investments periods of this strategy, so that as a consequence the Sharpe ratio becomes quite high by definition. In this case, by not investing during volatile periods, the added value of possible profitable IRDs is outweighed by the loss made on extreme IRDs.

Our last conditional strategy, CONDIRD, solely conditions on extreme IRDs, i.e. this strategy always avoids the currencies with the highest absolute IRDs, regardless of the FX volatility regime. Only the standard deviation and skewness show better performance relative to the base-case, all other indicators are doing worse than the conventional CT strategy. Therefore, extreme IRDs during low volatility regimes do add value to CT performance.

Fig. 3 visualizes how our four CT strategies perform over time. In the first few years of the graph, some lines overlap. This is because during this low FX volatility regime, the returns of the CONDFXIRD and CONDFX strategies are equal to the

[14] Although many studies report the average profitability of portfolios of currencies (e.g., Clarida et al., 2009; Lustig et al., 2011; Menkhoff et al., 2012), only some also report the excess returns for individual currencies, see e.g. Table 1 in Brunnermeier et al. (2009). Our statistics confirm the ones reported in these studies.

Table 1

Summary statistics. Table shows for each country the start and end date for which data is available for both exchange rate and interest rate, the number of observations, the annualized average exchange rate change relative to the USD ($\overline{\Delta S}$) in %, the annualized average interest rate differential (IRD) relative to the USD ($\overline{IRD}$) in %, and the number of times each currency was employed in the conventional CT strategy (base-case) and our three conditional CT strategies. See Section 2 or Fig. 1 for the implementation of the CT strategies. The monthly data set covers 25 currencies spanning the 1985 to mid-2015 period.

Country	Start date	End date	Nr. obs	$\overline{\Delta s}$	$\overline{IRD}$	$\bar{r}$	Conventional		CONDFXIRD		CONDFX		CONDIRD	
				(ann. %)	(ann. %)	(ann. %)	Long	Short	Long	Short	Long	Short	Long	Short
Australia	30-04-97	29-05-15	218	0.09	2.17	2.07	59	0	71	0	55	0	120	0
Austria	28-06-91	31-12-98	91	−0.67	0.52	1.18	0	0	6	0	0	0	6	0
Belgium	31-01-85	31-12-98	168	−4.13	0.82	4.95	1	0	1	0	0	0	1	0
Canada	31-01-85	29-05-15	365	−0.20	0.74	0.95	14	0	14	0	14	0	15	0
Czech Republic	30-04-92	29-05-15	278	−0.76	2.19	2.95	52	0	53	2	49	0	37	15
Denmark	28-06-85	29-05-15	360	−1.61	0.92	2.53	15	29	18	29	11	29	23	29
Finland	30-04-92	31-12-98	81	2.27	0.85	−1.43	6	3	6	3	0	3	6	6
France	31-01-85	31-12-98	168	−3.68	1.50	5.19	4	0	7	0	3	0	9	0
Germany	31-01-85	29-05-15	365	−1.93	−0.33	1.60	0	49	1	49	0	49	1	51
Greece	29-04-94	29-12-00	81	6.08	7.27	1.19	78	0	78	0	78	0	77	0
Hong Kong	30-04-97	29-05-15	218	0.00	−0.06	−0.06	1	44	1	44	1	39	8	44
Iceland	31-07-98	29-05-15	203	3.65	6.38	2.73	187	0	163	0	162	0	71	0
Ireland	31-01-85	31-12-98	168	−2.73	2.90	5.63	24	0	24	0	19	0	34	0
Italy	31-01-85	31-12-98	168	−0.96	4.27	5.23	75	0	57	0	57	0	58	0
Japan	31-01-85	29-05-15	365	−2.41	−2.24	0.17	0	242	0	237	0	218	0	195
Netherlands	31-01-85	31-12-98	168	−4.38	−0.32	4.05	0	3	0	3	0	3	0	3
New Zealand	30-04-97	29-05-15	218	0.14	2.62	2.48	95	0	88	0	75	0	115	0
Norway	30-04-97	29-05-15	218	0.55	1.17	0.62	30	0	39	0	29	0	34	0
Portugal	30-11-92	31-12-98	74	3.41	3.93	0.52	20	0	24	0	12	0	24	0
Singapore	29-01-88	29-05-15	329	−1.47	−1.02	0.45	0	103	0	103	0	54	0	111
South Korea	30-07-04	29-05-15	131	−0.43	1.58	2.01	0	0	1	0	0	0	25	0
Spain	30-04-92	31-12-98	81	5.11	3.58	−1.54	10	0	20	0	4	0	20	0
Sweden	30-04-97	29-05-15	218	0.31	−0.03	−0.34	0	0	0	0	0	0	0	2
Switzerland	31-01-85	29-05-15	365	−3.46	−1.48	1.98	0	221	4	224	0	213	4	238
United Kingdom	31-01-85	29-05-15	365	−1.08	1.95	3.03	59	0	54	0	47	0	42	0
United States	–	–	–	–	–	–	0	36	0	36	0	8	0	36
							730	730	730	730	616	616	730	730

Table 2

Performance. Table shows the performance measured by annualized mean return in % $\overline{R}$, annualized standard deviation in % S.D., skewness Skew., kurtosis Kurt., Sharpe ratio S, and holding period return (end-of-period accumulated wealth of \$1 investment) HPR for the conventional CT strategy (base-case) and our three conditional CT strategies. See Section 2 or Fig. 1 for the implementation of the CT strategies. The monthly data set covers 25 currencies spanning the 1985 to mid-2015 period. All economic outperformances relative to the base-case strategy (higher $\overline{R}$, lower S.D, higher Sharpe, less negative Skew, lower Kurt, and higher HPR) are highlighted in gray.

	$\overline{R}$	S.D.	Skew.	Kurt.	Sharpe	HPR
Conventional	4.74	10.14	−1.60	8.41	0.47	3.59
CONDFXIRD	5.88	9.98	−1.16	8.36	0.59	5.11
	(0.09)	(0.72)	(0.21)	(0.97)	(0.11)	(0.09)
	[0.06]	[0.67]	[0.08]	[0.96]	[0.04]	[0.06]
CONDFX	6.00	7.67	−0.74	5.95	0.78	5.64
	(0.29)	(0.00)	(0.02)	(0.10)	(0.04)	(0.29)
	[0.39]	[0.02]	[0.04]	[0.12]	[0.11]	[0.39]
CONDIRD	4.40	9.75	−1.29	8.88	0.45	3.28
	(0.73)	(0.40)	(0.42)	(0.81)	(0.88)	(0.73)
	[0.74]	[0.35]	[0.31]	[0.72]	[0.87]	[0.74]

base-case strategy because they do not condition during these regimes (see Section 2.2.3). Note that the CONDFX strategy has a flat performance during high FX volatility periods because no investments are made in such periods. Our main conditional CT strategy that conditions on *both* FX volatility *and* extreme IRDs outperforms the base-case/conventional CT strategy over the entire sample period in terms of holding period returns. The outperformance of the CONDFX strategy is even better, but this outperformance is less stable in the alternative settings (as we will show throughout the paper). The performance of the last conditional CT strategy in Fig. 3 which conditions on extreme IRDs only does not seem to add value.

In the following sections we implement several robustness checks to test the sensitivity of our results to different settings.

