Good Carry, Bad Carry *

Geert Bekaert[a†] George Panayotov[b‡]

[a]*Columbia Business School, New York, NY 10027, USA. NBER ....*

[b]*School of Business, Hong Kong University of Science and Technology, Clearwater Bay, Hong Kong*

This version, January 16, 2018

Abstract

We distinguish between "good" and "bad" carry trades constructed from G-10 currencies. The good trades exhibit higher Sharpe ratios and sometimes positive return skewness, in contrast to the bad trades that have both substantially lower Sharpe ratios and highly negative return skewness. Surprisingly, good trades do not involve the most typical carry currencies like the Australian dollar and Japanese yen. The distinction between good and bad carry trades significantly alters our understanding of currency carry trade returns, and invalidates, for example, explanations invoking return skewness and crash risk.

KEY WORDS: currency carry trade, predictability, currency risk factors

JEL CLASSIFICATION CODES: C23, C53, G11,

*Comments, including references to related papers that we may have inadvertently overlooked are welcome. We acknowledge many fruitful suggestions by Raymond Kan and discussions with Gurdip Bakshi, Utpal Bhattacharya, Abhiroop Mukherjee, Steve Riddiough, Giorgio Valente and Jialin Yu, who also helped us obtain some of the data used in the paper. Participants at the 2016 Annual Conference in International Finance and the Bank of England, Banca dItalia and ECB 6th Workshop on Financial Determinants of Foreign Exchange Rates, and seminars at Nanyang Technology University, UNSW, SAIF and Xiamen University provided helpful insights. Any remaining errors are our responsibility alone.

†*E-mail address:* gb241@columbia.edu

‡*E-mail address:* panayotov@ust.hk

This article was published in *Journal of Financial and Quantitative Analysis,* and is available in its final form at https://doi.org/10.1017/S0022109019000887

1 Introduction

The currency carry trade, which goes long (short) currencies with high (low) yields continues to attract much research attention, as it has been shown to earn high Sharpe ratios, while its returns are largely uncorrelated with standard systematic risks (e.g., Burnside, Eichenbaum, Kleshchelski, and Rebelo (2011), Lustig, Roussanov, and Verdelhan (2011)). Prototypical carry currencies among the liquid G-10 currencies are the Swiss franc (CHF) and Japanese yen (JPY), which almost always exhibit the lowest yields and hence a typical G-10 carry trade would short them, and the New Zealand dollar (NZD) and Australian dollar (AUD), which typically have the highest yields and would be held long. These four currencies feature prominently in extant explanations of the returns of the carry trade. One such explanation invokes crash risk (e.g. Brunnermeier, Nagel, and Pedersen (2009)) and thus relies implicitly on the fact that JPY and AUD provide the most "skewed" return perspectives from the view point of a US investor. Another is based on the differential exposure to global productivity shocks of producers of final goods, such as Japan and Switzerland, versus commodity producers, such as Australia and New Zealand (Ready, Roussanov, and Ward (2017)). In a similar spirit, Verdelhan (2017) states that "... the Japanese yen appreciates in bad times, while the Australian dollar depreciates, and this difference is at the heart of any risk-based explanation of carry trades".

While the prior literature takes for granted that the prototypical carry currencies indeed drive carry trade profitability, in this article we document the existence of "good" and "bad" currency carry trades. First, we consider an investor who sequentially tests whether reducing the set of G-10 currencies improves the historical Sharpe ratio, and then implements equally weighted carry trades with fewer currencies. We find that such trades from fewer currencies improve the return profile (in terms of *both* Sharpe ratio and skewness) relative to the carry trade which employs all G-10 currencies, and denote them as "good" carry trades. Most surprisingly, these good trades almost never include the AUD and JPY, or the NOK - another commodity currency. Following this insight, we then construct carry trades using fixed subsets of the G-10 currencies over the full sample. We find that trades that focus only on prototypical currencies tend to have much lower Sharpe ratios and more negatively skewed returns, and denote them as ”bad” carry trades. The trades using the remaining currencies preserve the desirable features of ”good” carry trades. Our results indicate that much of the recent literature attempting to explain carry trade returns focusing on prototypical carry currencies may be misguided.

Providing a first glimpse on the issue, Figure 1 contrasts the return properties of carry trades that involve various subsets of the G-10 currencies. In particular, the figure plots (with black dots) skewness versus Sharpe ratio for all carry trades constructed from *five* currencies that use three of the prototypical currencies (AUD, CHF and JPY), together with any possible pair from the remaining seven currencies. The currencies enter each trade with equal weights, as is common in the literature and finance industry. Strikingly, these 21 trades show worse Sharpe ratios (from at best two thirds to slightly negative), and also substantially lower skewness (three to five times more negative) than the strategy that uses all G-10 currencies. Therefore, trades constructed predominantly from the prototypical carry currencies appear to be ”bad” carry trades. We subsequently refer to the trade from all G-10 currencies as ”standard carry” and denote it as SC.

[Figure 1 about here.]

Probing further, Figure 1 also displays (with unfilled circles) the skewness versus Sharpe ratio of the complements of the previous 21 carry trades, which are constructed with the remaining five currencies in each case, again with equal weights. It is noteworthy that 14 out of the 21 complement trades feature higher Sharpe ratios than that of the standard carry (SC) trade (in one case almost double that ratio), and 16 show higher (less negative or positive) skewness. Furthermore, half of the complement trades improve *both* on the skewness and Sharpe ratio of the SC trade, qualifying them as ”good” carry trades. These findings cast doubt on efforts to explain carry trade returns by focusing on properties of the prototypical carry currencies, and undermine the practice of associating the carry trade predominantly with such currencies.

In Section 4 we investigate the ability of good carry trades to function as risk factors for certain cross sections of currency returns (see, e.g. Lustig, Roussanov, and Verdelhan (2011) and Menkhoff, Sarno, Schmeling, and Schrimpf (2012)). We find that good carry trades perform at least as well as previously suggested currency market risk factors, and sometimes drive out such factors in a horse race. We also re-examine the predictability findings in Bakshi and Panayotov (2013) and Ready, Roussanov, and Ward (2017), and find that previously identified carry return predictors strongly predict the returns of bad, but not of good carry trades. In Section 5 we revisit several interpretations of carry returns that have been advanced in the recent literature, including the explanatory ability of factor models with equity market risk factors, a crash risk explanation of their returns as in Brunnermeier, Nagel, and Pedersen (2009), and the peso problem hypothesis of Burnside, Eichenbaum, Kleshchelski, and Rebelo (2011). Almost invariably, the results differ greatly across good versus bad carry trades.

In Section 6, we further explore the properties of good and bad carry trades to kindle research on economic models that may explain the apparently strong differences between the two types of trades. We show, for example, that the returns of good trades are still mostly driven by carry, but also derive part of their returns from exchange rate changes, whereas the returns of bad trades are eroded by exchange rate changes. We also examine the relationship between good carry trades and the ”dollar carry” factor introduced in Lustig, Roussanov, and Verdelhan (2014), which goes long (short) all currencies relative to the US dollar when the average foreign interest rate differential relative to the dollar is positive (negative). Because our good carry trades always involve the dollar, some of their returns can be traced to dollar exposure. However, while they do have substantial return correlation with dollar carry, they clearly present a distinct currency and economic risk, as we demonstrate.

Before introducing ”good” and ”bad” carry trades in Section 3, we describe the data in Section 2 and discuss some important concepts regarding the design of carry trades. The remainder of the article demonstrates how the good-bad trade distinction fundamentally alters our thinking about carry trades.

2 Data and carry trade design

Following previous work, we employ currency spot and forward contract quotes to construct carry trade returns. Using one-month forward quotes on the last trading day of each month in the sample, and spot quotes on the last day of the following month, we calculate one-month carry trade returns over the sample period from 12/1984 till 06/2014 (354 monthly observations). The return calculations take into account transaction costs, exploiting the availability of bid and ask quotes. The data comes from Barclays Bank, as available on Datastream, and have been used in Burnside, Eichenbaum, Kleshchelski, and Rebelo (2011) and Lustig, Roussanov, and Verdelhan (2011, 2014), among many others.

Our results are reported for percentage returns and equal (absolute) weights of the currencies entering a trade. On two occasions we report instead results with logarithmic returns or weights proportional to forward differentials, to facilitate comparability with previous studies.

Let S_t^i denote the spot exchange rate of currency i at time t, quoted as foreign currency units per one U.S. dollar. That is, the U.S. dollar is the benchmark currency and all trades are implemented relative to the dollar. Similarly, let F_t^i denote the time t one-month forward exchange rate, quoted in the same way. Then, if t is the end of a given month, and $t+1$ is the end of the following month, the percentage excess one-month return at $t+1$ of one dollar invested at t in a long (short) forward foreign currency contract is:

$$Rx_{t+1}^{i,long} = F_t^{i,bid}/S_{t+1}^{i,ask} - 1 \qquad \text{and} \qquad Rx_{t+1}^{i,short} = 1 - F_t^{i,ask}/S_{t+1}^{i,bid}, \tag{1}$$

whereby bid and ask quotes are denoted in the superscript.[1]

[1]These can also be seen as the payoffs to forward contracts in the foreign currency per ”forward” dollar.

We employ the G-10 currencies, which are the New Zealand dollar (NZD), Australian dollar (AUD), British pound (GBP), Norwegian krone (NOK), Swedish krona (SEK), Canadian dollar (CAD), US dollar (USD), Euro (EUR), Swiss franc (CHF) and Japanese yen (JPY), whereby prior to 1999 the German mark (DEM) is used instead of the Euro. These currencies represent the most liquid traded currencies, and are most often used both in the academic literature and professional practice to construct carry trades.

As indicated above, the carry trades that we consider go long and short an equal number of currencies relative to the USD, with equal weights. Various alternative weighting schemes are possible, mostly based on the magnitude of the interest rate differentials (see Table 1 for concrete examples from practice and the academic literature), but we prefer to keep the trade as simple as possible. Moreover, the total investment each period is one dollar, that is, the sum of all long and short positions (in absolute value) equals one. Specifically, when the trade uses all G-10 currencies, the five currencies with the lowest interest rates are shorted, and the remaining five are held long. In practice, we rank the currencies based on their forward differentials relative to the U.S. dollar, defined as $FD_t = F_t/S_t - 1$ at time t and calculated using mid-quotes. The weight of currency i held long (short) is $\omega_t^i = \frac{1}{10}$ ($\omega_t^i = -\frac{1}{10}$). The percentage excess return of this trade from t to $t+1$ is:

$$Rx_{t+1}^{carry} = \sum_{i=1}^{10} \left\{ \mathbf{1}_{\omega_t^i > 0} \omega_t^i \, Rx_{t+1}^{i,long} - \mathbf{1}_{\omega_t^i < 0} \omega_t^i \, Rx_{t+1}^{i,short} \right\} \tag{2}$$

where $\mathbf{1}_{(.)}$ is an indicator function.

When a subset of N currencies is used to construct a carry trade and N is even, we set $\omega_t^i = \frac{1}{N}$ or $-\frac{1}{N}$ in (2) and substitute N for 10. If N is odd, the currency with the median forward differential is dropped from the trade, and we use $N-1$ instead of N in the definition of ω_t^i and the summation in (2). Finally, whether using all G-10 currencies or subsets of them, we re-balance the carry trades at the end of each month, and liquidate them at the end of the following month.

We follow the previous literature and consider carry trades that are *symmetric*, in that they have an equal number of short and long positions, with equal total weights on the long and short side. Currencies are ranked according to their interest rates, and only the rank determines whether the position taken is short or long, while the signs of the interest rate differentials are irrelevant for the trade design, as these change with the currency perspective (see also Clarida, Davis, and Pedersen (2009)). Importantly, our carry trade design also ensures (approximate) *numeraire independence*, as we do *not* give a special role to the benchmark currency, and hence the positions taken in the various participating currencies are the same, regardless of the benchmark. Numeraire independence is an attractive property, and implies that only one currency trade must be defined for the world at large. Moreover, the returns on such a trade are very similar from any currency perspective, because the translation from one currency to another simply introduces cross-currency risk on currency returns, which is a second-order effect. In fact, the logarithmic returns of our strategies are *exactly* the same from any perspective, by triangular arbitrage (see Maurer, Tô, and Tran (2016) for further discussion). Numeraire independence may also be a desirable trait for a global risk factor. The major commercial investable carry products delivered by the major players in the foreign exchange market, such as Deutsche Bank or Citibank, described in Table 1, are symmetric and numeraire-independent as per our definition. They do not all assign equal weights to all positions however, e.g. the well-known tradeable Deutsche Bank carry strategy takes only the three highest- and lowest-yielding currencies among the G-10 currencies.

Non-symmetric trade designs are also possible, and have been considered, for example in a recent well-recognized article by Burnside, Eichenbaum, Kleshchelski, and Rebelo (2011), where all currencies with interest rates that are higher (lower) than the US dollar interest rate are bought (sold) in equal proportions. Such a strategy is obviously not symmetric, and also may deliver very different results depending on the benchmark currency. Another example of a non-symmetric trade is the ”dollar carry” trade, studied in

Lustig, Roussanov, and Verdelhan (2014) and Hassan and Mano (2015).

In Appendix OA-I we discuss these issues more formally, as the particular design of carry trades affects the resulting return profiles. For example, when considering popular asymmetric strategies such as the one used in Burnside, Eichenbaum, Kleshchelski, and Rebelo (2011), the currency perspective matters greatly. Daniel, Hodrick, and Lu (2017) show that the USD perspective produces the most attractive Sharpe ratio by far. In addition, by creating a trade with equal weights and equal number of short/long positions, but without USD exposure, they further demonstrate that much of the return of an asymmetric carry trade can be attributed to its USD exposure. Such a trade is symmetric, but not numeraire-independent. We do not consider such ”dollar-neutral” trades, and our symmetric trades are *not* dollar-neutral: positions in the benchmark currency are explicitly included, even though the excess return for such positions is automatically equal to zero.

3 Good and bad carry trades: using sub-sets of the G-10 currencies

The symmetric carry trades that we consider go long high-yield currencies and short an equal number of low-yield currencies, with equal weights. Our interest is to explore the implications of excluding certain currencies altogether from the carry trade, or, equivalently, of constructing carry trades only from certain subsets of currencies.

In this section we first propose a disciplined approach to create symmetric carry trades from subsets of the G-10 currencies that have yielded attractive historical performance. The procedure is simple and exploits all available foreign exchange history at each point in time. It is implemented dynamically and hence yields out-of-sample results. Essentially, it evaluates on each trading date whether excluding currencies can improve on the standard carry trade (SC) that uses all G-10 currencies, and if so, which currencies should be excluded. Next, we exploit the information garnered in his exercise to create fixed subsets of ”good” and ”bad” currencies that do not change over time, and use them to construct carry trades over the full sample period.

3.1 Enhancing the currency carry trade

Imagine an investor starting to trade at the end of December 1994 ($t = T_1$). On this date, and at the end of each month going forward till May 2014 ($t = T_2$), he uses all available return information for the period since December 1984 ($t = T_0$) and first calculates the Sharpe ratio (denoted ”benchmark Sharpe ratio”) of the standard carry trade (SC) that employs all G-10 currencies. The trade ranks these currencies according to their forward differentials at the end of each month between T_0 and $t-1$, for $t = T_1, ..., T_2$, and goes long (short) over the following month the five currencies with the highest (lowest) forward differentials, all with equal weights.

To create an enhanced carry trade with nine currencies, on date t the investor excludes one by one each of the G-10 currencies, and computes the Sharpe ratios over T_0 to t of the ten possible trades that involve only nine currencies. These trades exclude the currency with the median forward differential at the end of each month between T_0 and $t-1$ and go long (short) the four currencies with the highest (lowest) differentials. If the highest of the ten Sharpe ratios obtained in this way exceeds the benchmark Sharpe ratio, an *enhanced trade* is implemented over the following month (t to $t+1$) using the nine currencies corresponding to this highest Sharpe ratio, while the one currency left out of the trade is the first to be excluded on date t. If, on the other hand, all ten Sharpe ratios are lower than the benchmark ratio, then no currency is excluded and the enhanced trade for this date has the return of the standard trade.

Note that the dynamic and real-time nature of this enhanced trade could, in principle, result in a substantially different currency mix used at different points of time. Further, our enhancement rule is intentionally simple and uses an easily understood and popular performance measure, whereas a wide range of other, more sophisticated optimization rules could be applied as well (see, e.g., Bilson (1984) and

Barroso and Santa-Clara (2014) on the use of optimization techniques for currency selection). Mimicking the construction of the enhanced trade that uses nine currencies, we proceed to design other enhanced trades that exclude more than one of the G-10 currencies in a similar manner. In particular, on a date t when one currency has been excluded, we use the remaining nine currencies to find the highest Sharpe ratio across the nine possible trades that involve only eight currencies. Again, if this highest ratio exceeds the benchmark Sharpe ratio, the currency that was omitted to achieve it is the second currency to be excluded for this date, whereas if all Sharpe ratios are lower than the benchmark one, no further currency is excluded and the enhanced trade that uses eight currencies has the return of the standard trade for the date. In the same way we attempt to exclude up to seven of the G-10 currencies on this date t, and thus obtain seven enhanced trades, which use a decreasing number of currencies. Importantly, we record the exact order in which currencies have been excluded. The above procedure is repeated on each date in the sample to obtain time series of returns of the seven enhanced carry trades.

For completeness of the search algorithm, we have postulated that if no improvement on the benchmark Sharpe ratio can be achieved for a certain date and number of excluded currencies, then no further currencies are excluded on this date, and all enhanced trades with fewer currencies have the return of the standard trade. In practice, however, this choice is inconsequential, since it turns out that improvement is possible on each trading date in our sample and for each number between one and seven of excluded currencies.

3.2 Return patterns for enhanced trades

Table 2 presents results for the enhanced carry trades that allow excluding from one to as much as seven currencies on each trading date. Returns are computed as described in Section 2, with equal currency weights and in percent. For comparability, we also show results for the standard carry trade (SC).

Panel A of Table 2 reports the annualized average returns, annualized Sharpe ratios, and return skewness for each carry trade (standard and enhanced). Interest in skewness is justified because occasional highly negative returns are an essential feature of currency carry trades, with possible implications for the understanding of their nature (e.g., Brunnermeier, Nagel, and Pedersen (2009), Jurek (2014)). Also reported are p-values for the hypothesis that the respective Sharpe ratio or skewness does not exceed that for the standard trade SC. We note here that by being explicitly concerned about return skewness we presume a possible departure from normality of the return distributions, and hence cannot rely on standard tests for the difference between Sharpe ratios, as for example in Jobson and Korkie (1981) or Memmel (2003). Therefore, we resort to bootstrap tests, and follow Ledoit and Wolf (2008) for Sharpe ratios, and Annaert, Van Osselaer, and Verstraete (2009) for skewness. Because the enhanced carry trades are designed with the goal to improve on SC, the reported p-values are based on one-sided bootstrap confidence intervals. Details on the bootstrap procedures are provided in Appendix OA-II.

The SC trade has an annualized Sharpe ratio of 0.32, and return skewness of -0.33. The benchmark Sharpe ratio is thus close, for example, to the value of 0.31 for the HML trade reported in Lustig, Roussanov, and Verdelhan (2014, Table 1) for their set of developed countries, over a similar sample period and using equal weighting and bid and ask quotes.

When one currency is excluded from the carry trade on each trading day, practically no change is observed: the Sharpe ratio is still 0.32 and skewness is -0.26. When two currencies are excluded, the Sharpe ratio increases to 0.41, but skewness drops to -0.57. When three to six currencies are excluded, the Sharpe ratios remain somewhat higher than the benchmark ratio (between 0.41 and 0.46), but the differences are not statistically significant. However, skewness improves sharply in three out of these four cases and turns positive on two occasions (and as high as 0.21 on one), whereby two of the associated p-values are below 5% and another one equals 10%. These findings indicate a possible two-dimensional beneficial effect of excluding three or more of the G-10 currencies, given that *both* the Sharpe ratio and skewness improve, albeit not always in a statistically significant way. This effect is further confirmed by the enhanced trade that excludes seven currencies: the Sharpe ratio is now 0.61, while skewness is positive, and both are marginally significantly different from the benchmark values.

Our findings echo previous results, where significant improvement in skewness is obtained with no or only a minor change in the Sharpe ratio (for example, the option-hedged carry trades in Burnside, Eichenbaum, Kleshchelski, and Rebelo (2011, Table 5). However, what is surprising in our case is that the improvement of the return profile is (i) in both dimensions at the same time, and (ii) achieved simply by excluding currencies from the symmetric carry trade. It is also noteworthy that the two-dimensional improvement is achieved by a procedure that maximizes the Sharpe ratios alone, with no explicit concern about the skewness of the returns so obtained.

3.3 Identity of the excluded currencies

While on each trading date the enhanced trade re-considers the available return history and thus can potentially deliver a different set of currencies to be excluded, we observe the same currencies to be excluded over and over, with amazing consistency. Panel B of Table 2 shows the number of months, over the 234-month sample period of enhanced trading, each G-10 currency is excluded by the enhancement rule from Section 3.1. In particular, it shows how many times the respective currency is the first, or among the first two, or among the first three, etc., to be excluded from the carry trade.

The consistency is observed most clearly with respect to the first three currencies excluded. Specifically, AUD is the first to be excluded on 135 out of the 234 trading dates in the sample. Furthermore, it is among the first three currencies to be excluded on a total of 192 dates. Similarly, NOK is among the first three excluded on 219 occasions, and JPY is among them on 214 occasions. It appears that these three currencies are found to be by far the most detrimental to carry trade Sharpe ratios by our enhancement rule - no other currency is ever excluded first, and practically only one other currency has been excluded second or third (EUR: 15 and 59 times out of 234, respectively).

The next currencies to be the most often excluded are the EUR, NZD, CAD and CHF, and while the order of their exclusion is somewhat ambiguous, these are still the obvious further candidates for exclusion by the enhancement rule. At the same time, the remaining three currencies are clearly found valuable by the rule: GBP and SEK are among the first seven to be excluded only on about 60 occasions each, and in fact are never among the first four excluded. Most conspicuously, however, the USD is never among even the first seven excluded currencies, possibly consistent with previous studies which discuss the special role of the USD in the carry trade (e.g., Lustig, Roussanov, and Verdelhan (2014), Daniel, Hodrick, and Lu (2017)) from various perspectives.

These findings are surprising, as the enhancement rule consistently excludes from the carry trade precisely the prototypical carry trade currencies, like the JPY and AUD, which have been perpetually among the lowest- or highest-yielding G-10 currencies, and used most commonly as examples in various carry trade discussions. Importantly, this consistency refers to the *entire* period since 1994, therefore the results are not driven, for example, by the recent financial crisis and the related drastic changes in certain exchange rates that are often cited to illustrate the risks inherent in the carry trade.

Two further observations can be made at this point. First, the surprise is not limited to the JPY and AUD. The fact that the enhancement rule also tends to exclude NOK and CHF is similarly unexpected, given that these have also been among the few highest- or lowest-yielding currencies over our sample period.

Second, the design of the enhancement procedure, as described in Section 3.1, leaves open the possibility that at some step no improvement of the Sharpe ratio can be achieved, whereby no further currencies are excluded on this date and the respective enhanced trades are assigned the SC return for the next trading period. This possibility is of some concern, as it could blur the distinction between enhanced trades that exclude a different number of currencies. However, Panel B in Table 2 reveals that this has never happened in our sample, as evidenced by the fact that the sum of the numbers in the first row equals 234, the sum of those on the second row equals 234 $\times$ 2, and so on. Therefore, on each date in the sample period the enhancement rule has identified seven currencies to be sequentially excluded, and hence seven distinct enhanced trades to be implemented.

In sum, the conclusion from Table 2 is that the enhancement rule consistently excludes the same few currencies from the carry trade, among which are those epitomizing the essential concept underlying carry trades that low (high) yield currencies should be sold (bought). The surprising evidence presented in Table 2 thus calls for a re-consideration of this concept and/or its implementation.

3.4 Good and bad carry trades from fixed subsets of the G-10 currencies

Prompted by the finding that the dynamic enhancement rule excludes the same currencies over and over, we now examine carry trades constructed with fixed subsets of the G-10 currencies. While staying close to the spirit of the enhanced trades, the fixed subsets allow for better comparison with previous carry trade results, which are similarly obtained using fixed sets of currencies over fixed sample periods. Our choice of the fixed subsets is informed by the order of exclusion implied by Panel B of Table 2, which showed that (i) the three currencies that are the least often excluded by the enhancement rule are the GBP, SEK and USD, (ii) the next three least often excluded are the CAD, NZD and CHF, whereas (iii) the AUD, NOK and JPY are the most often excluded currencies.

In particular, we construct five carry trades from fixed subsets, which (i) exclude only the AUD, NOK and JPY, (ii) include the GBP, SEK and USD, together with any of the three possible pairs from the CAD, NZD and CHF, and (iii) keep only the GBP, SEK and USD. These carry trades are designed to illustrate the properties of enhanced carry trades, and we denote them by G1 to G5, a notation we shall clarify shortly. The first column of Table 3 displays the codes of the currencies included in each of these five trades. We also consider the trades complementary to G1-G5, which include the currencies that are left out of each of these trades, and denote these complements by B1 to B5, respectively, with currency codes again displayed in the first column of Table 3. For example, only the three most often excluded currencies (AUD, NOK and JPY) enter the B1 carry trade.

In addition, we consider a larger set of trades which can represent more broadly the enhanced carry trades: it consists of 18 trades from five currencies each, and is denoted by GC, whereby each trade includes the three least often excluded currencies (GBP, SEK and USD), together with any possible pair from the remaining currencies which has *none or only one* of the three most often excluded currencies (AUD, NOK and JPY). This choice yields a reasonably large cross section of trades which maintains the predominant presence of currencies that are preferred by the enhancement rule. Again, we also consider the 18 complementary trades, and denote them by BC. Despite creating many carry trades from only five currencies, the average correlation among the returns of the 18 good trades is 0.66, and thus lower, for example, than the average correlation among the 25 value-weighted Fama-French portfolios sorted on size and book-to-market for the same period, which is 0.80.

Table 3 presents results for the SC trade, the G1-G5 and B1-B5 trades, and the GC and BC trades described above, using the entire sample period from 12/1984 till 6/2014. Shown are annualized average returns, return standard deviations and Sharpe ratios, as well as skewness. For the GC and BC trades we show *averages* of these quantities. Also reported are p-values for tests of differences between the Sharpe ratios and skewness coefficients, similar to those in Table 2. In the last two lines, the first (second) number in parentheses shows how many of the 18 corresponding individual estimates for the GC or BC trades are significant at the 5% (10%) level. Where p-values are not in square brackets, the null hypothesis is that the respective Sharpe ratio or skewness does not exceed the one of the SC trade. Where p-values are in square brackets, the null is that the Sharpe ratio or skewness of a G1-G5 trade or GC trade does not exceed that of the corresponding B1-B5 trade or BC trade. Note that over the full sample period the benchmark Sharpe ratio and skewness remain close to those reported in Panel A of Table 2 for the shorter period since 1994.

The G1-G5 trades exhibit invariably higher average returns than the SC trade. In addition, their average returns and return standard deviations tend to increase as the number of currencies in a trade decreases. The Sharpe ratios of the G1-G5 trades all exceed the benchmark Sharpe ratio (in two cases by a factor of about two), with the difference statistically significant at the 5% significance level in three cases out of five. Skewness increases in three cases for the G1-G5 trades, even though this increase is significant only for G1. Overall, these five trades reproduce, and possibly make more salient the features that characterize the enhanced trades in Table 2.

In contrast, the complementary trades B1-B5 fare much worse. The average returns are often two to three times lower than those of the SC trade, whereas the standard deviations are on average twice higher, leading to much lower annualized Sharpe ratios, which are between 0.04 and 0.18. In addition, the return skewness is markedly more negative for these complementary trades, averaging -0.77 (versus -0.11 for the G1-G5 trades). Furthermore, the p-values shown in square brackets, pertaining to tests of the differences in Sharpe ratios and skewness between the corresponding G1-G5 and B1-B5 trades are below 0.02 for four out of five Sharpe ratios, and show three (one) rejections at the 5% (10%) level for skewness.

The relatively high Sharpe ratios and slightly negative or positive skewness of the G1 to G5 trades earn them the label ”good” carry trades (”G” for good). Analogously, we refer to the B1 to B5 trades with low Sharpe ratios and strongly negative skewness as ”bad” carry trades (”B” for bad), from now on.

Turning to the larger sets of GC and BC trades, each constructed from five currencies, they broadly replicate the distinction between the good and bad carry trades above. The GC average returns (Sharpe ratios) are on average three (three and a half) times higher than those for the BC ones, and the GC skewness is on average twice lower (in absolute terms), whereby the differences are statistically significant in about half the cases. The Sharpe ratios for the GC trades are significantly higher than those for the SC trade in one third of the cases, in line with what was observed for the comparable G2 to G4 trades.

To further illustrate the properties of the GC and BC carry trades, Panel A of Figure 2 plots their Sharpe ratios versus skewness, similar to Figure 1, with unfilled circles and black dots, respectively. The distinction is sharp and clear in the Sharpe ratio dimension, where, with *no* exception, the GC trades dominate the BC trades, thus justifying their classification as good trades. On the other hand, a few GC (BC) trades display low (relatively high) skewness, hence the distinction is not as clear in this dimension, even though on average the skewness of the BC trades is still twice lower, consistent with the bad trades classification.

[Figure 2 about here.]

With some data mining even more striking distinctions can be obtained. For example, in Panel B of Figure 2, the unfilled circles refer to trades from five currencies which again involve the three least often excluded currencies (GBP, SEK, USD) in each case, but are now combined with any other pair that excludes the JPY. As before, black dots refer to the complementary trades. These two sets of 15 trades deliver striking separation in *both* the Sharpe ratio and skewness dimension. In what follows we retain the GC and BC trades as described above and in Panel A of Figure 2, which mimic closely the dynamic enhancement rule of Section 3.1.

Finally, we examine the return correlations between various carry trades, and present the results in the last two columns in Table 3. The first of these columns shows the correlations between the returns of the SC trade and the good and bad trades constructed from subsets. The second one shows similar correlations, but now calculated using only the *negative* SC returns and the corresponding returns of the good and bad trades. As do the Sharpe ratios and skewness coefficients, these correlations also exhibit a strong pattern. First, with the exception of the G1 and B1 trades, the correlations are always higher for the bad trades (on average 0.80, versus 0.67 for the good trades). The difference in correlations is large for the trades using the fewest currencies (0.72 for B1 versus only 0.40 for G5). The SC trade thus likely reflects more strongly the traits of the bad carry trades. Second, the distinction is even more pronounced with respect to the correlations with the down-moves in the SC trade. On average these correlations are about 0.70 for the bad trades and less than 0.50 for the good trades. Therefore, explanations of the carry trade that hinge on moves in the left tail of its return distribution are probably more relevant for the bad carry trades, and less valid for good carry trades, even though the latter may be of more interest to investors.

In sum, eliminating some typical carry trade currencies, such as the AUD, JPY and NOK, from the currency set leads to good carry trades, with Sharpe ratios and skewness often higher than those of the SC trade, and certainly higher than those of the complementary bad carry trades that involve mostly the typical carry trade currencies. Good trades reduce tail risk considerably.

3.5 Statistical significance of the distinction between good and bad carry trades

Table 3 shows that the distinction between good and bad carry trades is economically and statistically important. However, the statistical evidence must be interpreted with caution. In particular, the reported p-values rely on the block bootstrap procedure under the alternative, developed by Ledoit and Wolf (2008) (as in Section 3.2 and Appendix OA-II). While this procedure accounts for certain finite-sample properties of the distribution of currency returns, it does not reflect two aspects of our good carry trades. First, they are constructed using information from the enhancement procedure, as reported in Panel B of Table 2, and thus the procedure suffers from look-back bias. Second, the enhancement procedure applied to a finite sample is bound to lead to improved Sharpe ratios, even if in population all 10 currencies are necessary to attain optimal results. Therefore, a modified test is needed to assess the statistical contribution as fairly as possible. We emphasize, however, that the results in Table 3 need not be statistically significant to profoundly impact carry research: the finding that the prototypical carry trade currencies, if anything, worsen or certainly do not provide a positive contribution to carry returns, suffices.