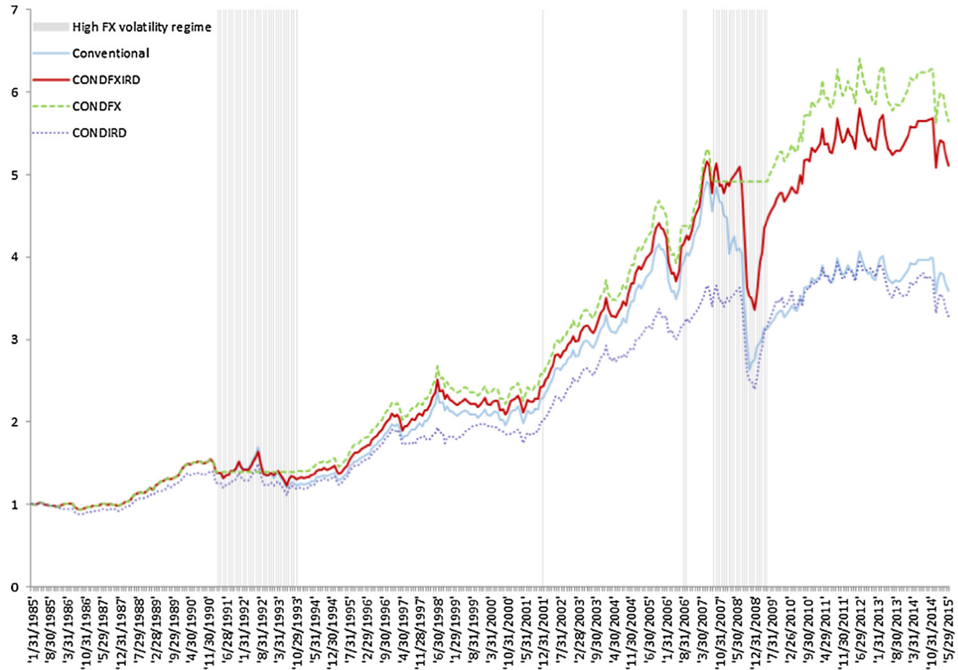

Fig. 3. Holding period return. Figure shows the accumulated wealth (holding period return) of a $1 investment for the conventional CT strategy (base-case) and our three conditional CT strategies. See Section 2 or Fig. 1 for the implementation of the CT strategies. The monthly data set covers 25 currencies spanning the 1985 to mid-2015 period. High FX volatility regimes are indicated by the shaded areas.

3.3. Performance for various extreme IRD cutoff rates

In this section we test the sensitivity of our findings by changing the threshold to determine whether an IRD is extreme or not. The performance of the CONDFXIRD and CONDIRD strategies for different cutoff values ranging from 1 to 10% are reported in Table 3.[15]

In the previous section we already showed that our main conditional CT strategy that conditions on both high FX volatility and extreme IRDs outperforms the conventional CT strategy when a cutoff of 10% is used and the same conclusion holds for other cutoff values (see panel B of Table 3). The outperformance of the CONDFXIRD strategy translates into increases in average returns of about 0.9 to 1.3 percent point and in holding period returns of about 1.2 to 1.8 additional dollars of 1 dollar invested at the start of the sample period. The Sharpe ratio increases about 0.10 to 0.14 which translates into a 21% to 30% increase relative to the base-case. The results for our CONDIRD strategy are mixed in terms of average returns, holding period returns, Sharpe, and kurtosis. In some cases this strategy outperforms the base-case while it underperforms in others although the size of the underperformance is lower than the outperformance. The CONDIRD strategy does outperform the base-case in terms of skewness and standard deviation for all cutoff values. These findings again show the importance of conditioning on both high FX volatility and extreme IRDs instead of conditioning on one dimension only. Therefore extreme IRDs should only be ignored when FX volatility is high.

3.4. Impact of the definition of 'extreme' IRDs

In our main analysis we have identified an IRD as extreme when it belongs to the upper 10th percentile of the empirical IRD distribution determined over the previous five years. In this section we test whether our findings are robust to other definitions. We consider four alternative measures for which the results are reported in panels B and C of Table 4.[16]

The first measure considers *positive* extremes instead of *absolute* extremes. The so-called 'extreme' sampling literature (see e.g. Huisman et al., 1998; Lothian and Wu, 2011) focuses on absolute IRDs and shows that UIP holds in these extremes; however, Bansal (1997) and Wu and Zhang (1996) have shown that UIP only holds for *positive* extremes.[17] We test the performance of the CONDFXIRD and CONDIRD strategies when positive (instead of absolute) extreme IRDs are removed from the investment opportunity set. Using this alternative definition the CONDFXIRD strategy, that conditions on both extreme IRDs and

[15] By definition, changing the cutoff value does not have an impact on CONDFX (and on the base-case) as this strategy does not condition on extreme IRDs.

[16] The main results using extreme absolute IRDs are reported for ease of comparison.

[17] Both Bansal (1997) and Wu and Zhang (1996) use direct quotations of foreign exchange rates, and find that UIP only holds for negative extreme IRDs. Since we use indirect quotations for exchange rates their findings translate into positive extreme IRDs for our data.

Table 3
Impact of the IRD truncation percentile Table shows the impact of the IRD truncation percentile on the performance in terms of annualized mean return $\overline{R}$, annualized standard deviation S.D., skewness Skew., kurtosis Kurt., Sharpe ratio S, and holding period return (end-of-period accumulated wealth of $1 investment) HPR for the conventional CT strategy (base-case) and the two conditional CT strategies that condition on IRDs. See Section 2 or Fig. 1 for the implementation of the CT strategies. Panel A displays the performance of the base-case. For the CT strategies that condition on IRDs the performance for various percentiles ν of truncation, ranging from 1 to 10%, are reported in panels B (CONDFXIRD) and C (CONDIRD). The monthly data set covers 25 currencies spanning the 1985 to mid-2015 period. All economic outperformances relative to the base-case strategy (higher $\overline{R}$, lower S.D, higher Sharpe, less negative Skew, lower Kurt, and higher HPR) are highlighted in gray. Statistical significance on the basis of HAC estimators is indicated by * and ** respectively (90% and 95% confidence intervals). See Section 3.11 for a discussion on statistical significance testing. The CONDFX strategy does not condition on extreme IRDs and is therefore ignored here.

	Panel A: Conventional CT strategy					
	$\overline{R}$	S.D.	Skew.	Kurt.	Sharpe	HPR
	4.74	10.14	−1.60	8.41	0.47	3.59
	Panel B: CONDFXIRD					
ν	$\overline{R}$	S.D.	Skew.	Kurt.	Sharpe	HPR
1	5.68*	9.99	−1.13**	8.11	0.57*	4.81*
2	5.68*	9.99	−1.13**	8.11	0.57*	4.81*
3	5.78*	9.97	−1.14*	8.17	0.58**	4.96*
4	5.87*	9.98	−1.15*	8.16	0.59**	5.10*
5	6.09**	10.01	−1.12**	8.14	0.61**	5.44**
6	5.96*	10.05	−1.16*	8.15	0.59*	5.22*
7	5.92*	10.03	−1.16*	8.20	0.59*	5.16*
8	5.97*	9.97	−1.14*	8.29	0.60**	5.25*
9	5.97*	9.95	−1.15*	8.34	0.60**	5.25*
10	5.88*	9.98	−1.16*	8.36	0.59**	5.11*
	Panel C: CONDIRD					
ν	$\overline{R}$	S.D.	Skew.	Kurt.	Sharpe	HPR
1	5.48	10.01	−1.12**	8.11	0.55	4.53
2	5.49	9.92	−1.18	8.29	0.55	4.55
3	5.41	9.87	−1.21	8.39	0.55	4.45
4	5.23	9.89	−1.17	8.29	0.53	4.20
5	5.13	10.02	−1.15*	8.07	0.51	4.06
6	4.66	9.96	−1.21	8.31	0.47	3.53
7	4.30	9.81	−1.23	8.61	0.44	3.18
8	4.22	9.77	−1.22	8.66	0.43	3.11
9	4.23	9.74	−1.24	8.76	0.43	3.12
10	4.40	9.75	−1.29	8.88	0.45	3.28

high FX volatility, again outperforms the base-case across all performance measures. The CONDIRD strategy only marginally outperforms the base-case for almost all performance indicators.

The second alternative measure considers the other side of the IRD spectrum and defines an IRD as extreme when it belongs to the *lower* 10th percentile of the empirical IRD distribution, i.e. the negative extreme IRDs. Following the same logic (see footnote [17]) we expect that negative extreme IRDs are profitable investment opportunities and, as a consequence, removing them leads to *under*performance relative to the base-case. This explains why removing absolute extreme IRDs or positive extreme IRDs leads to larger outperformance than removing negative extreme IRDs. This is in line with Bansal and Dahlquist (2000) who find that UIP is particularly violated in periods during which US interest rates exceed foreign interest rates.