With respect to the look-back bias, starting the sample in 1994, rather than 1984 weakens the statistical significance somewhat, but we still retain significance at the 10% level for the majority of the trades (not reported). A full correction for this bias would require a much longer sample where we actually let the procedure choose which currencies to exclude ex-ante, before we record trading results.

Incorporating the selection procedure into a test of statistical significance is harder, because it requires creating a benchmark world in which carry trades still have realistic attractive returns, but somehow the identity of the currencies contributing to these returns is randomized. Appendix OA-III describes in detail a procedure creating entirely randomized individual currency returns which nonetheless reproduces exactly the returns of standard carry (SC) in each randomized sample. We then apply our enhancement strategy to 1000 such randomized samples, finding that:

- the selection procedure biases the Sharpe ratios of the good trades upwards by about 0.15.
- while the observed Sharpe ratios are, on average, higher still, they are only in the 10% right tail of the simulated distribution for two good trades
- in a proper test using t-statistics, only the G5 (at the 5% level) and G3 (at the 10% level) deliver statistically significant improvements in Sharpe ratios.

When comparing the Sharpe ratios of good versus bad trades, the bias is worse and only the Sharpe ratios of G5 and B5 are significantly different under the test procedure controlling for selection bias.

In sum, we do find that statistical significance is retained for the G5 carry trade, and perhaps for G3 as well. Overall there is no strong statistical evidence that the enhancement procedure delivers significantly higher Sharpe ratios. However, it remains the case that the prototypical, "skewed" carry currencies can be removed from the trade without worsening performance.

4 Good carry trades as currency market risk factors

Lustig, Roussanov, and Verdelhan (2011) suggest as a key currency market risk factor the return of a trading strategy that each month goes long (short) a portfolio with the highest (lowest) forward differentials. This is obviously a symmetric carry trade strategy and is denoted here as "HML^{FX}" (to be distinguished from the Fama-French HML factor used in Section 5). Creating test portfolios by ranking currencies on forward differentials, they find that the covariation with HML^{FX} largely explains the difference in average returns between these portfolios. Furthermore, they propose HML^{FX} as a proxy for a global risk factor in a no-arbitrage model explaining the results, and show that it is also related to a measure of aggregate stock market volatility. Menkhoff, Sarno, Schmeling, and Schrimpf (2012) conduct a similar exercise using a global exchange rate volatility factor as a proxy for the global risk factor. We now revisit these findings by considering the good carry trades as risk factors and comparing their performance with that of the previously used currency market factors.

4.1 Test assets and risk factors

The test assets in our pricing tests are the currency portfolios studied in Lustig, Roussanov, and Verdelhan (2011), and kindly made available on Adrien Verdelhan's website for the period ending in 12/2013. They form five portfolios of currencies of developed countries (denoted "Developed"), and six portfolios which also include emerging market currencies (denoted "All"), by sorting the respective set of currencies on forward differentials. We consider these 11 portfolios *together* in our tests, and not the "All" and "Developed" separately, as they do. The larger cross section poses a higher hurdle to the various risk factors that are examined and compared.[2] We take the "net" versions of the 11 portfolios, which account for transaction costs.

[2] While using some currencies twice in each test, the average correlation between the six "All" portfolios and five "Developed" portfolios is only 0.74, which is just slightly higher than the average correlation among the "Developed" (0.72) or the "All" portfolios (0.68). Moreover, we have verified that the relative performance of the risk factors separately on the "All" and "Developed" portfolios remains largely the same.

Moving to the risk factors, first we use HML^{FX} ("All" version) as in Lustig, Roussanov, and Verdelhan (2011) (and available at Verdelhan's website). Next, we use a mimicking portfolio for the innovations in foreign exchange volatility (denoted "FXVol") as in Menkhoff, Sarno, Schmeling, and Schrimpf (2012).[3] Finally, we also consider as risk factors the good carry trades G1-G5 to contrast their performance, particularly with HML^{FX}. We note that the correlations between the G1-G5 trades and HML^{FX} (FXVol) are on average 0.39 (-0.29), and do not exceed 0.55 in magnitude, hence using them together does not raise multi-collinearity concerns. The respective correlations for the B1-B5 trades average 0.64 (-0.71), which hints at a much closer relation between the previously considered currency market factors and our bad carry trades.

4.2 Design of asset pricing tests

We adopt a standard asset pricing framework, following Cochrane (2005, Chapters 12 and 13), and consider linear factor models, both in their beta representation and stochastic discount factor (SDF) form, assuming SDF's specified as:

$$m_{t+1} = 1 \, - \, b' \left(f_{t+1} - E[f] \right). \tag{3}$$

In (3) f_{t+1} is a $K \times 1$ vector of risk factors and b is a conformable constant vector of SDF coefficients. Normalization is required when excess returns of the test assets are used (and hence the expectation of the SDF is not identified), and the adopted specification assumes $E\left(m_{t+1}\right) = 1$.

Given (3), the SDF form of a pricing model is $E[rx^i_{t+1} m_{t+1}] \, = \, 0$, where rx^i_{t+1} are the excess percentage returns of the test assets, indexed by i. In addition, we denote by Rx_{t+1} the $N \times 1$ vector of the rx^i_{t+1}'s.

[3] This factor is obtained as follows: First, we construct a proxy for exchange rate volatility, and for this purpose calculate the absolute daily log changes of the exchange rates of the G-10 currencies (except the USD) against the USD, using mid-quotes. We average these log changes over each month, and then across the nine currencies, to get a proxy of the monthly volatility of the G-10 currencies. Next, we take first differences of the volatility proxy. Finally, we regress these first differences on the percentage monthly returns (in USD) of long positions in the G-10 currencies (except the USD), calculated as in equation (1), but using mid-quotes. The factor we use is the fitted value from this regression (multiplied by 100).

The model is estimated with GMM, where the moment conditions are:

$$g = \begin{bmatrix} E[Rx_{t+1}m_{t+1}] \\ E[f_{t+1} - E[f_{t+1}]] \end{bmatrix} = \begin{bmatrix} E[Rx_{t+1} - Rx_{t+1}(f'_{t+1} - E[f'_{t+1}])b] \\ E[f_{t+1} - E[f_{t+1}]] \end{bmatrix}. \tag{4}$$

The weighting matrix defining which moments are set to zero is $a = \begin{bmatrix} d & 0 \\ 0 & I_K \end{bmatrix}$, where $d = E[Rx_{t+1}f'_{t+1} - Rx_{t+1}E[f'_{t+1}]]$.

If we denote $\mu = 1/T\sum_{t=1}^{T} f_t$ and $\overline{Rx} = 1/T\sum_{t=1}^{T} Rx_t$, where T is the length of the return time series, then the GMM estimates of b are $(d'd)^{-1}d'\overline{Rx}$, and that of $E[f_{t+1}]$ is μ. The standard errors of the b estimates are obtained from the covariance matrix $1/T(d'd)^{-1}d'Sd(d'd)^{-1}$, where S is an estimator of $\sum_{j=-\infty}^{\infty} E[u_{t+1}u'_{t+1-j}]$ and $u_{t+1} = \begin{bmatrix} Rx_{t+1}m_{t+1} \\ f_{t+1} - \mu \end{bmatrix}$. As in Lustig, Roussanov, and Verdelhan (2011), we use one Newey-West lag throughout to estimate S.

The beta representation of the pricing model is $E[rx^i_{t+1}] = \lambda' \beta^i$, with systematic risk exposures for asset i given by the vector β^i, and λ a vector of factor risk prices. The vectors β^i are estimated with GMM from time-series regressions of returns rx^i_{t+1} on the factors, and λ is estimated from a cross-sectional regression (without a constant) of average returns on the β's. We report the SDF coefficients b and factor risk prices λ with corresponding p-values, as well as p-values for the χ^2 statistic testing if the pricing errors are jointly equal to zero (see, e.g., Cochrane (2005, page 237)).

4.3 *Good carry trades in competition with other currency market risk factors*

Table 4 shows the results of tests which compare the performance of HML^{FX} and the good carry trades as risk factors. As in Lustig, Roussanov, and Verdelhan (2011), each test also includes the dollar factor, denoted RX, which is the average excess return of their basket of currencies held long against the USD. In each of the two panels of the table, the first line refers to a model with the RX and HML^{FX} factors alone, the next five lines to models with RX and each of the good carry trades G1-G5, and the remaining five lines to models combining RX, HML^{FX} and each of the G1-G5. The top panel summarizes results from time-series regressions of each of the test assets, and reports average coefficient estimates and (in parentheses) the number of respective estimates that are significant at the 5 or 10% confidence levels. (The regression results for each individual test asset are shown in the Online Appendix, Table OA-7.) The bottom panel reports both the prices of risk λ and the SDF coefficient estimates b. The latter are key in evaluating the relative importance of alternative factors for pricing a given cross section (see, e.g., Cochrane (2005, Chapter 13.4)). It is of obvious interest to determine whether good carry trades can replace HML^{FX} as the dominant currency risk factor, or whether they represent a distinct risk factor in the case of joint significance.

The top panel of the table does not reveal important differences between HML^{FX} and the good carry trades: the slope coefficients β in the time-series regressions are similarly significant; the R^2 related to HML^{FX} is slightly higher, but so are the respective intercepts α. When entering the regression jointly, the two factors also show similar significance, with the HML^{FX} coefficients remaining negative on average, but the coefficients on the good trades turning all positive on average. The RX factor always has a statistically significant slope coefficient of around 1.1.

In the bottom panel of Table 4, all two-factor models (RX with either HML^{FX} or a good carry trade) show significant prices of risk λ for HML^{FX} and the good trades (at the 5% level), but not for the RX factor. However, in the three-factor models the p-values increase somewhat for HML^{FX}, and in three cases become significant only at the 10% level, while the significance remains unaffected for the λ's of the good trades.

An essential difference, however, is observed with respect to the SDF coefficients b. In the two-factor models, the b-coefficient for HML^{FX} is significant at the 10% level only, but at the 5% level for all good trades. However, in the three-factor models the b coefficients turn highly insignificant for HML^{FX}, whereas for the good carry trades they remain significant at the 5% level in three of the five cases, and at the 10% level in one case. Moreover, the test for the pricing errors being jointly equal to zero rejects in this sample for the two-factor model with RX and HML^{FX} with a p-value of zero, while the corresponding p-values for the models with RX and a good trade are all above 0.10, except for G3 where the p-value is 0.09. The test fails to reject the three-factor models at the 5% level for all specifications. In addition, the b coefficients appear similar across different specifications for the good trades, but not for the HML^{FX} factor, where the sign switches across specifications. The results in the bottom panel of Table 4 clearly favor the good carry trades as risk factors explaining the returns of the interest rate-sorted currency portfolios.

Table OA-1 in the Online Appendix shows results from analogous tests, but with the currency volatility factor FXVol replacing HML^{FX}. The conclusions remain robust: the good carry trades again win the horse race, with p-values for all their SDF coefficients equal to 0.01 or lower, while these p-values are never below 15% for FXVol.

4.4 Return predictability of good and bad carry trades

The cross-sectional tests we have conducted follow the extant literature and assume constant prices of risk and betas. It is surely conceivable that these assumptions are violated and thus that additional factors may affect the unconditional cross section of currency returns (see e.g. Jagannathan and Wang (1996)). There is, in fact, evidence of carry return predictability.

Bakshi and Panayotov (2013) document that commodity index returns and exchange rate volatility strongly predict carry trade returns. Further, relying on a model which postulates a link between commodity trading and currency dynamics, Ready, Roussanov, and Ward (2017) find time-series predictive ability of an index of shipping costs, the Baltic Dry Index (BDI), for carry trade returns. They primarily investigate an unconditional carry strategy that is always long the currencies of commodity exporters (commodity-producing countries) and short those of commodity importers (countries producing final goods), which is a key component of their theoretical model.

In Table 5 we reconsider the evidence for time-series predictability from the perspective of good and bad carry trades. To follow closely the empirical design in the two studies cited above, we use *log* returns of equally-weighted good and bad carry trades. The commodity index predictor is defined as the three-month log change in the CRB index. To construct the exchange rate volatility predictor, we first calculate the annualized standard deviation of the daily log changes over each month t for each of the G-10 spot exchange rates against the USD. The cross-sectional average of these for month t is denoted σ_t^{avg}. The value of the volatility predictor available at the end of month t (and used to predict the return for month $t+1$) is the three-month log change $\ln\left(\sigma_t^{avg}/\sigma_{t-3}^{avg}\right)$. The shipping cost predictor is the three-month log change in the BDI. As in Bakshi and Panayotov (2013), we show in-sample predictive slope coefficients β and their p-values, using Hodrick (1992) standard errors, adjusted R^2's, and p-values for the MSPE-adjusted statistic of Clark and West (2007).[4] In addition, we show an (out-of-sample) measure of the economic significance of predictability, based on a prediction-based trading strategy. This strategy enters a carry trade at the end of month t only if its predicted return for month $t+1$ is positive. If a negative return is predicted, then no position is taken and the strategy's return for month $t+1$ is zero. The reported measure "Δ SR" of economic significance equals the difference between the Sharpe ratio of the prediction-based strategy, implemented with the respective subset of G-10 currencies, and the corresponding carry trade as shown in the first column. The predictive regressions use an expanding window with initial length of 120 months.

The main feature that stands out from Table 5 is the sharp difference between the return predictability

[4]MSPE stands for "mean squared prediction error". The statistic is obtained using $f_{t+1} = (y_{t+1} - \mu_{t+1})^2 - [(y_{t+1} - \widehat{\mu}_{t+1})^2 - (\mu_{t+1} - \widehat{\mu}_{t+1})^2]$, where $\widehat{\mu}_{t+1}$ is the prediction for month $t+1$ from a predictive regression $y_{t+1} = a + bx_t + \varepsilon_{t+1}$, and μ_{t+1} is the historical average of y. Both $\widehat{\mu}_{t+1}$ and μ_{t+1} are estimated using data up to month t. The null hypothesis is that $\widehat{\mu}_{t+1}$ does not improve on the forecast which uses μ_{t+1} as the predictor. The test statistic is the t-statistic from the regression of f_{t+1} on a constant, for which we report one-sided p-values.

results for good and bad carry trades. Out of 15 possible combinations with the three predictors, the G1-G5 trades show a significant predictive slope on three occasions, whereas the B1-B5 trades record 13 occasions with p-values not higher than 0.05 and another one with a p-value below 0.10. The average predictive R^2 is 0.7% for the G1-G5 trades and 2.2% for the B1-B5 trades. The MSPE statistics show significant out-of-sample predictability in one case (out of 15) for the G1-G5 trades and in 13 cases for the B1-B5 trades. Finally, exploiting the predictability in dynamic trading does not materially impact the Sharpe ratio for the G1-G5 trades (the average change is -0.005), while it mostly improves the Sharpe ratio for the B1-B5 trades, on average by 0.10. This is economically quite a large improvement, because the Sharpe ratios for bad carry trades are often below 0.10 (see Table 3).

The above patterns are confirmed by the results from the GC and BC trades, where again the GC trades show insignificant predictive slopes for two of the predictors, twice smaller predictive R^2's, rarely significant out-of-sample predictability, and on average a reduction in Sharpe ratios by 0.01 from exploiting predictability, in contrast to the BC trades which exhibit an increase by 0.09 in Sharpe ratios on average.

We recognize that our predictability results echo some findings in Ready, Roussanov, and Ward (2017). The trade (denoted IMX) for which they find the strongest predictability with the CRB commodity index and the BDI overlaps to a large extent with some of our bad trades. In particular, their Figure 5 shows that this trade would be long AUD, NZD and NOK, and short JPY and CHF, which is often true for our bad trades.[5] Therefore, our Table 5 confirms the predictive ability of the CRB and BDI for a "commodity focused" carry trade as implied by their commodity trade model. Our contribution here, however, is to highlight the similarity between the commodity-based trade and our bad carry trades, and the fact that a commodity-based interpretation of carry trade returns reflects mostly features of bad carry trades.

[5] They also show that a complement to the IMX trade (denoted CHML) is not predictable at all by the CRB or BDI, and also has practically zero skewness, similar to some of our good trades. However, the Sharpe ratio of CHML is still below that of their version of the standard carry trade (0.85 vs. 0.95 in their sample and without transaction costs), hence their orthogonalization procedure fails to identify a good carry trade.

Essentially, our article provides an interesting qualification to the prevailing carry return predictability story. A carry trade that focuses on the prototypical high (low) interest currencies such as the AUD (JPY) is a rather unattractive strategy, but its return properties can be substantially enhanced by exploiting return predictability. This predictability may also resurrect the performance of bad trades in the cross-section, either through hedging terms or through correlation between conditional betas and the conditional expected carry return. In contrast, our good carry trades have attractive properties which, however, cannot be enhanced by the predictors previously identified in Bakshi and Panayotov (2013) and Ready, Roussanov, and Ward (2017). It remains, of course, conceivable they are predictable by other variables.

5 Good and bad carry trades and previous carry interpretations

This section investigates whether the conclusions of previous studies of carry trades are valid for good and/or bad trades, and whether the good-bad carry trade perspective may affect such conclusions.

5.1 Explaining carry trade returns with equity market risk factors

We start by re-examining a key result in the literature stating that standard (linear) equity market factor models cannot explain the time variation in carry returns, which appear uncorrelated with these risk factors in normal times, but correlate highly with them in crisis times (e.g., Melvin and Taylor (2009), Christiansen, Ranaldo, and Söderlind (2011)).

We examine three models: (i) the Fama-French three-factor model, following Burnside, Eichenbaum, Kleshchelski, and Rebelo (2011) and Daniel, Hodrick, and Lu (2017), (ii) a three-factor model with the market factor, the global equity volatility factor used in Lustig, Roussanov, and Verdelhan (2011), and their *product*, and (iii) a model with two factors which explicitly distinguish the down- and up-moves of the equity market, in the spirit of Lettau, Maggiori, and Weber (2015). The latter two models effectively exhibit a non-linearity that may capture the time-variation in the correlation mentioned above. To conserve space, we relegate detailed results to the Online Appendix, summarizing the key results here.

Let's start with the model featuring a market factor (denoted MKT), proxied by the total return of the MSCI-World equity index, in excess of the risk-free rate and expressed in USD, an equity volatility factor (EqVol) constructed as in Lustig, Roussanov, and Verdelhan (2011), and the interaction term (the product of MKT and EqVol). Table OA-2 shows that in time-series regressions of carry trade returns on the three risk factors the main difference between good and bad trades is in their loadings on the product factor. These are typically negative, albeit rarely significant, for the good trades, while they are positive, much larger in magnitude, and almost always significant at the 5% significance level for the bad trades.

Given that increases in volatility tend to characterize periods of market downturns (the correlation between MKT and EqVol is -0.24 in our sample), these findings imply that the market risk exposure of the bad trades increases substantially in bad times. Thus, bad carry trades under-perform in times of crisis, while good trades are less affected.

We also perform GMM-based cross-sectional tests on the GC and BC return cross sections. For the GC trades, the risk price for the MKT factor is significant at the 5% level, while for the BC trades no risk price is significant, although the model is not rejected for either of the two cross sections. When we run a simple OLS regression of actual average returns on a constant and the model-based expected returns, we obtain an R^2 of 0.67 for the GC trades, and 0.29 for the BC trades. The combined evidence suggests that this three-factor model does not adequately describe the returns of the bad carry trades, but still saliently reveals the high exposure of these trades to the equity market during high-volatility periods. In contrast, a significant price of risk for the market factor and tighter link between model expected returns and average returns show the promise of the model to provide a risk-based interpretation of good carry trades.

The Online Appendix further shows quite similar results for the model with an Up- and Down-market factors. Table OA-3 shows that good (bad) carry trades load primarily on the Up (Down)-market factor, with beta exposures being economically and statistically very different across the two types of trades. In the cross-sectional tests, the prices of risk for both factors are significant; the pricing errors are not statistically different from zero and the model generates expected returns highly (weakly) correlated with good (bad) carry trades.

The Online Appendix and Table OA-4 also report analogous results for the Fama-French three-factor model. Here the time-series regressions reveal that good carry trades do not load much on any of the three factors, and retain significant alphas relative to the model. In contrast, the bad carry trades feature significantly higher regression slope coefficients on all three factors and it is striking that their SMB and HML exposures are positive and economically meaningful (often even above 0.10). However, the Fama-French model fails to fit expected returns cross-sectionally, with all prices of risk being insignificantly different from zero for both good and bad carry trades.

In sum, the evidence from Tables OA-2, OA-4 and OA-3 provides some (weak) support for the ability of risk factors from the equity market to explain the returns of the good carry trades. Our results are not directly comparable to extant studies which analyse numeraire-dependent carry trades. While Burnside, Eichenbaum, Kleshchelski, and Rebelo (2011) and Daniel, Hodrick, and Lu (2017) find that an asymmetric carry trade cannot be explained by the Fama-French model, the latter study also finds this model to render insignificant the alphas for the dollar-neutral symmetric carry trade.

5.2 Currency crashes as an explanation for the carry return puzzle

One established explanation for the carry trade's profitability is that it reflects compensation for the negative return skewness or crash risk, inherent to these trades. For example, Brunnermeier, Nagel, and Pedersen (2009) argue that ”investment currencies are subject to crash risk, that is, positive interest rate differentials are associated with negative conditional skewness of exchange rate movements.... The skewness cannot easily be diversified away, suggesting that currency crashes are correlated across different countries .... This correlation could be driven by exposure to common, crash-risk factors". If agents exhibit a preference for positive skewness, an equilibrium model may generate negatively skewed returns and high Sharpe ratios for the carry trade.

However, the crash risk hypothesis is not consistent with our findings from good and bad carry trades (see Figure 1 and Table 3): good carry trades have relatively high Sharpe ratios and slightly negative (or even positive) skewness. The assertion in Brunnermeier, Nagel, and Pedersen (2009) that the negative skewness in carry trade returns cannot be diversified away must also be qualified. We have demonstrated that, in fact, skewness can be dramatically improved by judiciously removing currencies from the carry trade, without impairing profitability.

Studies relying on option market data (e.g., Burnside, Eichenbaum, Kleshchelski, and Rebelo (2011), Jurek (2014)) have criticized the crash-risk hypothesis before, because options can essentially hedge away the crash risk without affecting the carry trade's profitability very much. In the Online Appendix we directly examine option-hedged good and bad carry trades. We find that while skewness also improves across the board, the Sharpe ratio differences between hedged and unhedged trades deliver mixed results. These differences are relatively small for bad carry trades, but we observe a substantial decrease (increase) in Sharpe ratios for the G1-G3 (G4-G5) good trades. Our contribution here is to demonstrate that return skewness can be improved without recourse to currency options, undermining the crash risk hypothesis.

6 Further exploration of good and bad carry trades

In this section we embark on a more detailed examination of the good and bad carry trades, trying to set the stage for future work that will hopefully clarify fully the economic interpretation of our findings. First, we reflect on the return components of various carry trades and how they contribute to the differential performance of the good and bad trades.

Second, the good carry trades always include the USD (it is never excluded in our selection procedure). There is a burgeoning literature stressing the special nature of the USD in international financial markets, so it is perhaps not surprising that relatively sizable currency premiums are attached to dollar exposure. For example, Adrian, Etula, and Shin (2015) associate increased global dollar funding with expected currency depreciations; Hassan (2013) argues that economies representing a larger share of world wealth have low interest rates and low risk premiums, whereas Maggiori (2013) ascribes a low premium to holding the USD to its role as a reserve currency. The importance of the USD in currency trading was also emphasized by Lustig, Roussanov, and Verdelhan (2014), who explore a new trade, denoted ”dollar carry”, which goes long (short) in all foreign currencies against the USD with equal weights when their average interest rate differential relative to the USD is positive (negative). This trade also featured prominently in Hassan and Mano (2015). The dollar carry trade has a very attractive Sharpe ratio, substantially higher than that of the standard carry trade (SC). These traits raise the issue that we may have simply repackaged dollar carry into our good carry trades. We show in this section that this is not the case, and these two types of trades, while correlated, are economically distinct.

Finally, Lustig, Roussanov, and Verdelhan (2011, 2014) have proposed several economic factors that are correlated with HML (standard carry) and dollar carry. We revisit this analysis from the perspective of our good and bad carry trades.

6.1 The sources of good and bad carry returns

In Table 6, we decompose carry trade returns into the carry component (i.e., the forward differential) and exchange rate change component. The SC trade derives more than 100% of its returns from carry. That is, the investment currencies do depreciate and/or the funding currencies do appreciate, but the exchange rate component is sufficiently small relative to carry to leave an attractive return on the table. Bad carry trades have higher carry return components, both in absolute terms (and the difference is statistically significant) and in relative terms, but even more negative exchange rate components, so that lower returns than those for standard carry are obtained. In contrast, good carry trades derive their returns *both* from the carry and exchange rate components. Their carry component is on average about 20 basis points lower than that of SC in three cases (and statistically significant at the 5% level for the G4 trade), while it is significantly higher for the G3 trade (even if still lower than the carry of any bad trade).

The contrast between the carry contributions to the returns of bad and good trades is illustrated in Panel A of Figure 3. The graph plots total average return on the horizontal axis, and the ratio of carry to total return on the vertical axis for all trades considered in this paper (18 GC trades, 18 BC trades, as well as the G1 and G5, and B1 and B5 trades). The graph also includes the standard and dollar carry trades. Bad carry trades have lower returns and much higher carry-to-return ratios; good carry trades have higher returns, and derive between 50 and 100% of their returns from carry (the G5 trade being the only exception).

[Figure 3 about here.]

Panel B of Figure 3 shows the distinction in an alternative way. We plot, again for all good and bad carry trades, the carry component on the horizontal axis and the total return on the vertical axis. In this plot, points at or near the 45-degree line correspond to a random walk model where the average carry return equals the carry itself, while points at or near the horizontal axis would imply that the unbiasedness hypothesis holds. Note that while the good carry trades plot barely to the left of the bad trades (that is, there is still a significant carry component embedded in their returns), their total returns are much higher because the carry component is not at all eroded by exchange rate depreciation/appreciation. Whereas all but one of the good trades plot above the 45-degree line, the bad trades all plot below the 45-degree line, and are often not far away from the horizontal axis.

In addition, we run a cross-sectional regression of our various carry trade returns (G1, G5, B1, B5, 18

GC and 18 BC, a total of 40 trades) onto a constant and the carry, accommodating an interaction effect for the good trades.

$$\overline{ret}_i = \alpha_0 + \alpha_1 DUMMY_i + (\beta_0 + \beta_1 DUMMY_i)\overline{carry}_i + \varepsilon_i, \tag{5}$$

where $\overline{ret}_i$ is average return of a carry trade, $DUMMY_i$ is one for a good trade and zero otherwise, and $\overline{carry}_i$ is average carry. We find the dummy to be insignificant for the slope, but highly significant for the constant. Leaving out the interaction for the slope coefficient, we find $\alpha_0 = -0.16$ (0.47), $\alpha_1 = 1.63$ (0.20) and $\beta_0 = 0.42$ (0.23), with standard errors in parentheses. Thus, the returns of both good and bad carry trades are significantly related to carry (but not one for one), whereas the significant constant for good trades reflects the positive exchange rate component of their returns. We add the corresponding parallel regression lines (top for good trades and bottom for bad trades) to Panel B of Figure 3.

The above results suggest that the unbiasedness hypothesis may not be strongly rejected for bad carry currencies, which include the prototypical carry currencies. Recall that a necessary condition for a carry trade to deliver excess returns is that the unbiasedness hypothesis does not hold, at least over some time periods (see Bekaert, Wei, and Xing (2007) for recent tests of the hypothesis). However, when examining standard regressions testing the unbiasedness hypothesis for the four pairs containing prototypical carry currencies, AUD/JPY, NZD/JPY, AUD/CHF and NZD/CHF (see Table 7), we find no strong rejections of the hypothesis. In particular, we regress future exchange rate changes onto a constant and the current forward differential, and the null hypothesis is that the constant is zero and the slope coefficient is one. The constants in all four regressions are insignificantly different from zero, and the slope coefficients are insignificantly different from one. Most saliently, the slope coefficient for the AUD/JPY regression is 0.92, and thus remarkably close to one. However, our analysis reveals the NOK also to be a ”bad” currency, more so than the CHF and the NZD. Interestingly, Table 7 shows that the slope coefficient in unbiasedness regressions of the NOK relative to the CHF and JPY is either not significantly different from one, or exceeds one by a large amount, indicating an expected depreciation of the NOK relative to the JPY when the NOK interest rate exceeds the JPY one. This is partially counteracted by a positive and significant constant.

The decomposition and the regression results above also suggest that the good carry trades are likely to be more ”active” than the bad or standard carry trades, i.e., they likely involve more frequent re-balancing. The insightful paper by Hassan and Mano (2015) decomposes the carry trade into a ”static” trade (which goes long (short) currencies with unconditionally low (high) forward differentials) and a dynamic trade, which also helps explain deviations from unbiasedness. Such deviations, driven by the slope coefficient in the unbiasedness regressions being different from one, lead to dynamic trades when forward premiums are high or low relative to their unconditional means. However, carry trades can also be profitable simply through non-zero constants in the unbiasedness regressions (see also Bekaert and Hodrick (2012, Chapter 7)).

Table 8 provides some evidence regarding the dynamic nature of the various carry trades. We create a dummy variable that records the proportion of currencies that change position (from long to short or vice versa) at each point of time. For example, for a completely static trade this proportion would be zero, whereas a trade where half the currencies switch positions at each point of time would record 50% on this measure. Furthermore, since the dynamic nature may be related to the number of currencies in the trade, it is important to take sampling error into account. The table shows the sample averages of these proportions, together with 95% confidence intervals, computed using the bootstrapped carry trade returns that we have employed in the exercise reported in Section 3.5.

Clearly, the good carry trades are more ”dynamic” than the bad trades and the SC trade. Note that the average proportions for good trades are invariably above the confidence interval for SC, and in turn, that for SC is always below the good carry confidence intervals. Yet, the proportions for the good carries and SC are relatively highly correlated in the time series, suggesting that these trades switch currency positions at roughly similar times (the correlations in the fourth column of Table 8 range between 49% and 82%). For bad carry trades the proportion of switches is typically within the confidence interval of SC, except for the B3 trade.

The last three columns in Table 8 reports results related to the decomposition in Hassan and Mano (2015). In particular, we show the ratios between the average returns of various static carry trades (which are never re-balanced), and the returns of the corresponding good or bad carry trade. Hassan and Mano (2015) find that static trade returns account for about 70% of carry trade returns (but the standard error on that estimate is substantial), whereas according to Lustig, Roussanov and Verdelhan (2011) this proportion is between one third and one half.

To obtain the results in the last three columns of Table 8 we mimic the Hassan and Mano (2015) methodology, and now use log returns and carry weights equal to the demeaned and normalized forward differentials at each trading date over 1/1995-6/2014. The trades are constructed with all G-10 currencies, as well as with the currencies entering the G1-G5 and B1-B5 trades. The various ”static trades” thus obtained use as weights the average forward differentials over the 12/1984-12/1994 period (the first 120 months of our sample), demeaned and normalized to have absolute values that sum to one. The weights are kept fixed for the entire sample period 1/1995-6/2014, without ever re-balancing.

The average return of the static SC trade in Table 8 is about half of that of the original SC trade, confirming results in the literature. Importantly, there is a clear distinction between the relative performance of the static versions of the good and bad carry trades, with the ratios between the corresponding average returns never exceeding 0.30 (and sometimes going negative) for the good trades, but ranging between 0.60 and 1.2 for the bad trades. The distinction is even clearer in terms of Sharpe ratios (see the last two columns), which for the good static trades rarely exceed 0.15, much worse than their re-balanced counterparts. In contrast, the Sharpe ratios of the re-balanced and static ”bad” trades are quite close to one another.