The third and fourth alternative measures do not focus on the extreme IRDs over the entire sample (as is done in the main analysis) but identify extremes for each individual currency and for each time period, respectively. Notably, the third (fourth) alternative measure identifies an IRD for a particular currency (time period) as extreme when it belongs to the upper 10th

Table 4

Impact of the definition of 'extreme' IRDs. Table shows the impact of the definition of extreme IRD on the performance measured by annualized mean return in % $\overline{R}$, annualized standard deviation in % S.D., skewness Skew., kurtosis Kurt., Sharpe ratio S, and holding period return (end-of-period accumulated wealth of $1 investment) HPR for the conventional CT strategy (base-case) and the two conditional CT strategies that condition on IRDs. See Section 2 or Fig. 1 for the implementation of the CT strategies. Panel A displays the performance of the base-case. For the CT strategies that condition on IRDs the performance for various definitions of an 'extreme' IRD are reported in panels B (CONDFXIRD) and C (CONDIRD). The 'ABS' definition is used in our main analysis and is based on the absolute extreme IRDs of the panel of all currencies over the past 5 years (rolling window). The 'POS' definition does the same but considers positive extreme IRDs only while the 'NEG' definition only considers negative extreme IRDs. The 'Currency' extremes considers the upper 10th percentile extreme IRDs per currency. The 'Time' definition considers the upper 10th percentile time period extreme IRDs. The monthly data set covers 25 currencies spanning the 1985 to mid-2015 period. All economic outperformances relative to the base-case strategy (higher $\overline{R}$, lower S.D, higher Sharpe, less negative Skew, lower Kurt, and higher HPR) are highlighted in gray. Statistical significance on the basis of HAC estimators is indicated by * and ** respectively (90% and 95% confidence intervals). See Section 3.11 for a discussion on statistical significance testing. The CONDFX strategy does not condition on extreme IRDs and is therefore ignored.

	$\overline{R}$	S.D.	Skew.	Kurt.	Sharpe	HPR
Panel A: Conventional CT strategy						
	4.74	10.14	−1.60	8.41	0.47	3.59
Panel B: CONDFXIRD						
	$\overline{R}$	S.D.	Skew.	Kurt.	Sharpe	HPR
ABS	5.88*	9.98	−1.16*	8.36	0.59**	5.11*
POS	5.90**	9.99	−1.11**	8.07	0.59**	5.14**
NEG	4.89	9.80	−1.41	7.57	0.50	3.80
Currency	5.28	9.90	−1.09**	6.88**	0.53	4.27
Time	5.87*	9.58	−1.22*	8.12	0.61**	5.16*
Panel C: CONDIRD						
ABS	4.40	9.75	−1.29	8.88	0.45	3.28
POS	4.72	9.89	−1.19	8.24	0.48	3.60
NEG	3.09	9.11**	−1.47	8.49	0.34	2.25
Currency	4.92	9.74	−1.09**	7.03**	0.50	3.84
Time	5.27	9.09**	−1.32	8.82	0.58	4.35

percentile of the empirical IRD distribution of that particular currency (time period). Again, the CONDFXIRD strategy is able to outperform the base-case under both alternative measures across all performance indicators. The CONDIRD is also able to outperform the base-case across virtually all dimensions when extremes are identified per currency or time period.

As in Section 3.3, when varying the definition of 'extremes', our findings support the notion that both extreme IRDs and high FX volatility should be included in the conditional CT strategy in order to outperform the conventional CT strategy.

3.5. Impact of FX volatility thresholds

In this section we test whether our main findings change when the threshold to determine the 'high' FX volatility regime is altered.[18] We consider a threshold range from 0.5 (i.e. median) to 0.99 (i.e. the 99th percentile of the FX volatility distribution, see Section 2.2.2 for further details). The results are reported in Table 5. Once again, our CONDFXIRD strategy outperforms the base-case in virtually all cases, while the results are more mixed for CONDFX. This again supports our earlier claim that both dimensions are important to include in the conditional CT strategy. Our CONDFX strategy follows the base-case/conventional CT strategy in the 'low volatility' regimes, and is passive (i.e., no investment, and thus a zero return) in the 'high volatility' regimes. Hence, if our FXVOL incorrectly classifies a regime as 'high' it foregoes income (because it does NOT invest), and if it incorrectly labels a regime as 'low' volatility then its returns suffer from losses that could have been avoided. Therefore, the outperformance of our CONDFX strategy is very sensitive to the choice of the FX volatility threshold.

3.6. Impact of the definition of FX volatility

In this section we apply several robustness checks for the definition of FX volatility. We vary the decay parameter in the exponentially weighted moving average (EWMA) measure to calculate FX volatility. Furthermore, as our FX volatility measure in the main analysis is backward-looking, we also test whether our findings are robust to the use of a forward-looking measure.

[18] By definition, changing the threshold for the 'high' FX volatility regime does not have an impact on the CONDIRD stratgey and the base-case as these strategies do not condition on FX volatility. For ease of comparison we still report the results for the base-case.

Table 5
Impact of FX volatility thresholds. Table shows the impact of the FX volatility threshold on the performance measured by annualized mean return in % $\overline{R}$, annualized standard deviation in % S.D., skewness Skew., kurtosis Kurt., Sharpe ratio S, and holding period return (end-of-period accumulated wealth of $1 investment) HPR for the conventional CT strategy (base-case) and the two conditional CT strategies that condition on FX volatility. See Section 2 or Fig. 1 for the implementation of the CT strategies. Panel A displays the performance of the base-case. For the CT strategies that condition on FX volatility the performance for various FX volatility thresholds η, ranging from 0.5 to 0.99, are reported in panels B (CONDFXIRD) and C (CONDFX). The case of $\eta = 0.90$ is used in our main analysis. The monthly data set covers 25 currencies spanning the 1985 to mid-2015 period. All economic outperformances relative to the base-case strategy (higher $\overline{R}$, lower S.D, higher Sharpe, less negative Skew, lower Kurt, and higher HPR) are highlighted in gray. Statistical significance on the basis of HAC estimators is indicated by * and ** respectively (90% and 95% confidence intervals). See Section 3.11 for a discussion on statistical significance testing. The CONDIRD strategy does not condition on FX volatility and is ignored here.

	Panel A: Conventional CT strategy					
	$\overline{R}$	S.D.	Skew.	Kurt.	Sharpe	HPR
	4.74	10.14	−1.60	8.41	0.47	3.59
	Panel B: CONDFXIRD					
η	$\overline{R}$	S.D.	Skew.	Kurt.	Sharpe	HPR
0.5	5.43	10.00	−1.22	8.28	0.54	4.46
0.6	5.37	9.99	−1.23	8.30	0.54	4.38
0.7	5.44	10.09	−1.15*	8.09	0.54	4.45
0.75	5.66	10.01	−1.14*	8.24	0.57	4.77
0.8	5.90*	9.99	−1.16*	8.33	0.59*	5.14*
0.9	5.88*	9.98	−1.16*	8.36	0.59**	5.11*
0.95	5.50	10.08	−1.37	9.26	0.55	4.54
0.99	5.04	10.07	−1.59	9.09	0.50	3.95
	Panel C: CONDFX					
η	$\overline{R}$	S.D.	Skew.	Kurt.	Sharpe	HPR
0.5	2.44	6.30**	−0.85	8.63	0.39	1.97
0.6	3.17	6.62**	−0.82	7.46	0.48	2.45
0.7	4.35	7.00**	−0.62**	6.81	0.62	3.48
0.75	4.40	7.36**	−0.76*	6.57	0.60	3.50
0.8	5.05	7.55**	−0.70**	6.18	0.67	4.25
0.9	6.00	7.67**	−0.74**	5.95	0.78	5.64
0.95	5.90	8.39**	−1.18**	8.14	0.70*	5.38
0.99	5.64	8.93*	−1.19**	7.28	0.63	4.90

In our main analysis the decay parameter λ was such that the weight of the midpoint of the rolling window equals 0.5 resulting in $\lambda \approx 0.977$. [19] Here we consider λ to be 0.9, 0.95, 0.99, and 1 where the latter implies an equally weighted average. The results are reported in panels B and C in Table 6 for respectively the CONDFXIRD and CONDFX strategy. For virtually all cases we see that both the CONDFXIRD and CONDFX strategies outperform the base-case.