Hence, good carry represents a dimension of standard carry that is not well explained by its static component. This is intuitive, because good carry trades tend to exclude currencies with either the highest or lowest forward differentials, and thus do not have stable short and long positions. In contrast, the currencies involved in bad carry trades typically switch less often from long to short positions and vice versa, as also indicated more directly above. Our results thus completely confirm the Hassan and Mano (2015) decomposition for ”bad carry trades”, but not for ”good carry trades”. Hassan and Mano (2015) split up carry trades in the static carry trade we studied above and a ”dynamic” trade, which essentially exploits time-variation in the relative ranking of currencies in terms of their forward differentials, relative to their unconditional counterparts. This dynamic trade must necessarily be relatively more important for good trades, which feature currencies with less extreme interest rate differentials relative to the dollar, and for which the unbiasedness hypothesis does not hold. The dynamic trade therefore also contributes positively to the trade exploiting deviations from unbiasedness (what they called the ”forward premium trade”). Do note that our results are not entirely comparable to Hassan and Mano (2015) because they do not impose symmetry on their carry trade, while we do.

6.2 Good carry versus dollar carry

As indicated before, the dollar carry trade and our good carry trades should show some relation, since they share a large USD exposure. In this section, we fully characterize the differences and similarities.

First, note that dollar carry (hereafter DC for short) does not satisfy the standard conditions for a carry trade as discussed in Section 2. Carry trades go long (short) high (low) yield currencies, whereas DC combines high and low yield currencies on one side of the trade. Going back to Table 6, the last column reports the carry and exchange rate change components for DC, and the last row of the table reports p-values for a test of equality between the carry components of the trade in the respective column and DC. The DC trade derives most of its substantial returns from currency appreciation, and only 22% from interest rate differentials. This proportion is significantly lower than that of any other carry trade. Perhaps not surprisingly, the G5 trade, only featuring three currencies and including the USD comes closest to DC. In Figure 3, the DC trade also represents somewhat of an outlier. Moreover, given its small carry component, the DC trade has a return significantly above (at the 1% level) the return predicted by the regression in Figure 3 (panel B). To drive home the different nature of DC relative to our good and bad carries, we also constructed versions of DC from ”good” and ”bad” currencies separately, and find that their Sharpe ratios are very similar.

Second, DC is much less dynamic than the good trades: the last row of Table 8 shows that it switches positions more rarely than any other trade (in spite of the fact that any switch involves all currencies). Besides, the switching proportions are in fact not too correlated with those for the good trades (at most 39%) or the SC trade (47%). Furthermore, Table 8 (column ”days w/o switch”) shows that while the typical proportion of days when not a single currency switches position from long to short or vice versa in carry trades ranges between 0.60 and 0.80, it is 0.93 for DC.

Third, because good carry trades eliminate some non-dollar exposure, they should be more correlated with DC than SC or the bad carries are. Table 9 confirms this intuition, showing in Panel A that DC has the highest correlation with good trades. However, for the G1, G2 and G3 trades the correlation is less than 50%, and it does not exceed 70% for the G5 trade. Besides, the correlations between the good trades (except G5) and SC are higher than those for DC. Good carries thus preserve their close link with the SC trade, and remain distinct from DC.

In Panel B of Table 9 we report on regressions which have the returns of SC, DC or good carries either as dependent or independent variables. Both SC and good carries have explanatory power for the DC trade, but SC is insignificiant in two specifications. The R^2's are relatively low, however, less than 40% in all but one case. The DC trade mostly delivers significant alphas relative to these two factors, which is not surprising, given its very attractive return profile.

Next, both SC and DC have explanatory power for good carry trades, with almost all slope coefficients featuring p-values below 1% and R^2's ranging from 43 to 75%. The coefficients on SC, however, are typically much larger than the coefficients on DC. The G2, G3 and G5 trades still show significant alphas with respect to these two factors. Finally, the return of SC is explained with R^2s between 18 and 75% and all intercepts are insignificantly different from zero. The explanatory power in this case comes predominantly from the good carries, as evidenced by the small coefficients and some high p-values on DC.

The main message of Table 9 is that neither good carries, nor DC can be perfectly spanned by other trades, despite being correlated with them. In contrast, SC *is* spanned, and this is mostly due to the good carry trades. Being proper carry trades, the good carries should therefore be viewed, unlike DC, as better versions of SC.[6]

6.3 Revisiting the factor pricing of currency returns

We now compare the performance of DC and good carries as pricing factors for the interest rate-sorted portfolios, discussed in Section 4.3 above. Table 10 reports results from pricing tests, which juxtapose DC with the G1-G5 trades, but do not include the RX factor (which only shorts the USD and has correlation of 0.53 with DC).[7] The top panel summarizes, as in previous tables, results from time-series regressions on individual portfolios, while the bottom panel presents results from cross-sectional tests. (The first-pass

[6]For further validation of this claim, in unreported results we consider creating a mean-variance efficient portfolio from DC, SC and one of the G1-G5 portfolios. When we do so, the good carries invariably get a large positive weight, always exceeding that of DC, sometimes by factor of two or three, and SC is always shorted (except if paired with the G5 trade, when its weight is close to zero). In other words, good carries dominate DC and SC is pushed out.

[7]DC is short the dollar about 70% of the time in our sample, and its profitability is entirely driven by these short dollar positions. Note that this is not true for the good carry trades, which gain both when the dollar is short and long.

regression results for each individual test asset are shown in the Online Appendix, Table OA-8.)

DC alone explains reasonably well the time-series behavior of the test assets - none of the intercepts and all slope coefficients are significant, even at the 5% confidence level, with a relatively high R^2 of 21%. The performance of the good trades alone is similar with respect to the intercepts, while the slopes are not always significant and the R^2's are much lower in three out of five cases. Moving to the cross-section, however, we observe that the price of risk for DC is significant only with a p-value of 0.07, while three of the good trades show significance at the 5% level. In addition, the test for the pricing errors being jointly equal to zero rejects with a p-value of 0.03 for DC, but never rejects for individual good carries, even at the 10% confidence level.

We also perform cross-sectional tests which include both DC and a good trade, as reported at the bottom of Table 10. The price of risk λ is now statistically significant for DC at the 5% level in only two out of five cases, whereas four out of the five p-values for the good carries are at 2% or below. The p-values for the SDF coefficients b, which provide the proper horse race test, are all above 0.20 for DC. In contrast, four of these p-values for the good trades are below or equal to 0.10. While the statistical significance in favor of good trades is borderline, this evidence points to an important distinction between DC and the good carry trades. Also note that the b-coefficients for the good trades are quite stable (see also Table 4), whereas the b-coefficients for DC even switch sign across specifications. The conjecture that good trades may be simply reflecting features of DC is thus not supported here. We also note that the relatively high correlation between DC and the G5 trade likely leads to insignificant coefficients.

6.4 Economic correlations

It appears that the extant academic literature does not have a complete economic explanation for the risk/return profile exhibited by the standard and dollar carry trades. Adding the distinction between good and bad carry trades makes the task even more challenging. We now provide some summary analysis linking currency returns to economic factors, largely building on the work of Lustig, Roussanov, and Verdelhan (2011, 2014), who estimate reduced-form pricing kernel models and perform a host of interesting regressions.

In particular, we run multivariate regressions to investigate the explanatory power of three economic factors for various carry trades, including DC. The first factor is the global equity market volatility, proposed by Lustig, Roussanov, and Verdelhan (2011) as a global risk factor to help explain carry returns. Next, we consider two real variables: global industrial production growth, proxied by the OECD total growth[8], and the residual from regressing the US industrial production growth onto the global growth variable; Lustig, Roussanov, and Verdelhan (2014, Section 6.2) suggest that the expected return of the DC trade ought to be correlated with such a US-specific growth variable.

To conserve space the results are reported in Online Appendix Table OA-6, but they are summarized here. First, none of the three variables has explanatory power for DC in our sample period. In contrast, there are non-zero adjusted R^2's for all carry returns, standard, good and bad. Second, for SC and most of the bad carry trades, the equity volatility factor is significant at the 5% level, while among the good trades this factor is significant at the 5% level only for G3. OECD industrial production growth is statistically relatively more important for SC and good carries than for bad carries. Finally, the US-specific component of industrial production growth is not significantly related to any of the trades in our sample period. In summary, the various carry trades are similarly related to the macro factors considered, except for DC, which appears to have no significant link to these factors.

To gain further intuition, we also simulated the Lustig, Roussanov, and Verdelhan (2014) model, using the calibrated parameters provided in their paper to investigate whether it could generate the distinction between good and bad carry returns that we find in the data. While the model does generate realistic HML

[8] from stats.oecd.org, Monthly economic indicators, Production of total industry excluding construction, growth rate over the previous month, seasonally adjusted

returns, our simulation results reveal that it does not reproduce the good-bad carry phenomenon. In fact, good carry trades perform worse than standard carry in the simulated data.

Linking this finding to a possible economic interpretation, we note that the model is essentially Gaussian and does not incorporate skewed returns. At the same time, good carry trades invariably exhibit less negative return skewness than standard carry or bad carries. This seems essential to understand the good /bad carry return distinction. Consider the small table underneath, which shows average forward differentials against the USD for the remaining G-10 currencies over our full sample period 12/1984 to 6/2014 (together with the first three moments of the percentage returns of long positions in each currency against the USD, not adjusted for transaction costs). All numbers are annualized and in percent, except for skew-

	NZD	AUD	NOK	GBP	SEK	CAD	EUR	CHF	JPY
avg. forw. differ.	4.40	3.26	1.98	2.19	1.63	0.83	-0.41	-1.58	-2.51
avg. return	7.23	4.42	3.82	4.11	3.27	1.81	2.85	2.76	1.21
stand. dev.	12.40	11.88	10.24	10.83	11.31	7.10	10.92	11.82	11.46
skewness	-0.137	-0.585	-0.374	-0.048	-0.321	-0.331	-0.139	0.109	0.497

ness. The bad carry currencies (JPY, AUD and NOK) do not only have the highest forward differentials, they also are the most skewed. The good carry trades essentially remove these currencies and thereby do not worsen and mostly improve the return-risk properties of the trade. Therefore, this skewness must be idiosyncratic and not priced, or it must be endogenously generated by carry traders. Why this is the case remains an important open question for further research, but it surely undermines any explanation of attractive carry returns based on priced ”crash” risk. The DC and good carry trades use very different mechanisms to eliminate the impact of bad carry currencies. While the good trades simply remove the currencies, DC puts such naturally ”long” and ”short” currencies on the same side of the trade.

7 Conclusion

This paper introduces ”good” and ”bad” carry trades, which are all constructed from subsets of the G-10 currencies, but exhibit markedly different properties. The differences are evident in return features like Sharpe ratios and skewness, and also change previous interpretations of carry trades. Surprisingly, trades that just exclude some of the typical carry trade currencies *do* perform significantly better than the benchmark SC trade, while trades that only include the typical carry currencies have inferior return profiles. These findings challenge the conventional wisdom on the construction of carry trades from an investor’s view point. Furthermore, the trades from subsets also challenge some of the available conceptual interpretations of the carry trade. We document that several of these interpretations appear to be mostly consistent with the bad carry trades, but are less applicable to good trades.

We find that good carry trades can serve as risk factors, able to explain a cross section of currency portfolio returns, and in this role can drive out previously suggested risk factors, such as the HML^{FX} factor of Lustig, Roussanov, and Verdelhan (2011). Further, the returns of good carry trades can be explained to a certain extent with risk factors from the global equity market. While good carry trades are more strongly correlated with the ”dollar carry” trade of Lustig, Roussanov, and Verdelhan (2014) and Hassan and Mano (2015) than is the standard carry trade, good trades remain symmetric carry trades, deriving the bulk of their returns from carry, and offer a distinct return profile.

The results in this paper, even though largely focused on the statistical properties of carry trade returns, should impact the study of carry trades in various directions. First, exploring crash risk or differentiating fundamental risks of commodity producers versus exporters are unlikely fruitful avenues of research. Second, our reported asset pricing tests can inform further risk-based interpretations of carry trade returns. Finally, it can be promising to explore, in the spirit of Koijen, Moskowitz, Pedersen, and Vrugt (2015), the notion of good and bad carry trades from financial assets other than currencies.

References

Adrian, T., Etula, E., Shin, H. S., 2015. Risk appetite and exchange rates. Staff reports. Federal Reserve Bank of New York.

Ang, A., Chen, J., 2010. Yield curve predictors of foreign exchange returns. Working paper. Columbia University.

Ang, A., Chen, J., Xing, Y., 2006. Downside risk. Review of Financial Studies 19, 1191–1239.

Annaert, J., Van Osselaer, S., Verstraete, B., 2009. Performance evaluation of portfolio insurance strategies using stochastic dominance criteria. Journal of Banking and Finance 33, 272–280.

Bakshi, G., Panayotov, G., 2013. Predictability of currency carry trades and asset pricing implications. Journal of Financial Economics 110, 139–163.

Barroso, P., Santa-Clara, P., 2014. Beyond the carry trade: Optimal currency portfolios. Journal of Financial and Quantitative Analysis 50, 1037–1056.

Bekaert, G., Hodrick, R., 2012. International Financial Management. Prentice Hall.

Bekaert, G., Wei, M., Xing, Y., 2007. Uncovered interest rate parity and the term structure. Journal of International Money and Finance 26, 1038–1069.

Bilson, J. F. O., 1984. Purchasing power parity as a trading strategy. Journal of Finance 39, 715–723.

Brunnermeier, M., Nagel, S., Pedersen, L., 2009. Carry trades and currency crashes. NBER Macroeconomics Annual 23, 313–347.

Bruno, V., Shin, H. S., 2015b. Capital flows and the risk-taking channel of monetary policy. Journal of Monetary Economics 71, 119–132.

Burnside, C., Eichenbaum, M., Kleshchelski, I., Rebelo, S., 2011. Do peso problems explain the returns to the carry trade?. Review of Financial Studies 24, 853–891.

Christiansen, C., Ranaldo, A., Söderlind, P., 2011. The time-varying systematic risk of carry trade strategies. Journal of Financial and Quantitative Analysis 46, 1107–1125.

Clarida, R., Davis, J., Pedersen, N., 2009. Currency carry trade regimes: Beyond the Fama regression. Journal of International Money and Finance 28, 1375–1389.

Clark, T., West, K., 2007. Approximately normal tests for equal predictive accuracy in nested models. Journal of Econometrics 138, 291–311.

Cochrane, J., 2005. Asset Pricing. Princeton University Press, Princeton, NJ.

Daniel, K., Hodrick, R., Lu, Z., 2017. The carry trade: Risks and drawdowns. Critical Finance Review (forthcoming).

DeMiguel, V., Nogales, F., Uppal, R., 2014. Stock return serial dependence and out-of-sample portfolio performance. Review of Financial Studies 27, 1031–1073.

Hassan, T., 2013. Country size, currency unions, and international asset returns. Journal of Finance 68, 2269–2308.

Hassan, T., Mano, R., 2015. Forward and spot exchange rates in a multi-currency world. Working paper. University of Chicago.

Hodrick, R., 1992. Dividend yields and expected stock returns: Alternative procedures for inference and measurement. Review of Financial Studies 5, 357–386.

Jagannathan, R., Wang, Z., 1996. The conditional CAPM and the cross-section of expected returns. The Journal of Finance 51, 3–53.

Jobson, J. D., Korkie, B. M., 1981. Performance hypothesis testing with the Sharpe and Treynor measures. Journal of Finance 36, 889–908.

Jordá, O., Taylor, A., 2012. The carry trade and fundamentals: Nothing to fear but FEER itself. Journal of International Economics 88, 74–90.

Jurek, J., 2014. Crash-neutral currency carry trades. Journal of Financial Economics 113, 325–347.

Koijen, R., Moskowitz, T., Pedersen, L., Vrugt, E., 2015. Carry. Unpublished working paper. London Business School.

Ledoit, O., Wolf, M., 2008. Robust performance hypothesis testing with the Sharpe ratio. Journal of Empirical Finance 15, 850–859.

Lettau, M., Maggiori, M., Weber, M., 2015. Conditional risk premia in currency markets and other asset classes. Journal of Financial Economics 114, 197–225.

Lustig, H., Roussanov, N., Verdelhan, A., 2011. Common risk factors in currency markets. Review of Financial Studies 24, 3731–3777.

Lustig, H., Roussanov, N., Verdelhan, A., 2014. Countercyclical currency risk premia. Journal of Financial Economics 111, 527–553.

Maggiori, M., 2013. The U.S. Dollar safety premium. Unpublished working paper. Harvard University.

Maurer, T., Tô, T.-D., Tran, N.-K., 2016. Pricing risks across currency denominations. Unpublished working paper. Washington University in St. Louis.