Furthermore, we also test a forward-looking measure. For this we use the JP Morgan implied FX volatility index based on the G7 currencies which is available through Bloomberg. This measure is available from June 1992 onwards so the sample for this robustness test effectively starts in June 1997 as the first 5 years are used to determine the FX volatility regime (see Section 2.2.2). To be able to compare the results we also apply our own backward-looking FX volatility measure to the same window. The results of the backward-looking measure and the forward-looking measure are reported in respectively panel D and E of Table 6.[20] The CONDFX strategy that conditions on high FX volatility periods consistently outperforms the base-case for both backward- and forward-looking measure. The CONDFXIRD strategy also outperforms the conventional CT strategy, but less convincingly.

3.7. VIX as an alternative regime switch indicator

So far we focussed on FX volatility as our regime switch indicator. One might argue, however, that it is not FX volatility that drives our results, but instead it could be a general appetite for risk taking. It is often suggested to take VIX as a proxy for

[19] See Section 2.2.2.
[20] For ease of comparison we also report the results for the base-case and the CONDIRD strategy as these differ due to the change in sample period.

Table 6
Impact of FX volatility definition. Table shows the impact of the FX volatility definition on the performance measured by annualized mean return in % $\overline{R}$, annualized standard deviation in % S.D., skewness Skew., kurtosis Kurt., Sharpe ratio S, and holding period return (end-of-period accumulated wealth of $1 investment) HPR for the conventional CT strategy (base-case) and the two conditional CT strategies that condition on FX volatility. See Section 2 or Fig. 1 for the implementation of the CT strategies. Panel A displays the performance of the base-case. For the CT strategies that condition on FX volatility the performance for various decay parameters λ, ranging from 1.00 to 0.90, are reported in panels B (CONDFXIRD) and C (CONDFX). The case of $\lambda \approx 0.977$ is used in our main analysis. Panel E applies the forward-looking JP Morgan implied FX volatility index. Panel D recalculates our main results using the backward-looking EWMA measure for the same time window for the ease of comparison. The monthly data set covers 25 currencies spanning the 1985 to mid-2015 period. All economic outperformances relative to the base-case strategy (higher $\overline{R}$, lower S.D, higher Sharpe, less negative Skew, lower Kurt, and higher HPR) are highlighted in gray. Statistical significance on the basis of HAC estimators is indicated by * and *presup* $**$ respectively (90% and 95% confidence intervals). See Section 3.11 for a discussion on statistical significance testing. The CONDIRD strategy does not condition on FX volatility and is ignored here.

	Panel A: Conventional CT strategy					
	$\overline{R}$	S.D.	Skew.	Kurt.	Sharpe	HPR
	4.74	10.14	−1.60	8.41	0.47	3.59
λ	Panel B: CONDFXIRD					
1.00	5.98*	10.01	−1.16*	8.29	0.60**	5.26*
0.99	5.92*	9.99	−1.16*	8.35	0.59**	5.17*
0.95	5.40	9.92	−1.18*	8.42	0.54	4.42
0.90	4.88	10.33	−1.49	9.45	0.47	3.73
λ	Panel C: CONDFX					
1.00	5.51	7.64**	−0.70**	5.99	0.72	4.86
0.99	6.12	7.68**	−0.75**	5.93	0.80*	5.85
0.95	5.42	7.89**	−0.71**	5.62*	0.69	4.71
0.90	4.81	9.00*	−1.49	9.13	0.53	3.79
	Panel D: backward-looking EWMA measure (mid-1997 to mid-2015)					
Conventional	4.47	10.93	−1.51	7.77	0.41	2.01
CONDFXIRD	6.07	10.76	−1.10	8.53	0.56*	2.69
CONDFX	6.05	8.61	−0.55*	4.94*	0.70	2.79
CONDIRD	4.08	10.26	−1.28	9.68	0.40	1.90
Panel E: forward-looking JP Morgan implied FX volatility index (mid-1997 to mid-2015)						
Conventional	4.47	10.93	−1.51	7.77	0.41	2.01
CONDFXIRD	5.14	10.99	−1.33	9.32	0.47	2.26
CONDFX	5.52	9.46	−1.09**	7.30	0.58	2.50
CONDIRD	4.08	10.26	−1.28	9.68	0.40	1.90

this general appetite for risk taking.[21] A high VIX means that investors' risk aversion is high, and thus the demand for assets (including FX) is low. Some earlier papers show that an increase in global risk or risk aversion as measured by the VIX equity-option implied volatility index coincides with reductions in speculator carry positions (unwind) and carry trade losses (Brunnermeier et al., 2009; Menkhoff et al., 2012; Egbers and Swinkels, 2015). We will therefore test whether VIX is also an appropriate regime switch indicator.

The VIX is only available from 1990 onwards.[22] We again sacrifice 5 years of data to determine the 'high' versus 'low' indicator (see Section 2.2.2), which effectively makes the series start in 1995. The correlation coefficient is only 0.42, which means that VIX cannot be readily used as a pure proxy for FX volatility. Let us consider what happens if one nevertheless applies VIX as the regime switch indicator. Table 7 shows the results.

From the results in Table 7 we conclude that using VIX as the regime switch indicator yields qualitatively similar results as using FX volatility as the switch indicator (compare Panels A and B of Table 7). The Panel B results are only modestly worse than the Panel A results, which suggests that VIX might also be used as a switch indicator. Panel C shows similar results as Panel A, which implies that combining FX volatility and VIX does not add to the performance.

3.8. Impact of the Lehman crisis

Taking Fig. 3 into consideration it seems that the Lehman crisis has the largest impact, and one can thus wonder to what extent our overall results are driven by this event alone. Lehman Brothers filed for Chapter 11 bankruptcy protection on

[21] The Chicago Board of Options Exchange (CBOE) volatility index, or VIX index, is a popular indicator for measuring investor risk aversion, see Brunnermeier et al. (2009), Clarida et al. (2009), Christiansen et al. (2011).

[22] VIX data is available at http://www.cboe.com/micro/vix/historical.aspx.

Table 7
Impact of the regime switch indicator. Table shows the impact of the regime switch indicator on the performance measured by annualized mean return in % $\overline{R}$, annualized standard deviation in % S.D., skewness Skew., kurtosis Kurt., Sharpe ratio S, and holding period return (end-of-period accumulated wealth of $1 investment) HPR for the conventional CT strategy (base-case) and our three conditional CT strategies. See Section 2 or Fig. 1 for the implementation of the CT strategies. Panel A displays the performance using FX volatility (EWMA) as regime switch indicator, panel B applies the forward-looking VIX index as regime switch indicator, and panel C reports the findings when both regime switch indicators are used. The monthly data set covers 25 currencies spanning the 1995 to mid-2015 period. All economic outperformances relative to the base-case strategy (higher $\overline{R}$, lower S.D, higher Sharpe, less negative Skew, lower Kurt, and higher HPR) are highlighted in gray. Statistical significance on the basis of HAC estimators is indicated by * and ** respectively (90% and 95% confidence intervals). See Section 3.11 for a discussion on statistical significance testing.

	Panel A: FX volatility (EWMA) is regime switch indicator					
	$\overline{R}$	S.D.	Skew.	Kurt.	Sharpe	HPR
Conventional	5.34	10.72	−1.54	7.81	0.50	2.63
CONDFXIRD	6.75	10.55	−1.16	8.50	0.64*	3.53
CONDFX	6.74	8.63	−0.68*	4.95*	0.78	3.66
CONDIRD	5.00	10.11	−1.33	9.49	0.49	2.49
	Panel B: VIX is regime switch indicator					
Conventional	5.34	10.72	−1.54	7.81	0.50	2.63
CONDFXIRD	5.97	10.63	−1.42	9.68*	0.56	3.00
CONDFX	5.88	8.81**	−1.21	8.49	0.67	3.06
CONDIRD	5.00	10.11	−1.33	9.49	0.49	2.49
	Panel C: Both FX volatility and VIX is regime switch indicator					
Conventional	5.34	10.72	−1.54	7.81	0.50	2.63
CONDFXIRD	6.70	10.62	−1.13	8.27	0.63*	3.48
CONDFX	6.79	8.54	−0.62*	4.89*	0.80	3.70
CONDIRD	5.00	10.11	−1.33	9.49	0.49	2.49

September 15, 2008 but Melvin and Taylor (2009) show that the carry trade returns already started to decline in the summer of 2008. Therefore, as a robustness check we exclude the subperiods from July 2008 until respectively October 2008, December 2008, and June 2009. Table 8 shows the results. Our main conclusions still hold. The CONDIRD strategy, that conditions on extreme IRDs only, outperforms the base-case for some settings but underperforms for others. The CONDFXIRD and CONDFX strategies, that condition on both FX volatility and extreme IRDs and on solely FX volatility respectively, outperform the conventional CT strategy in all cases.

3.9. Number of long and short positions

In our main analysis, we have used 2 long and 2 short positions (i.e. [+2, −2]) in all our CT strategies. Table 9 report the results for portfolios from [+1, −1] up to [+5, −5]. As a change in the number of positions affect all CT strategies the results are reported for all CT strategies. The performance of all strategies are qualitatively in line with all earlier results, i.e. the CONDFXIRD and CONDFX strategies outperform the base-case in virtually all settings while the CONDIRD strategy shows mixed performance. However, the difference between both the CONDFXIRD and CONDFX strategies relative to the base-case are more pronounced for the [+1, −1] and [+2, −2] strategies than for the strategies that have a bigger number of long and short positions. The 'problem' here is that the larger the number of short and long positions, the more diversified the base-case conventional CT portfolio (and also the lower its return). As a consequence, the (lowered) returns of the conventional CT strategy also become more stable, which in turn means that our volatility indicator (based on the volatility of the conventional CT returns) increasingly fails to pick up the 'high volatility' regimes. We conclude that, when using the current FX volatility indicator, conditioning on extreme IRDs during high FX volatility regimes works best for smaller CT portfolios.

3.10. Choice of currencies

In this subsection we test the sensitivity of our results to the choice of currencies. On the basis of Table 1, one can conclude our total results may heavily depend on a few currencies only. For example, the Icelandic Krona seems by far the most favourite long position, just as the most popular short positions are the Japanese Yen, and the Swiss Franc. In Panels B, C, and D of Table 10 we test whether our main results are robust to leaving out one of these currencies at a time. In addition, we also use a totally restricted sample using the so-called G10 currencies which are the most liquidly traded currencies (see e.g.,

Table 8
Impact of the Lehman crisis. Table 8 shows the impact of the Lehman crisis on the performance measured by annualized mean return in % $\overline{R}$, annualized standard deviation in % S.D., skewness Skew., kurtosis Kurt., Sharpe ratio S, and holding period return (end-of-period accumulated wealth of $1 investment) HPR for the conventional CT strategy (base-case) and our three conditional CT strategies. See Section 2 or Fig. 1 for the implementation of the CT strategies. Panel A displays the results of our main analysis for the ease of comparison. Panels B, C, and D report the results when respectively July 2008-October 2008, July 2008-December 2008, and July 2008-June 2009 are excluded from the sample. The monthly data set covers 25 currencies spanning the 1985 to mid-2015 period. All economic outperformances relative to the base-case strategy (higher $\overline{R}$, lower S.D, higher Sharpe, less negative Skew, lower Kurt, and higher HPR) are highlighted in gray. Statistical significance on the basis of HAC estimators is indicated by * and ** respectively (90% and 95% confidence intervals). See Section 3.11 for a discussion on statistical significance testing.

	Panel A: Full sample					
	$\overline{R}$	S.D.	Skew.	Kurt.	Sharpe	HPR
Conventional	4.74	10.14	−1.60	8.41	0.47	3.59
CONDFXIRD	5.88*	9.98	−1.16*	8.36	0.59**	5.11*
CONDFX	6.00	7.67**	−0.74**	5.95	0.78	5.64
CONDIRD	4.40	9.75	−1.29	8.88	0.45	3.28
	Panel B: Excluding July 2008-October 2008					
Conventional	5.78	9.44	−1.21	6.61	0.61	4.96
CONDFXIRD	6.96*	9.27	−0.59**	5.44*	0.75*	7.08*
CONDFX	6.07	7.71**	−0.74	5.90	0.79	5.64
CONDIRD	5.25	9.02	−0.69*	5.81	0.58	4.27
	Panel C: Excluding July 2008-December 2008					
Conventional	6.08	9.20	−1.10	6.35	0.66	5.40
CONDFXIRD	6.92	9.20	−0.63**	5.51	0.75	6.95
CONDFX	6.22	7.76**	−0.74	5.82	0.80	5.85
CONDIRD	5.34	8.99	−0.71	5.88	0.59	4.36
	Panel D: Excluding July 2008-June 2009					
Conventional	5.69	9.20	−1.12	6.36	0.62	4.69
CONDFXIRD	6.49*	8.96	−0.87**	5.23	0.72	5.97*
CONDFX	6.06	7.79**	−0.74	5.79	0.78	5.42
CONDIRD	4.65	8.68	−1.03	5.56	0.54	3.50

Christiansen et al., 2011; Clarida et al., 2009; Jordà and Taylor, 2012; Ready et al., 2013).[23] Table 10 reports the findings for the G10 currencies in panel E. For the ease of comparison, panel A shows the results for the set of currencies used in our main analysis. Again, the CONDFXIRD strategy outperforms the base-case in almost all cases. The CONDFX also outperforms the base-case, while the CONDIRD strategy shows mixed results in comparison to the conventional CT strategy.

3.11. Statistical significance of our results

We follow Roon et al. (2012) and Ledoit and Wolf (2008) to check the statistical significance of the difference in performance of our conditional CT strategies versus the base-case. Ledoit and Wolf (2008) test for the difference in Sharpe ratios while Roon et al. (2012) do the same but also for other performance measures (standard deviation, skewness, and kurtosis). We report the regular p-values, the p-values using a heteroscedasticity and autocorrelation consistent (HAC) covariance matrix, and the p-values using a pre-whitened HAC covariance matrix in Table 11.[24]

The return of CONDFXIRD is statistically different from the base-case at the 10% level for all tests while the skewness and the Sharpe ratio is significantly different at resp. 10% and 5% when the HAC and pre-whitened HAC are used. The standard deviation and kurtosis are statistically insignificant for CONDFXIRD. For CONDFX we have that standard deviation and skewness are significantly different at 5% for all tests while the Sharpe ratio is significant at 5% for the statistical significance test using regular p-values, but insignificant when the HAC and pre-whitened HAC are used. The return and kurtosis are insignificant across all tests.

For the robustness checks we only focussed on the HAC estimators since our data displays serial correlation (and regular p-values would thus be biased). As the results for the HAC and pre-whitened HAC estimators are qualitatively the same we only report statistical significance based on the HAC estimators. We first reflect on the statistical significance of the CON-

[23] These countries are Australia, Canada, Germany/eurozone, Japan, New Zealand, Norway, Sweden, Switzerland, the United Kingdom, and the United States.

[24] We have implemented the tests in Matlab and have used part of the Matlab software provided by Ledoit and Wolf (2008) (available at http://www.econ.uzh.ch/en/people/faculty/wolf/publications.html) to estimate the covariance matrix denoted by Ω in Roon et al. (2012) and by Ψ in Ledoit and Wolf (2008). Both HAC and pre-whitened HAC estimators are based on the Parzen-Gallant kernel; see Ledoit and Wolf (2008) and references therein for further details.