Melvin, M., Taylor, M., 2009. The crisis in the foreign exchange market. Journal of International Money and Finance 28, 1317–1330.

Memmel, C., 2003. Performance hypothesis testing with the Sharpe ratio. Finance Letters 1, 21–23.

Menkhoff, L., Sarno, L., Schmeling, M., Schrimpf, A., 2012. Carry trades and global foreign exchange volatility. Journal of Finance 67, 681–718.

Ready, R., Roussanov, N., Ward, C., 2017. Commodity trade and the carry trade: A tale of two countries. Journal of Finance 72, 2629–2684.

Verdelhan, A., 2017. The share of systematic variation in bilateral exchange rates. Journal of Finance (forthcoming).

Table 1

Some recent approaches to carry trade construction

This table summarizes three aspects of carry trade constructions, as adopted in recent studies. It shows which currencies are employed in the trade, are currencies given equal weights (possibly among other weighting schemes), and whether the total weights of the long and short sides of the trade are equal. The bottom part of the table provides similar details on several investable carry trade indexes (source: FX Week (Feb, 2008)).

	Which currencies?	Weights equal?	Long and short equal?
Brunnermeier et al. (2009)	G-10	yes	yes
Clarida et al. (2009)	G-10	yes	yes
Jorda and Taylor (2009)	G-10 ex SEK	yes	no
Ang and Chen (2010)	G-10 plus 13 other	yes	yes
Burnside et al. (2011)	20 developed	yes	no
Lustig et al. (2011)	{ nine to 34, plus a smaller subset of 15 developed	yes	yes
Lustig et al. (2014)		yes	no
Menkhoff et al. (2012)		yes	yes
Ready et al. (2013)		yes	yes
Bakshi and Panayotov (2013)	G-10	yes	yes
Hassan and Mano (2015)	nine to 39, plus a smaller subset of 15 developed	no (deviations of forward differentials from their mean)	yes
Daniel et al. (2017)	G-10	yes and no (various weighting schemes)	yes and no
Jurek (2014)	G-10	yes and no (various weighting schemes)	yes and no
Barroso and Santa-Clara (2014)	27 developed	no (vol. scaled deviations of forward differentials from their mean)	yes
Some investable carry trade indexes:			
Barclays Capital: Intelligent Carry Index	G-10	no (portfolio optimization)	yes
Citigroup: Beta1 range	G-10	yes (uses 9 or 13 most liquid tradeable pairs)	yes
Credit Suisse: Rolling Optimised Carry Indices	G-10 and G-18	no (portfolio optimization)	yes
Deutsche Bank: G-10 Harvest Index	G-10	yes (three highest- and lowest-yielding)	yes
JP Morgan: IncomeFX	G-10	yes and no (four pairs selected each month to optimize the risk-return ratio)	yes

Table 2

Carry trades constructed by sequentially excluding G-10 currencies

Using mid-quotes for spot (S_t) and one-month forward (F_t) exchange rates of the G-10 currencies against the USD (Euro spliced with DEM before 1999) from Datastream, we calculate forward differentials as $F_t/S_t - 1$. With these we construct carry trades either using all G-10 currencies, or excluding one up to seven of these currencies, as explained in Section 3.1. Panel A reports their averages (denoted "avg. ret.", annualized an in percent) and SR's (both annualized), as well as skewness (denoted "skew"). The first column of the table shows how many currencies have been excluded. The columns denoted "p-val" are obtained using bootstrap confidence intervals for the hypothesis that the respective SR or skewness does not exceed the benchmark one (see Appendix OA-II). The sample period is 12/1984 to 6/2014 (354 months), and the initial window for calculating SR's that are used to decide which currencies should be excluded is 120 months. Each column in Panel B shows how many months, out of the 234 months in which we search for the highest SR's, was a currency with code displayed in the first row, the first to be excluded (row starting with "1"), or among the first two excluded (row starting with "2"), etc.

	A. Average returns, Sharpe ratios, skewness					B. Frequency and order of exclusion									
	avg. ret.	SR	p-val	skew	p-val	NZD	AUD	NOK	GBP	SEK	CAD	USD	EUR	CHF	JPY
SC	1.00	0.32		-0.33											
1	1.05	0.32	0.51	-0.26	0.24	0	135	50	0	0	0	0	0	0	49
2	1.23	0.41	0.22	-0.57	0.87	0	184	219	0	0	0	0	15	0	50
3	1.25	0.41	0.25	0.01	0.03	0	192	219	0	0	0	0	74	3	214
4	1.34	0.46	0.21	0.21	0.01	14	224	234	0	0	82	0	152	10	220
5	1.56	0.40	0.32	-0.01	0.10	106	224	234	0	41	82	0	230	19	234
6	1.47	0.44	0.29	-0.35	0.53	171	234	234	36	58	175	0	230	32	234
7	3.17	0.61	0.10	0.04	0.10	172	234	234	63	58	175	0	234	234	234

Table 3

Carry trades constructed with fixed subsets of the G-10 currencies

This table shows averages (denoted ”avg. ret.”, in percent), standard deviations (”st.dev.”, in percent) and Sharpe ratios (”SR”), all annualized, as well as skewness (denoted ”skew”) for the monthly excess returns of several carry trades, all with equal weights. ”SC” denotes the standard carry trade constructed with all G-10 currencies. The G1 to G5 trades are constructed using the currencies with the displayed codes, and represent combinations of currencies that are less often excluded by the enhancement rule discussed in Section 3.1 and Table 2. B1 to B5 are the complementary trades, each using the currencies that have been left out of one of the G1-G5 trades. ”GC” denotes the set of 18 carry trades, each constructed from the three least often excluded currencies (USD, GBP and SEK), combined with any possible pair of the remaining G-10 currencies that contains *none* or *only one* of the three most often excluded currencies (AUD, NOK and JPY). ”BC” denotes the set of 18 complementary carry trades, each using the five currencies left out of one of the 18 trades in GC. The rows corresponding to the GC and BC trades show *averages* of the average returns, standard deviations, Sharpe ratios and skewness across the respective 18 carry trades. The columns denoted ”p-val” show p-values obtained using bootstrap confidence intervals (see Appendix OA-II). The first (second) number in parentheses shows how many of the 18 corresponding individual estimates are significant at the 5% (10%) confidence level. Where p-values are not in square brackets, the null hypothesis is that the respective SR or skewness does not exceed the one of the SC trade. Where p-values are in square brackets, the null is that the SR or skewness of a G1-G5 trade or GC trade does not exceed that of the corresponding B1-B5 trade or BC trade. The last two columns show correlations (denoted ”corr”) with the returns of the SC trade, and correlations with only the negative returns of the SC trade (denoted ”down”). The rows corresponding to the GC and BC trades show averages of the respective correlations. The sample period is 12/1984 to 6/2014 (354 months).

		avg. ret.	st.dev.	SR	p-val	skew	p-val	corr	down
SC		1.02	3.30	0.31		-0.22			
G1	NZD, GBP, SEK, CAD, USD, EUR, CHF	1.67	3.29	0.51	0.02	0.07	0.02	0.87	0.70
G2	GBP, SEK, CAD, USD, CHF	1.70	3.47	0.49	0.13	-0.17	0.41	0.63	0.39
G3	NZD, GBP, SEK, USD, CHF	2.49	4.09	0.61	0.01	-0.21	0.48	0.77	0.51
G4	NZD, GBP, SEK, CAD, USD	2.22	4.39	0.51	0.12	-0.01	0.19	0.61	0.39
G5	GBP, SEK, USD	3.97	5.71	0.69	0.03	-0.23	0.50	0.40	0.18
B1	AUD, NOK, JPY	0.68	7.50	0.09	[0.01]	-0.92	[0.01]	0.72	0.70
B2	NZD, AUD, NOK, EUR, JPY	0.98	5.54	0.18	[0.07]	-0.60	[0.06]	0.84	0.75
B3	AUD, NOK, CAD, EUR, JPY	0.21	4.91	0.04	[0.01]	-0.85	[0.04]	0.84	0.79
B4	AUD, NOK, EUR, CHF, JPY	0.28	4.96	0.06	[0.02]	-0.87	[0.04]	0.78	0.73
B5	NZD, AUD, NOK, CAD, EUR, CHF, JPY	0.61	4.66	0.13	[0.01]	-0.63	[0.17]	0.88	0.79
GC		1.96	4.11	0.47	(6/7)	-0.33	(0/0)	0.68	0.49
BC		0.66	5.13	0.13	[(8/12)]	-0.66	[(5/11)]	0.79	0.69

Table 4

Currency HML vs. good carry trades as currency market pricing factors

Several factor pricing models are estimated with GMM, as described in Section 4, over 12/1984 to 12/2013. The 11 test assets (six "All" and five "Developed" interest-rate sorted currency portfolios (net returns)) and the RX (dollar) and currency HML (denoted "HML^{FX}") factors ("All" version, net) are as in Lustig et al. (2011), and kindly made available at Adrian Verdelhan's website. The good carry trades G1-G5 are as in Table 3. All models include the RX factor, and either the HML^{FX} or a good carry trade (as indicated in the first column), or both. The top panel reports *averages* of the 11 annualized average returns, time-series regression coefficients and adjusted R^2's (in percent). Standard errors are estimated with GMM and account for one Newey-West lag. The first (second) number in parentheses shows how many of the 11 corresponding estimates are significant at the 5% (10%) confidence level. The bottom panel shows, for the same models, factor risk prices λ and SDF coefficients b with p-values, as well as p-values for the χ^2 statistic testing that the pricing errors are jointly equal to zero. Average returns, α's and λ's are reported annualized and in percent. β_{Good}, λ_{Good} and b_{Good} refer to the good carries G1-G5, as shown in the first column.

	avg. ret.	p-val	α	p-val	β_{RX}	p-val	$\beta_{HML^{FX}}$	p-val	β_{Good}	p-val	R^2
	2.33	(3/4)	0.21	(1/2)	1.12	(11/11)	-0.05	(8/8)			81.1
G1			0.06	(1/1)	1.11	(11/11)			-0.01	(10/10)	78.3
G2			0.06	(1/1)	1.11	(11/11)			-0.01	(9/10)	75.9
G3			0.06	(0/1)	1.12	(11/11)			-0.01	(9/9)	77.3
G4			0.01	(1/1)	1.11	(11/11)			0.02	(8/9)	75.3
G5			-0.01	(0/3)	1.11	(11/11)			0.02	(7/7)	74.0
G1			0.15	(0/1)	1.12	(11/11)	-0.06	(5/6)	0.08	(6/8)	83.0
G2			0.16	(0/1)	1.12	(11/11)	-0.05	(8/8)	0.05	(10/10)	81.9
G3			0.12	(0/1)	1.11	(11/11)	-0.06	(7/8)	0.07	(5/5)	82.2
G4			0.15	(0/1)	1.1	(11/11)	-0.05	(8/8)	0.06	(5/6)	82.1
G5			0.13	(0/0)	1.11	(11/11)	-0.05	(8/8)	0.03	(7/7)	81.6

	λ_{RX}	p-val	$\lambda_{HML^{FX}}$	p-val	λ_{Good}	p-val	b_{RX}	p-val	$b_{HML^{FX}}$	p-val	b_{Good}	p-val	$\chi^2_{pr.err.}$
	2.29	0.08	4.24	0.03			4.03	0.18	4.75	0.06			0.00
G1	2.13	0.10			2.09	0.02	3.43	0.24			17.80	0.03	0.17
G2	2.13	0.10			2.90	0.03	3.08	0.30			22.47	0.04	0.50
G3	2.12	0.10			2.91	0.02	1.70	0.59			15.91	0.03	0.09
G4	2.03	0.12			4.29	0.01	-4.38	0.35			24.98	0.02	0.22
G5	1.99	0.13			8.64	0.01	-8.54	0.13			29.83	0.01	0.47
G1	2.13	0.11	3.23	0.08	2.10	0.01	3.42	0.25	-0.08	0.98	18.03	0.03	0.27
G2	2.17	0.10	3.68	0.05	2.46	0.05	3.24	0.27	1.11	0.71	18.10	0.12	0.27
G3	2.11	0.11	3.29	0.07	3.03	0.01	1.49	0.64	-0.55	0.86	17.45	0.06	0.05
G4	2.09	0.11	3.31	0.07	3.74	0.00	-2.85	0.50	1.31	0.60	20.26	0.02	0.24
G5	2.13	0.10	3.88	0.04	6.18	0.01	-4.57	0.30	2.66	0.25	20.05	0.01	0.97

Table 5

Carry return predictability

This table shows results from univariate predictive regressions for the log returns of the SC trade, the G1-G5 and B1-B5 trades, and the GC and BC trades as described in Table 3. The three predictors shown in the first row of the table have been found to predict carry trade returns in Bakshi and Panayotov (2013, Table 2) and Ready, Roussanov, and Ward (2017, Table 10), and are designed as in those studies (see also Section 4.4). The table displays the in-sample estimates of the predictive slope coefficients β, two-sided p-values, based on the Hodrick (1992) 1B covariance matrix estimator, and adjusted R^2's (in percent). Next are shown one-sided p-values (denoted "MS") for the MSPE-adjusted statistic (Clark and West (2007), see also footnote 4), obtained with an expanding window with initial length of 120 months. The columns denoted "Δ SR" report a measure of the economic significance of predictability, based on a prediction-based trading strategy, which enters into a carry trade at the end of month t only if the trade's return predicted for month $t+1$ is positive (if a negative return is predicted, the carry trade return for month $t+1$ is zero). Specifically, the measure equals the difference between the Sharpe ratio of the prediction-based strategy, implemented with the respective subset of G-10 currencies, and the corresponding unconditional carry trade (using an expanding window with initial length of 120 months). For the GC and BC trades the table shows averages of the respective 18 predictive slope coefficients β, adjusted R^2's, and changes in Sharpe ratios. The first (second) number in parentheses shows how many of the 18 corresponding individual estimates are significant at the 5% (10%) confidence level. The sample period is 12/1984 to 6/2014.

	Commodity returns					Change in exchange rate volatility					Change in BDI				
	β	p-val	R^2	MS	Δ SR	β	p-val	R^2	MS	Δ SR	β	p-val	R^2	MS	Δ SR
SC	0.015	0.12	0.8	0.45	-0.04	-0.005	0.00	2.4	0.01	0.10	0.003	0.15	1.0	0.18	0.08
G1	0.006	0.53	0.1	0.99	-0.04	-0.003	0.12	0.8	0.22	-0.04	0.003	0.16	0.9	0.21	0.10
G2	0.013	0.19	0.5	0.85	-0.05	-0.003	0.21	0.6	0.30	0.00	0.002	0.39	0.3	0.79	0.00
G3	0.008	0.50	0.1	0.82	0.01	-0.005	0.05	1.6	0.13	-0.05	0.005	0.02	2.3	0.02	0.09
G4	0.012	0.52	0.3	0.88	-0.03	-0.006	0.04	1.9	0.21	-0.03	0.002	0.48	0.4	0.71	-0.01
G5	-0.005	0.82	0.0	0.94	0.00	-0.003	0.35	0.4	0.60	0.03	-0.002	0.68	0.1	0.97	-0.05
B1	0.069	0.01	3.3	0.02	0.29	-0.010	0.02	1.7	0.01	0.22	0.010	0.05	2.3	0.04	-0.14
B2	0.043	0.04	2.3	0.06	0.15	-0.008	0.01	2.1	0.01	0.13	0.010	0.01	4.2	0.01	0.09
B3	0.032	0.04	1.6	0.10	-0.02	-0.007	0.01	2.3	0.01	0.32	0.004	0.15	1.0	0.13	-0.15
B4	0.040	0.03	2.5	0.06	0.12	-0.007	0.02	1.9	0.02	0.23	0.007	0.06	2.3	0.05	0.09
B5	0.029	0.05	1.4	0.12	-0.01	-0.006	0.01	1.9	0.01	0.14	0.007	0.02	2.8	0.01	0.06
GC	0.010	(0/3)	0.5	(0/0)	-0.04	-0.005	(11/14)	1.4	(3/5)	0.00	0.002	(2/3)	0.8	(2/3)	-0.02
BC	0.033	(7/14)	1.7	(1/9)	0.03	-0.007	(15/17)	1.8	(17/18)	0.17	0.008	(12/14)	2.9	(12/15)	0.07

Table 6
Carry components in carry trade returns
This table shows the average returns, and the components of these returns due to carry and currency depreciation, for the standard (SC) trade, the good and bad carry trades G1 to G5 and B1 to B5, and the dollar carry trade (DC) of Lustig, Roussanov, and Verdelhan (2014), which we consider in more detail in Section 6.2. The respective values are shown annualized and in percent. Also shown are in curly brackets p-values from a test for equality of the respective average carry component to that of the SC and DC trades (Newey-West standard errors, with 12 lags).