Table 9
Impact of the number of long and short positions. Table shows the impact of the number of long and short positions on the performance measured by annualized mean return in % $\overline{R}$, annualized standard deviation in % S.D., skewness Skew., kurtosis Kurt., Sharpe ratio S, and holding period return (end-of-period accumulated wealth of \$1 investment) HPR for the conventional CT strategy (base-case) and our three conditional CT strategies. See Section 2 or Fig. 1 for the implementation of the CT strategies. Panels A through E display the results for various [long, short] positions. The monthly data set covers 25 currencies spanning the 1985 to mid-2015 period. All economic outperformances relative to the base-case strategy (higher $\overline{R}$, lower S.D, higher Sharpe, less negative Skew, lower Kurt, and higher HPR) are highlighted in gray. Statistical significance on the basis of HAC estimators is indicated by * and ** respectively (90% and 95% confidence intervals). See Section 3.11 for a discussion on statistical significance testing.

	$\overline{R}$	S.D.	Skew.	Kurt.	Sharpe	HPR
Panel A: 1 long/1 short positions [+1,-1]						
Conventional	6.70	13.97	−1.61	10.29	0.48	5.60
CONDFXIRD	8.34	12.59	−1.18	7.87*	0.66*	9.79
CONDFX	8.63	10.16**	−0.43**	5.98*	0.85*	11.70
CONDIRD	6.74	12.69	−1.28	8.03	0.53	6.01
Panel B: 2 long/2 short positions [+2,-2]						
Conventional	4.74	10.14	−1.60	8.41	0.47	3.59
CONDFXIRD	5.88*	9.98	−1.16*	8.36	0.59**	5.11*
CONDFX	6.00	7.67**	−0.74**	5.95	0.78	5.64
CONDIRD	4.40	9.75	−1.29	8.88	0.45	3.28
Panel C: 3 long/3 short positions [+3,-3]						
Conventional	3.65	8.81	−1.71	9.27	0.41	2.69
CONDFXIRD	3.97	8.17**	−1.27**	8.17	0.49	3.01
CONDFX	4.48	6.09**	−0.79**	5.97	0.73	3.68
CONDIRD	3.42	8.02**	−1.24**	8.38	0.43	2.56
Panel D: 4 long/4 short positions [+4,-4]						
Conventional	3.78	7.72	−1.26	7.38	0.49	2.87
CONDFXIRD	3.80	7.17**	−0.86**	6.37	0.53	2.93
CONDFX	3.99	5.47**	−0.52*	5.23	0.73	3.21
CONDIRD	3.05	7.18*	−0.85**	6.38	0.43	2.34
Panel E: 5 long/5 short positions [+5,-5]						
Conventional	3.86	6.97	−1.14	6.76	0.55	3.00
CONDFXIRD	4.03	6.74	−0.62**	4.96	0.60	3.18
CONDFX	3.40	5.18**	−0.48	5.46	0.66	2.70
CONDIRD	3.31	6.77	−0.70	5.25	0.49	2.55

DFXIRD strategy for all the settings tested, after which we discuss the CONDFX strategy. Results for the CONDIRD strategy are not discussed (nor have they been reported) as that strategy does not show economic outperformance relative to the conventional CT strategy.

3.11.1. Statistical significance for CONDFXIRD

The economic outperformance of the CONDFXIRD strategy in Table 3 shows comparable statistical results as reported in Table 2, being that the return, skewness, and Sharpe ratio are statistically significant, mostly at the 10% level. The results in Table 4 show that the definition of ‘extreme’ IRDs impacts statistical significance: while for absolute, positive, and time period extreme IRDs the results are comparable to those reported in Table 2, but for negative and currency specific extreme IRDs the results are statistically insignificant. When it comes to the choice for an FX volatility threshold (see Table 5), we notice that statistical significance is only obtained for the 0.8 and 0.9 threshold values; for the 0.95 and 0.99 thresholds the CONDFXIRD strategy almost overlaps (in terms of number of observations) with the base case, and as a consequence its economic outperformance hardly has any statistical significance. For the lower FX volatility thresholds (in our testing, 0.75 and less) the economic outperformance seems to be mainly driven by the highest FX volatility observations and not by the observations for which FX volatility is merely higher than the sample median FX volatility. For our other robustness checks, the economic outperformance of our CONDFXIRD strategy has similar statistical significance as the main results reported in Table 2 for high values of the EWMA decay parameter λ (see Table 6), as well as in the case of the reduced sample period (that avoided the Lehman crisis, see Table 8) which excluded the July to October 2008 period, and also in case of the alternative samples that excluded the most ‘popular’ currencies as reported in Table 10 (i.e., those with the highest number of long or short positions as reported in Table 1). We did not find statistical significance for the economic outperformance of our CONDFXIRD strategy in the other cases.

Table 10
Impact of the choice of currencies. Table shows the impact of the choice of currencies on the performance measured by annualized mean return in % $\overline{R}$, annualized standard deviation in % S.D., skewness Skew., kurtosis Kurt., Sharpe ratio S, and holding period return (end-of-period accumulated wealth of $1 investment) HPR for the conventional CT strategy (base-case) and our three conditional CT strategies. See Section 2 or Fig. 1 for the implementation of the CT strategies. Panel A displays the results of our main analysis for the ease of comparison which uses the monthly data set covering 25 currencies spanning the 1985 to mid-2015 period. Panels B through E report the results when using a restricted number of currencies spanning the same 1985 to mid-2015 period. All economic outperformances relative to the base-case strategy (higher $\overline{R}$, lower S.D, higher Sharpe, less negative Skew, lower Kurt, and higher HPR) are highlighted in gray. Statistical significance on the basis of HAC estimators is indicated by * and ** respectively (90% and 95% confidence intervals). See Section 3.11 for a discussion on statistical significance testing.

	$\overline{R}$	S.D.	Skew.	Kurt.	Sharpe	HPR
Panel A: Main results						
Conventional	4.74	10.14	−1.60	8.41	0.47	3.59
CONDFXIRD	5.88*	9.98	−1.16*	8.36	0.59**	5.11*
CONDFX	6.00	7.67**	−0.74**	5.95	0.78	5.64
CONDIRD	4.40	9.75	−1.29	8.88	0.45	3.28
Panel B: Full sample excl. Iceland						
Conventional	5.33	9.91	−1.30	8.99	0.54	4.32
CONDFXIRD	6.02	9.43	−0.86*	6.12*	0.64**	5.42
CONDFX	5.44	7.47**	−0.87	5.94	0.73	4.79
CONDIRD	4.80	9.03**	−0.80**	6.26*	0.53	3.79
Panel C: Full sample excl. Japan						
Conventional	3.04	8.79	−1.63	9.06	0.35	2.24
CONDFXIRD	4.07*	8.36	−0.95**	7.28**	0.49**	3.09*
CONDFX	4.17	6.06**	−0.69**	7.51	0.69*	3.35
CONDIRD	2.94	8.46	−0.94**	7.18*	0.35	2.19
Panel D: Full sample excl. Switzerland						
Conventional	3.95	10.28	−1.33	7.66	0.38	2.82
CONDFXIRD	5.25**	10.12	−0.96	7.71	0.52**	4.20**
CONDFX	4.94	7.73**	−0.40**	4.81**	0.64	4.09
CONDIRD	3.77	10.09	−1.08	7.97	0.37	2.68
Panel E: G10 currencies only						
Conventional	3.31	10.47	−0.90	6.64	0.32	2.31
CONDFXIRD	3.69	10.19**	−0.92	6.84	0.36	2.61
CONDFX	3.28	9.05**	−0.73	4.80	0.36	2.39
CONDIRD	2.81	9.68**	−0.74	6.99	0.29	2.03

Table 11
Statistical significance. Table shows the statistical significance of our CT strategies. All data (except for the pre-whitened HAC p-values) stem from Table 2. Regular p-values in regular parentheses, p-values of HAC estimators in brackets, and p-values of pre-whitened HAC estimators in curly brackets. Results for the CONDIRD strategy are not reported as it does not perform better than the conventional CT strategy. The holding period return (HPR) is not reported as its significance is equivalent to the significance of the CT return.