	SC	B1	B2	B3	B4	B5	G1	G2	G3	G4	G5	DC
avg. total ret.	1.02	0.68	0.98	0.21	0.28	0.61	1.67	1.70	2.49	2.22	3.97	4.19
avg. deprec.	-0.37	-2.02	-1.25	-1.58	-1.65	-1.28	0.26	0.50	0.78	1.01	2.76	3.27
avg. carry	1.38	2.70	2.23	1.78	1.92	1.88	1.40	1.20	1.70	1.20	1.20	0.92
carry to total ret.	1.36	3.98	2.27	8.46	6.96	3.08	0.84	0.70	0.68	0.54	0.30	0.22
comparing carry components (p-values)												
to that of SC		{0.00}	{0.00}	{0.00}	{0.00}	{0.00}	{0.63}	{0.08}	{0.00}	{0.04}	{0.32}	{0.01}
to that of DC		{0.00}	{0.00}	{0.00}	{0.00}	{0.00}	{0.01}	{0.04}	{0.00}	{0.04}	{0.01}	

Table 7
Unbiasedness hypothesis regressions for pairs of prototypical carry trade currencies

For each of the six currency pairs NZD/CHF, AUD/CHF, NOK/CHF, NZD/JPY, AUD/JPY and NOK/JPY monthly log changes in spot rates are regressed against the corresponding forward differentials (all mid-quotes): $\ln(S_{t+1}/S_t) = \alpha + \beta \ln(F_t/S_t) + \varepsilon_{t+1}$. Estimates of α and β are shown, together with two-sided p-values, based on the Hodrick (1992) 1B covariance matrix estimator, for the null hypotheses $\alpha = 0$ and $\beta = 1$. The sample period is 12/1984 - 6/2014, and the data source is Barclays Bank, via Datastream.

	CHF				JPY			
	α	p-val	β	p-val	α	p-val	β	p-val
NZD	0.001	0.74	0.46	0.31	0.003	0.45	0.63	0.49
AUD	-0.002	0.58	0.15	0.30	0.002	0.57	0.92	0.92
NOK	-0.001	0.72	0.43	0.17	0.007	0.03	2.22	0.10

Table 8

Dynamic nature of carry trades

The table characterizes the dynamic behavior of the the standard carry trade (SC), the G1 to G5 and B1 to B5 trades, and, in the last row, the dollar carry trade (DC) of Lustig, Roussanov, and Verdelhan (2014), which we consider in more detail in Section 6.2. The first column of the table shows, for each trade, the time-series average of the proportion of currencies that change (switch) position, from long to short or vice versa, at each point of time. The average proportion is given in percent. The next two columns show bootstrapped 95% confidence intervals for these average proportions. The column denoted "days w/o switch" shows the proportion of dates in the sample when not a single currency changed position from short to long or vice versa. Next the table shows the correlation between the proportions for the G1 to G5 and B1 to B5 trades, and those for standard carry (SC). The last three columns aim to compare static and dynamic versions of our various trades. *Static* trades have been defined in Hassan and Mano (2015), and, to keep close to their setup, we use as weights the average forward differentials of the respective currencies over 12/1984-12/1994, demeaned and normalized to have absolute values that sum to one. These weights are kept fixed for the rest of the sample period for the static trades, without ever re-balancing. *Dynamic* trades are the usual (dynamically re-balanced) trades, as considered throughout this paper, but with weights again equal to the cross-sectionally demeaned forward differentials, normalized to have absolute values that sum to one. Shown are the ratios between the average returns of the respective static and dynamic carry trade, and the corresponding Sharpe ratios. Average returns and Sharpe ratios are calculated for the period 12/1994 to 6/2014.

	switch (%)	95% conf. int.		days w/o switch	correl. with SC	ratio of static to dynamic avg. ret.	Sharpe ratios: dynamic	Sharpe ratios: static
SC	8.22	[5.72	11.05]	0.68		0.45	0.41	0.21
G1	12.75	[9.27	16.47]	0.66	0.82	0.23	0.51	0.14
G2	12.97	[9.35	16.88]	0.73	0.75	-0.23	0.34	-0.10
G3	11.90	[8.73	15.18]	0.75	0.59	0.24	0.55	0.14
G4	17.22	[13.20	21.25]	0.63	0.63	0.28	0.50	0.13
G5	16.34	[11.99	20.96]	0.78	0.49	-0.01	0.69	-0.01
B1	9.82	[5.67	14.64]	0.86	0.53	1.07	0.21	0.23
B2	9.97	[6.35	13.94]	0.78	0.56	1.16	0.22	0.29
B3	12.18	[8.44	16.60]	0.73	0.59	0.75	0.25	0.22
B4	10.14	[7.25	13.20]	0.78	0.50	0.61	0.24	0.16
B5	10.12	[6.92	13.60]	0.71	0.65	0.82	0.27	0.24
DC	6.80	[3.40	11.05]	0.93	0.47			

Table 9

Correlations and spanning

Panel A shows return correlations between SC and DC and the G1 to G5 and B1 to B5 carry trades. In Panel B, ”Good” refers to the good carries G1-G5 shown in the first column, the LHS variable is that for which no regression coefficient is reported on the respective row. Intercepts are reported annualized and in percent.

A. Correlations

	DC	G1	G2	G3	G4	G5	B1	B2	B3	B4	B5
SC	0.40	0.87	0.63	0.77	0.62	0.41	0.72	0.84	0.84	0.78	0.88
DC		0.41	0.42	0.48	0.61	0.68	0.17	0.15	0.15	0.07	0.09

B. Return regressions

	interc.	p-val	β_{SC}	p-val	β_{DC}	p-val	β_{Good}	p-val	R^2
G1	-0.50	0.10			0.03	0.05	0.85	0.00	0.75
G2	-0.20	0.67			0.07	0.00	0.54	0.00	0.42
G3	-0.56	0.16			0.02	0.36	0.61	0.00	0.59
G4	-0.03	0.94			0.01	0.56	0.45	0.00	0.37
G5	0.02	0.97			0.10	0.00	0.14	0.00	0.18
G1	2.91	0.02	0.42	0.05			0.51	0.01	0.16
G2	2.71	0.02	0.47	0.00			0.59	0.00	0.20
G3	2.18	0.07	0.15	0.36			0.75	0.00	0.22
G4	1.97	0.06	0.07	0.56			0.97	0.00	0.37
G5	0.74	0.44	0.31	0.00			0.79	0.00	0.48
G1	0.69	0.03	0.84	0.00	0.03	0.02			0.75
G2	0.70	0.15	0.58	0.00	0.10	0.00			0.43
G3	1.13	0.02	0.86	0.00	0.12	0.00			0.63
G4	0.50	0.37	0.59	0.00	0.27	0.00			0.53
G5	1.63	0.03	0.27	0.00	0.49	0.00			0.48

Table 10

Dollar carry vs. good carry trades as currency market pricing factors

This table differs from Table 4 only in one aspect: instead of the RX and HML^{FX} factors, we now use the dollar carry factor (denoted ”DC”), as in Lustig et al. (2014) and Hassan and Mano (2015), calculated for the G-10 currencies. The test assets are again the 11 interest rate-sorted portfolios as in Lustig et al. (2011). ”Good” refers to the good carries G1-G5.

	avg. ret.	p-val	α	p-val	β_{DC}	p-val	β_{Good}	p-val	R^2
	2.33	(3/4)	-0.04	(0/0)	0.56	(11/11)			21.2
G1			1.88	(0/0)			0.26	(9/9)	5.3
G2			1.89	(0/2)			0.25	(7/7)	3.1
G3			1.02	(0/0)			0.52	(7/7)	8.6
G4			0.26	(0/0)			0.93	(11/11)	21.9
G5			-0.41	(0/0)			0.67	(11/11)	18.8
G1			0.21	(0/0)	0.55	(11/11)	-0.27	(7/7)	26.8
G2			0.23	(0/0)	0.56	(11/11)	-0.29	(5/6)	24.5
G3			-0.12	(0/0)	0.49	(11/11)	0.06	(7/8)	24.7
G4			-0.43	(0/0)	0.31	(8/9)	0.58	(10/10)	27.2
G5			-0.67	(0/0)	0.34	(10/11)	0.34	(7/9)	23.8

	λ_{DC}	p-val	λ_{GC}	p-val	b_{DC}	p-val	b_{Good}	p-val	$\chi^2_{pr.err.}$
	0.45	0.07			8.3	0.07			0.03
G1			0.32	0.01			28.4	0.01	0.99
G2			0.46	0.03			37.0	0.02	0.32
G3			0.35	0.04			21.6	0.02	0.13
G4			0.27	0.07			13.7	0.06	0.14
G5			0.37	0.07			11.1	0.06	0.95
G1	4.85	0.05	2.27	0.01	6.3	0.21	14.5	0.09	0.77
G2	5.19	0.04	2.93	0.02	5.8	0.27	18.1	0.10	0.99
G3	4.04	0.12	2.95	0.01	3.7	0.52	13.6	0.09	0.15
G4	1.15	0.75	3.49	0.01	-6.8	0.53	24.5	0.10	1.00
G5	3.12	0.31	3.88	0.13	-3.4	0.81	15.5	0.37	0.16

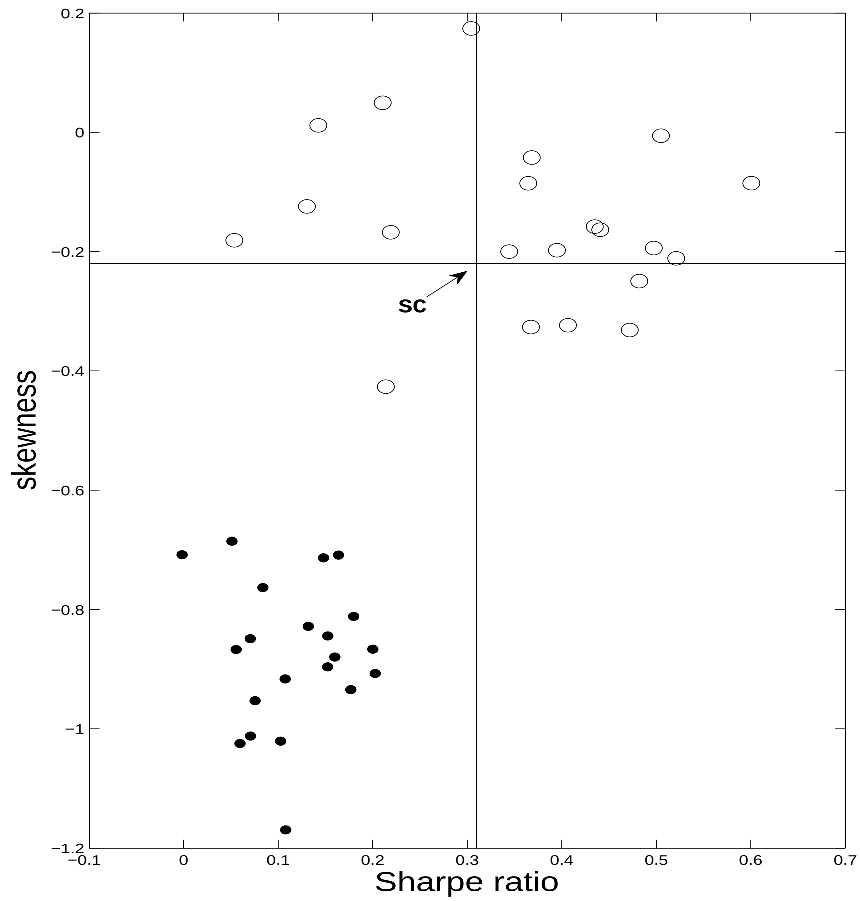

Figure 1: **Good and bad carry trades from sets of five G-10 currencies**

Large black dots plot skewness versus Sharpe ratio of all possible 21 carry trades trades constructed from five G-10 currencies, which include the AUD, CHF and JPY, together with any possible pair from the remaining seven currencies. Circles with no fill plot similarly skewness versus Sharpe ratio for the complementary trades, each including the five currencies left out of one of the previous 21 trades. For each trade, currencies are sorted on their forward differentials (against the USD) at the end of each month over the period 12/1984 to 6/2014, and the two currencies with highest differentials are held long over the next month, while the two with the lowest premiums are shorted, all with equal weights. A vertical and horizontal lines indicate the Sharpe ratio and return skewness of the standard carry trade (denoted SC), constructed with all G-10 currencies. Percentage carry trade returns are calculated with spot and forward quotes from Barclays Bank, available via Datastream, and with transaction costs taken into account.

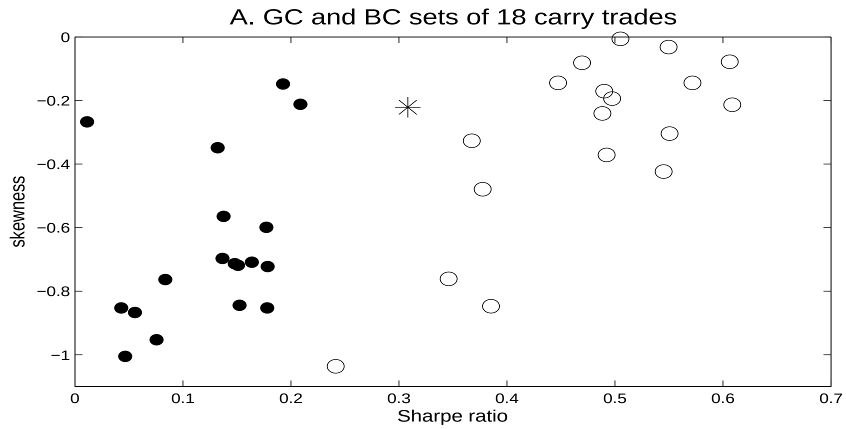

Figure 2: **Sharpe ratios versus skewness for sets of good and bad carry trades**

In Panel A, circles with no fill plot skewness versus Sharpe ratio for 18 carry trades, each constructed from five of the G-10 currencies, with equal weights. Each of these trades uses the three currencies (GBP, SEK and USD) which are least often excluded by the enhancement rule in Section 3.1 and Table 2. These three currencies are combined with any possible pair of the remaining G-10 currencies, which contains *none* or *only one* of the three most often excluded currencies (AUD, NOK and JPY). Large black dots plot similarly the skewness versus Sharpe ratio for the complementary carry trades, which use the five currencies left out of one of the previous 18 trades. These two sets of 18 trades are denoted in Section 3.4 and Table 3 and others as GC and BC. In Panel B, circles with no fill or black dots plot skewness versus Sharpe ratio for two alternative sets of 15 trades, representing good or bad carry trades and also constructed from five G-10 currencies with equal weights. The ”good” trades exclude the JPY and include the GBP, SEK, USD and any other pair of the remaining six currencies, while each of the complementary ”bad” trades includes the five currencies left out of one of the 15 ”good” trades. Both panels also show the standard carry trade (SC), denoted by a star. The sample period is 12/1984 to 6/2014.