	$\overline{R}$	S.D.	Skew.	Kurt.	Sharpe
Conventional	4.74	10.14	−1.60	8.41	0.47
CONDFXIRD	5.88	9.98	−1.16	8.36	0.59
	(0.09)	(0.72)	(0.21)	(0.97)	(0.11)
	[0.06]	[0.67]	[0.08]	[0.96]	[0.04]
	{0.07}	{0.68}	{0.06}	{0.96}	{0.05}
CONDFX	6.00	7.67	−0.74	5.95	0.78
	(0.29)	(0.00)	(0.02)	(0.10)	(0.04)
	[0.39]	[0.02]	[0.04]	[0.12]	[0.11]
	{0.45}	{0.02}	{0.05}	{0.15}	{0.13}

Table 12
Out-of-sample performance. Table reports the same metrics as in Table 11, however, this time with an extended sample period (ending January 2017, instead of mid-2015). As in Table 11, we also show statistical significance of our CT strategies. Regular p-values in regular parentheses, p-values of HAC estimators in brackets, and p-values of pre-whitened HAC estimators in curly brackets. Results for the CONDIRD strategy are not reported as it does not perform better than the conventional CT strategy. The holding period return (HPR) is not reported as its significance is equivalent to the significance of the CT return.

	$\overline{R}$	S.D.	Skew.	Kurt.	Sharpe
Conventional	5.38	10.01	−1.60	8.54	0.54
CONDFXIRD	6.47	9.86	−1.17	8.45	0.66
	(0.09)	(0.71)	(0.21)	(0.96)	(0.11)
	[0.06]	[0.66]	[0.07]	[0.93]	[0.05]
	{0.07}	{0.67}	{0.05}	{0.93}	{0.05}
CONDFX	6.58	7.65	−0.71	5.90	0.86
	(0.29)	(0.00)	(0.01)	(0.08)	(0.04)
	[0.39]	[0.02]	[0.04]	[0.11]	[0.10]
	{0.45}	{0.02}	{0.05}	{0.13}	{0.11}

3.11.2. Statistical significance for CONDFX

The CONDFX strategy, that solely conditions on FX volatility regimes, the economic outperformance (which is obtained in virtually any alternative setting) has almost always a similar statistical significance as reported in Table 2, except for the choice of regime switch indicator (see Table 7) and the restricted sample that avoids the Lehman crisis (see Table 8) where only its standard deviation or skewness (but not both simultaneously) are statistically significant.

3.12. Out-of-sample performance of our results

In the working-paper version of our paper, we used data up to mid-2015. In revising our paper, we realized we had the unique opportunity to do an out-of-sample robustness check of our results by adding new data. Hence, as a last robustness check of our paper, we stretch the sample period up to January 2017. The results (reported in Table 12) are qualitatively equal to the ones reported in Table 11, although slightly better for the extended sample.

4. Conclusions

This paper proposes to condition currency carry trade (CT) strategies such that they avoid regimes for which uncovered interest parity (UIP) is likely to hold. Prior research shows that UIP likely holds if absolute interest rate differentials (IRDs) are large, or when foreign exchange (FX) volatility is high. However, these prior findings concern contemporaneous effects that need not have predictive power.

We develop three conditional CT strategies, and analyze their performance relative to the base-case unconditional CT strategy. Our conditional CT strategies condition on extreme IRDs during high FX volatility regimes (CONDFXIRD), on FX volatility only (CONDFX), and on extreme IRDs only (CONDIRD). The CONDFXIRD strategy conditions on extreme IRDs during high FX volatility regimes, which means it truncates the investment opportunity set on an ex ante basis by ignoring the 10% extreme IRDs during periods when FX volatility is higher than the 90%-th quantile of the FX volatility distribution. A 5 year rolling window is used to determine the IRD distribution and the FX volatility distribution. The CONDFX strategy conditions on FX volatility only, and switches to the conventional CT strategy during low FX volatility regimes and to a passive (non-investment) strategy during high FX volatility regimes. The CONDIRD strategy conditions on the 10% extreme IRDs only, thereby ignoring the FX volatility regime.

Our main analysis is based on a CT portfolio that has 2 long and 2 short positions, which rebalances monthly. Using a data set of 25 currencies from 1985 to mid-2015, we find that both the CONDFXIRD and CONDFX strategies, that condition respectively on extreme IRDs during high FX volatility regimes and on FX volatility only, outperform the base-case in terms of mean return, standard deviation, skewness, kurtosis, Sharpe ratio, and holding period return in virtually any of the settings analyzed. Conditioning on very large IRDs only shows mixed findings. When varying the truncation percentile of IRDs (see Table 3 for the results), it appears that conditioning on IRDs during periods of high FX volatility consistently outperforms the base-case and peaks at about 5% truncation (see Panel B of Table 3). Conditioning on IRDs in any FX regime means that profitable investment opportunities are foregone. In our main analysis, we take absolute extreme IRDs for the entire panel of all currencies over the past 5 years (rolling window). When applying alternative definitions (see Table 4 for the results), the CONDFXIRD strategy consistently outperforms the base-case, but again the CONDIRD strategy shows mixed results. The CONDFX strategy does not truncate IRDs. In the main analysis we consider a regime as 'volatile' if the FX volatility exceeds the 90th percentile of the FX volatility distribution (determined over the prior 5-year rolling window). When varying this threshold from the 50th to the 99th percentile, the CONDFXIRD strategy consistently outperforms the base-case whilst peaking at the 80th percentile (see Table 5). The CONDFX strategy shows mixed results. When changing the definition of FX volatility (see Table 6 for the results), it appears that both the CONDFX and the CONDFXIRD strategy continue to outperform on virtually any dimension for all settings. The CONDIRD strategy does not switch on FX volatility regimes.

In our main analysis we condition on FX volatility. One could argue, however, that the carry trade is not driven by FX volatility but instead by a global appetite for risk-taking. It is often suggested to approximate that appetite for risk-taking by means of the VIX. It appears that our FX volatility measure has a correlation of 0.42 with the VIX. When applying the VIX as our switch indicator, it appears that the results are qualitatively comparable (see Table 7 for the results). Yet, prior research has shown that in times of high VIX, carry trades are loss-giving. This is not because UIP would hold, but simply because market liquidity dries up. Therefore, we have also analyzed the impact of jointly conditioning on FX volatility and VIX. In that restricted setting, our outperformance relative to the base-case is high.

The biggest crisis in our dataset is the Lehman crisis. Hence, as a robustness check we have left out some data to correct for it. Even though Lehman Brothers filed for Chapter 11 bankruptcy protection on September 15, 2008, the carry trade returns started to decline a little earlier. Therefore we have left out 4 months, 6 months, and 12 months, all starting from July 2008 onwards. Table 8 shows the results. It appears that leaving out the Lehman crisis does not qualitatively impact our findings.

Our main analysis works with CT portfolios that have 2 long and 2 short positions. It appears that if we change this to a [+1, −1] portfolio, that our outperformance becomes even stronger, whereas if we change to a [+3, −3] portfolio or even bigger, then our outperformance decreases. We interpret this result as follows. The larger the number of long and short positions in a CT portfolio, the larger the diversification effect, and thus the smaller the impact of our conditionings. For the CONDFX strategy, we can add here that (since it turns into a passive investment strategy during periods of 'high' FX volatility) the effect of 'not investing' (and foregoing income) is magnified for a larger portfolio, just as active investing in a period when you should not, has a bigger effect for the larger portfolios. In short: Our strategy works best for smaller CT portfolios.

We also assessed the impact of the choice of our currencies. One might argue that some of our currencies are less liquidly traded (e.g., the Czech kroner, or the Icelandic kroner). To tackle this concern, we rerun the entire analysis for the so-called G10 currencies only.[25] With less currencies in the dataset all returns, etc. decrease, but again our CONDFXIRD and CONDFX strategies outperform the base-case (except for CONDFXIRD having a slightly bigger negative skewness and higher kurtosis), whilst the CONDIRD strategy shows some outperformance only. These findings are not surprising, because the G10 currencies hardly offer real 'extreme' IRDs. Also, we analyzed to what extent our overall results were driven by the most frequently used currencies, namely the Japanese Yen, the Swiss Franc, and the Icelandic Krona (see Table 1 for the number of short and long positions). Though the results inevitably change when leaving out a currency, the outperformance of our conditional strategies relative to the conventional CT strategy remains.

Most carry trade papers focus on economic outperformance, and rarely test for statistical significance of their results. In Section 3.11 we tested the statistical significance of our outperformance. Although the CONDFXIRD and CONDFX strategies economically outperform the base-case in virtually any setting, the statistical significance of the CONDFXIRD strategy's outperformance is mostly visible for the return, skewness, Sharpe ratio, and holding period return. The CONDFX strategy's outperformance is mostly statistically significant in terms of standard deviation and skewness. We attribute the limited statistical significance to the limited number of observations.

We conclude that conditioning carry trade strategies is a simple way to reduce risk, which comes at a very low cost, whilst the performance actually improves. Smaller CT portfolios benefit more from conditioning. Both the CONDFXIRD and CONDFX strategies outperform the conventional CT strategy, but the statistical significance of the CONDFXIRD strategy is best. Though our results can be improved even more by means of an exact parametrization of our FX volatility indicator and 'extreme' IRD definition, the purpose of our paper is not to calibrate these indicators to obtain the best-possible outperformances (for one data set). Instead, we show that conditioning carry trades on volatility and extreme IRDs improve their performance.

References

Baillie, R.T., Chang, S.S., 2011. Carry trades, momentum trading and the forward premium anomaly. J. Financ. Markets 14, 441–464.
Baldwin, R.E., 1990. Re-interpreting the Failure of Foreign Exchange Market Efficiency Tests: Small Transaction Costs, Big Hysteresis Bands. NBER Working Paper Series No. 3319. NBER.
Bansal, R., 1997. An exploration of the forward premium puzzle in currency markets. Rev. Financ. Stud. 10, 369–403.
Bansal, R., Dahlquist, M., 2000. The forward premium puzzle: different tales from developed and emerging economies. J. Int. Econ. 51, 115–144.
Barroso, P., Santa-Clara, P., 2015. Beyond the carry trade: optimal currency portfolios. Journal of Financial and Quantitative Analysis 50, 1037–1056.
Brunnermeier, M.K., Nagel, S., Pedersen, L.H., 2009. Chapter 5: Carry trades and currency crashes. In: Acemoglu, D., Rogoff, K., Woodford, M. (Eds.), NBER Macroeconomics Annual 2008. MIT Press, pp. 313–347.
Burnside, C., Eichenbaum, M., Kleshchelski, I., Rebelo, S., 2011. Do peso problems explain the returns to carry trade? Rev. Financ. Stud. 23, 581–588.
Burnside, C., Eichenbaum, M., Rebelo, S., 2008. Carry trade: The gains of diversification. Journal of the European Economic Association 6, 853–891.
Christiansen, C., Ranaldo, A., Söderlind, P., 2011. The time-varying systematic risk of carry trade strategies. J. Financ. Quant. Anal. 46, 1107–1125.
Clarida, R., Davis, J., Pedersen, N., 2009. Currency carry trade regimes: beyond the Fama regression. J. Int. Money Financ. 28, 1375–1389.
Curcuru, S., Vega, C., Hoek, J., 2010. Measuring carry trade activity. In: BIS (Ed.), Initiatives to Address Data Gaps Revealed by the Financial Crisis: Proceedings of the Fifth IFC Conference, Basel, 25–26 August 2010. BIS, pp. 436–453.

[25] The G10 currencies refer to the 10 most liquidly traded currencies.

De Roon, F., Eiling, E., Gerard, B., Hillion, P., 2012. Currency Risk Hedging: No Free Lunch. Unpublished Working Paper 2012/87/FIN. INSEAD.
Egbers, T., Swinkels, L., 2015. Can implied volatility predict returns on the currency carry trade? J. Banking Financ. 59, 14–26.
Engel, C., 1996. The forward discount anomaly and the risk premium: a survey of recent evidence. J. Empirical Financ. 3, 123–192.
Evans, M.D., 2010. Order flows and the exchange rate disconnect puzzle. J. Financ. Econ. 80, 58–71.
Evans, M.D., Lyons, R.K., 2002a. Informational integration and FX trading. J. Int. Money Financ. 21, 807–831.
Evans, M.D., Lyons, R.K., 2002b. Order flow and exchange rate dynamics. J. Polit. Econ. 110, 170–180.
Galati, G., Heath, A., McGuire, P., 2007. Evidence of carry trade activity. BIS Quart. Rev. September. BIS.
Hochradl, M., Wagner, C., 2010. Trading the forward bias: are there limits to speculation? J. Int. Money Financ. 29, 423–441.
Hodrick, R.J., 1987. The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Markets. Academic Publishers, Chur, Switzerland and Harwood, New York.
Huisman, R., Koedijk, K., Kool, C., Nissen, F., 1998. Extreme support for uncovered interest parity. J. Int. Money Financ. 17, 211–228.
Jordà, O., Taylor, A.M., 2012. The carry trade and fundamentals: nothing to fear but FEER itself. J. Int. Econ. 88, 74–90.
Kohler, M., 2010. Exchange Rates During Financial Crises. BIS Quarterly Review March. BIS.
Koijen, R.S., Moskowitz, T.J., Pedersen, L.H., Vrugt, E.B., 2018. Carry. J. Financ. Econ. 127, 197–225.
Ledoit, O., Wolf, M., 2008. Robust performance hypothesis testing with the Sharpe ratio. J. Empirical Financ. 15, 850–859.
Lewis, K.K., 1995. 37: Puzzles in international financial markets. In: Grossman, G.M., Rogoff, K. (Eds.), Handbook of International Economics, vol. 3. Elsevier, Amsterdam, pp. 1913–1971.
Lothian, J.R., Wu, L., 2011. Uncovered interest-rate parity over the past two centuries. J. Int. Money Financ. 30, 448–473.
Lustig, H., Roussanov, N., Verdelhan, A., 2011. Common risk factors in currency markets. Rev. Financ. Stud. 24, 3731–3777.
Lustig, H., Verdelhan, A., 2007. The cross section of foreign currency risk premia and consumption growth risk. Am. Econ. Rev. 97, 89–117.
Lyons, R.K., 2001. The Microstructure Approach to Exchange Rates. MIT Press, Cambridge, MA.
Melvin, M., Taylor, M.P., 2009. The crisis in the foreign exchange market. J. Int. Money Financ. 28, 1317–1330.
Menkhoff, L., Sarno, L., Schmeling, M., Schrimpf, A., 2012. Carry trades and global foreign exchange volatility. J. Financ. 67, 681–718.
Ready, R., Roussanov, N., Ward, C., 2013. Commodity Trade and the Carry Trade: A Tale of Two Countries. NBER Working Paper Series No. 19371. NBER.
Sager, M., Taylor, M.P., 2008. Commercially available order flow data and exchange rate movements: Caveat emptor. J. Money Credit Banking 40, 583–625.
Sarno, L., 2005. Viewpoint: towards a solution to the puzzles in exchange rate economics: where do we stand? Can. J. Econ. 38, 673–708.
The Economist, 2006. Instant Returns: Why Investors Have Become Addicted to the Carry Trade. 05 October 2006.
Villanueva, M.O., 2007. Forecasting currency excess returns: can the forward bias be exploited? J. Financ. Quant. Anal. 42, 963–990.
Wu, Y., Zhang, H., 1996. Asymmetry in forward exchange rate bias: a puzzling result. Econ. Lett. 50, 407–411.