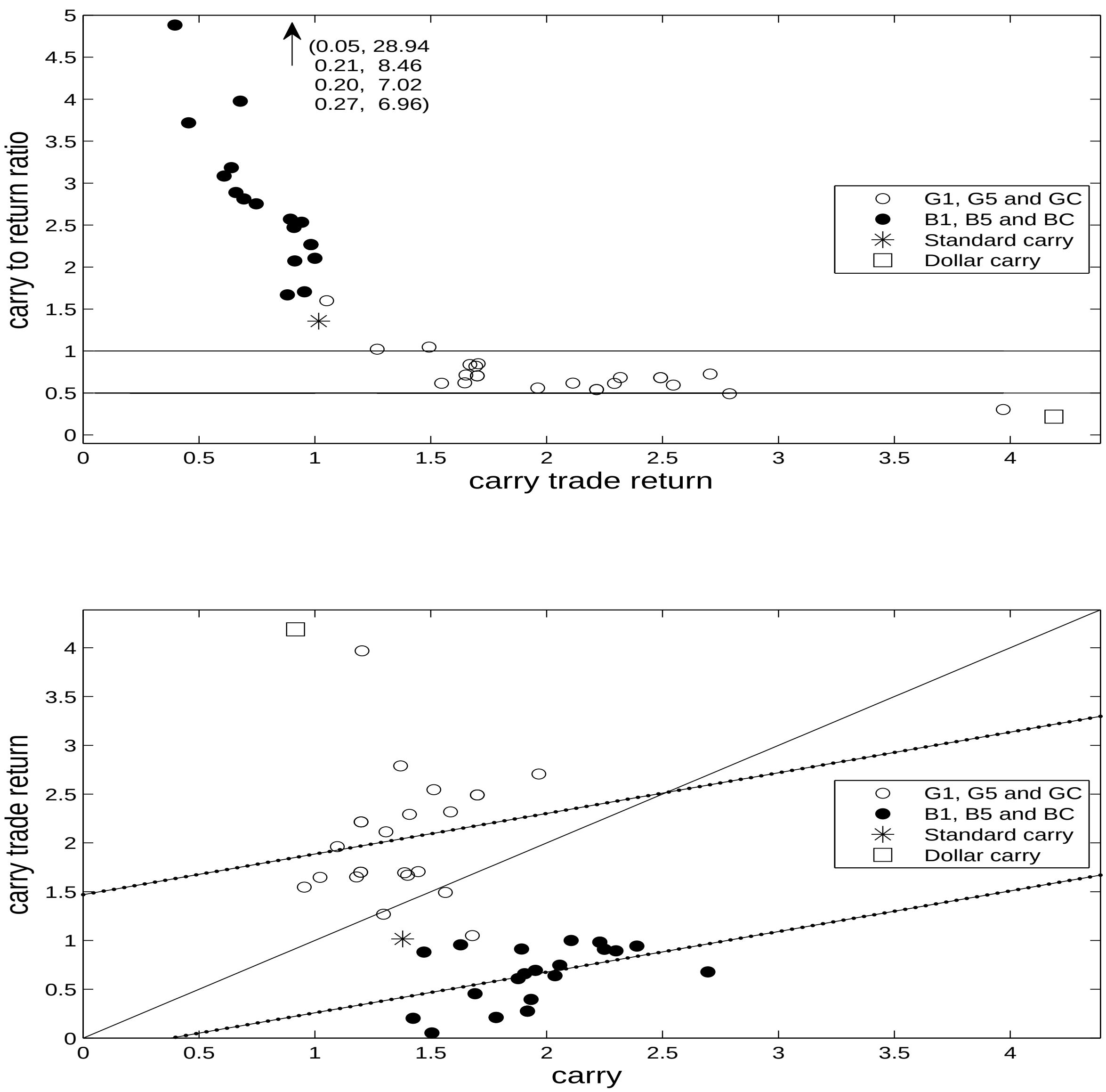

Figure 3: **Decomposing carry trade returns**

For all trades considered in this paper (18 GC trades, 18 BC trades, as well as G1 and G5, and B1 and B5) Panel A of the figure plots total average returns (horizontal axis) versus the ratios of average carry to total return (vertical axis). As previously, white (black) dots correspond to good (bad) trades. For visual clarity, four outlier points (all referring to bad carry trades) are not shown on this plot, but their coordinates are displayed in the top left corner. Panel A plots similarly the standard and dollar carry trades (SC and DC). Horizontal lines correspond to carry-to-return ratio of one (all average return comes from carry alone), and one half (return comes equally from carry and exchange rate changes). For the same trades, Panel B plots carry versus total return, together with the 45-degree line, corresponding again to carry-to-return ratio of one. The two additional parallel lines in Panel B are regression lines from equation (5), which includes a dummy variable for the intercept for good carry trades.

Good Carry, Bad Carry

Online Appendix: Not for Publication

OA-I Symmetry and numeraire neutrality of currency trades

This Appendix explains in detail the distinction between several designs of carry trades. Start with a set of N currencies, e.g. the G-10 currencies in our case. A currency trading strategy is a mapping between signals at time t and currency positions taken at this time, whereby positions are defined in terms of the weights of individual currencies. A trading strategy is formulated relative to a benchmark currency, i.e. positions are taken relative to a certain currency in the forward market. From this perspective, two properties seem important:

1. *Symmetry*: the number of short and long positions and their total weights are equal. A stronger version of symmetry would also require equal weights of the individual short or long positions.[1]

2. *Numeraire independence*: the positions taken in the various currencies are the same, regardless of which benchmark currency is considered. As a result, only one currency strategy must be defined for the world at large.

Symmetry and numeraire independence are well-established features of carry trades, and have been both adopted by recent academic studies, and implemented in investable products (see Table 1). Together, these properties imply that the trade's returns will be very similar from any currency perspective. This invariance follows from the fact that the translation of returns from one currency to another simply introduces cross-currency risk on currency returns, which is a second order effect. Conversely, if the ranking of a currency or the signal depends in any way on the identity of the benchmark currency, then defining the same strategy from another currency perspective will yield different currency positions and different currency weights, and this can result in quite different returns. A well-known example is the asymmetric carry strategy in Burnside, Eichenbaum, Kleshchelski, and Rebelo (2011), which has been shown in Daniel, Hodrick, and

[1] However, if weights are defined relative to a benchmark currency (e.g., based on forward differentials), they may differ on the long and short end, creating weight asymmetry. This would cause the trade to be numeraire-dependent.

Lu (2017) to produce very different (and worse) returns from other, non-USD currency perspectives. In fact, the USD-based version of this strategy is successful (at least partly) due to is its implicit exposure to a dollar-centric currency strategy, the "dollar carry" trade of Lustig, Roussanov, and Verdelhan (2014).[2]

We now formally show that symmetric, numeraire-independent strategies have largely equivalent returns across the world. Suppose first that the USD is the benchmark currency and define the weight of currency i as w_i. Spot and forward exchange rates are quoted here as USD per one unit of a foreign currency (reversing the notation from Section 2 above), and denoted as S_t^i and F_t^i. The return of a US-based currency trading strategy over the interval t to $t+1$ is:

$$r_{t+1}^{USD} = \sum_{i=1}^{N} w_i [S_{t+1}^i / F_t^i - 1]. \tag{OA-1}$$

If the strategy is numeraire independent, the weights w_i are *identical* for all currency perspectives. For example, if the trading strategy is based on interest rate signals, these signals should be independent of the benchmark currency.

Defining such a strategy relative, say, to the Japanese yen, with yen exchange rates denoted by $\overline{S}_t^i$ and $\overline{F}_t^i$ (JPY per one unit of currency i), its return (in yen) is:

$$r_{t+1}^{JPY} = \sum_{i=1}^{N} w_i [\overline{S}_{t+1}^i / \overline{F}_t^i - 1] = \sum_{i=1}^{N} w_i \overline{S}_{t+1}^i / \overline{F}_t^i - \sum_{i=1}^{N} w_i \tag{OA-2}$$

With symmetric strategies the weights sum to zero, hence the last term cancels, and we are left with $r_{t+1}^{JPY} = \sum_{i=1}^{N} w_i \overline{S}_{t+1}^i / \overline{F}_t^i$. By triangular arbitrage and symmetry, we can further derive:

$$r_{t+1}^{JPY} = [r_{t+1}^{USD} + \sum_{i=1}^{N} w_i] * \overline{S}_{t+1}^{USD} / \overline{F}_t^{USD} = r_{t+1}^{USD} * F_t^{JPY} / S_{t+1}^{JPY}. \tag{OA-3}$$

[2]The dollar carry weights are 1/(N-1) and are all either positive or negative depending on the average interest rate of the USD relative to other currencies. The weight on the USD itself is zero. The strategy is thus very asymmetric and yields entirely different results for other currency perspectives. That is, a "British pound carry" or "Swiss franc carry" need not be anything like dollar carry. Of course, dollar-centric strategies are of interest because of the importance of the dollar in international finance.

Cross-currency risk could drive, in principle, a wedge between the two currency perspectives, but in practice the returns and their properties will be rather similar (barring significant differences in transaction costs), because the forward to spot ratio in (OA-3) is close to one, and applies to returns. We have verified that standard carry strategies (as per our definition in Section 2) yield very similar returns from any currency perspective.

It is instructive to repeat the previous calculation, but for *log* returns. In this case:

$$\begin{aligned}
r_{t+1}^{JPY} &= \sum_{i=1}^{N} w_i \log\left(\overline{S}_{t+1}^{i}/\overline{F}_{t}^{i}\right) = \sum_{i=1}^{N} w_i \log\left(\frac{S_{t+1}^{i}}{F_t^{i}} \; \frac{\overline{S}_{t+1}^{USD}}{\overline{F}_{t}^{USD}}\right) \\
&= \sum_{i=1}^{N} w_i \log(S_{t+1}^{i}/F_t^{i}) + \sum_{i=1}^{N} w_i \log(\overline{S}_{t+1}^{USD}/\overline{F}_{t}^{USD}) \\
&= \sum_{i=1}^{N} w_i \log(S_{t+1}^{i}/F_t^{i}) + \log(\overline{S}_{t+1}^{USD}/\overline{F}_{t}^{USD}) \sum_{i=1}^{N} w_i \\
&= \sum_{i=1}^{N} w_i \log(S_{t+1}^{i}/F_t^{i}) + 0 = r_{t+1}^{USD},
\end{aligned} \tag{OA-4}$$

and therefore the log returns of symmetric, numeraire-independent trades are *identical* from any perspective; the differences between their percentage returns from different perspectives are of second order.

In sum, a symmetric carry trade, for any benchmark currency has similar returns for investors across the world. However, symmetry is not a sufficient condition for numeraire independence. It is important to emphasize this point, because a number of recent articles have considered ”currency-neutral” symmetric strategies, where no position is taken with respect the benchmark currency itself, or in other words, the weight assigned to the benchmark currency is always zero (this is implicitly true also for dollar carry). Let’s examine, following Daniel, Hodrick, and Lu (2017), a ”dollar-neutral” carry trade with weights $w_i = 1/(N-1)$ if the interest rate of currency i is in top half of the interest rates of the given set of currencies, and $w_i = -1/(N-1)$ otherwise (if N-1 is odd, the currency with the median interest rate is left out of the trade). This strategy is clearly symmetric. However, it is not numeraire independent because if we define it relative to another benchmark currency, say the yen, the weight function of this ”yen-neutral” trade will change, with now non-zero weights on the USD and zero weights on the JPY. Therefore, such ”currency-neutral” trades will produce different returns for different benchmark currencies, going beyond the differences induced by cross-currency risk.

We recognize that some numeraire-dependent strategies are of obvious interest, but care must be taken to define them in an international context. For example, the HML factor, introduced by Lustig, Roussanov, and Verdelhan (2011, 2014) is a carry trade which is symmetric, but not numeraire-independent as it goes long (short) an extreme portfolio based on an interest rate ranking (as the DB strategy does), but excludes the USD from any portfolio. This dollar neutrality makes the trade numeraire-dependent. Of course, when such a trade is defined for benchmark currencies with non-extreme interest rates, it should often yield similar returns across the different country perspectives.

Our preference for using symmetric, numeraire-independent carry trades is consistent with the best known investable indices, such as the Deutsche Bank (DB) Harvest Indexes. The DB strategy goes long (short) the G-10 currencies with the three highest (lowest) interest rates. Importantly, when the USD interest rate is among the top or bottom three, part of the trade automatically gets a zero return, because a position in the benchmark currency itself is taken, and hence the trade is *not* dollar-neutral. However, it is symmetric and numeraire-independent, which is an advantage for a global currency trading strategy, and may also be an advantage for a global risk pricing factor. In the trades that we consider, all participating currencies are given a non-zero weight, including the benchmark currency, which by design yields a zero return, whether it is held long or short.

Another way to see the fundamental difference between asymmetric, numeraire-dependent trades on the one hand, and numeraire-independent strategies on the other is to examine what would happen if, say, a yen-based investor would try to mimic, for example, dollar carry by taking exactly the same positions, but relative to the yen. That is, she will go long or short in all the currencies (including the yen) as dollar carry does, thus keeping the same weight function as in the original dollar trade, but for a different benchmark currency. This strategy would yield quite different returns as it would face *full* cross-currency risk, and not just profit and loss currency risk.

With the previous notation: $r_{t+1}^{JPY} = [r_{t+1}^{USD} + \sum_{i=1}^{N} w_i] * \overline{S}_{t+1}^{USD} / \overline{F}_t^{USD} - \sum_{i=1}^{N} w_i$. Since for dollar carry these weights add up to one, and not zero as in a symmetric trade, the yen-based return is now:

$$r_{t+1}^{JPY} = [r_{t+1}^{USD} + 1] * \overline{S}_{t+1}^{USD} / \overline{F}_t^{USD} - 1 = r_{t+1}^{USD} F_t^{JPY} / S_{t+1}^{JPY} + [F_t^{JPY} / S_{t+1}^{JPY} - 1], \quad \text{(OA-5)}$$

which, compared to the expression in (OA-3), adds a second return term that can well be of similar or even larger magnitude than the first term.

OA-II Tests for differences in Sharpe ratios and return skewness

Sharpe ratios

The statistical significance of the differences between the Sharpe ratio or skewness of the SC trade and those of trades from subsets is evaluated using bootstrap tests that follow Ledoit and Wolf (2008) or Annaert, Van Osselaer, and Verstraete (2009). Skewness difference can be tested in a ”direct” bootstrap that resamples from a distribution which respects the null hypothesis of no difference. In the case of Sharpe ratios, their difference does not easily admit such a distribution, hence the approach followed is ”indirect” and resamples from the observed data. A version of this approach to comparing Sharpe ratios has been applied recently, among others, in DeMiguel, Nogales, and Uppal (2014).

In implementing the test for a difference between Sharpe ratios, we depart in two minor ways from Ledoit and Wolf (2008). First, we only consider the i.i.d. case (their Section 3.2.1). We have verified that our carry trade return series have insignificant autocorrelations for lags up to 10. Furthermore, the suggested block size selection procedure (their Algorithm 3.1) results consistently in a selected block length of one, when using our data. Second, we consider one-sided bootstrap confidence intervals and p-values, since our null hypothesis is that carry trades obtained with the enhancement rule do not *improve* on the Sharpe ratio of the SC trade. We modify accordingly their equation (7).

Following the notation in Ledoit and Wolf (2008), let $\widehat{\mu}_S$ and $\widehat{\mu}_B$ denote the sample average returns of a carry trade from some subset of the G-10 currencies and the SC trade, respectively, while $\widehat{\gamma}_S$ and $\widehat{\gamma}_B$ are the sample second moments (uncentered) of the returns of these trades. Let also $\widehat{v} = (\widehat{\mu}_S, \widehat{\mu}_B, \widehat{\gamma}_S, \widehat{\gamma}_B)$, and assume that $\sqrt{T}(\widehat{v} - v) \xrightarrow{d} (0, \Psi)$, where v is the population counterpart, T is sample length and Ψ is some symmetric positive-definite matrix. The latter assumption holds under mild conditions. For the sample difference $\widehat{\Delta}$ between the Sharpe ratios of the carry trade from a subset of the G-10 currencies and the SC trade, and the deviation of this sample difference from the population value Δ, one can write

$$\widehat{\Delta} = f(\widehat{v}) = \frac{\widehat{\mu}_S}{\widehat{\gamma}_S - \widehat{\mu}_S^2} - \frac{\widehat{\mu}_B}{\widehat{\gamma}_B - \widehat{\mu}_B^2} \qquad \text{and} \qquad \sqrt{T}(\widehat{\Delta} - \Delta) \xrightarrow{d} (0, \nabla' f(v) \Psi \nabla f(v)), \tag{OA-6}$$

where $\nabla' f((a,b,c,d)) = \left(\frac{c}{(c-a^2)^{1.5}}, -\frac{d}{(d-b^2)^{1.5}}, -\frac{a}{2(c-a^2)^{1.5}}, \frac{b}{2(d-b^2)^{1.5}} \right)$ and (a, b, c, d) represent the elements in $\widehat{v}$. If $\widehat{\Psi}$ is a consistent estimator of Ψ, then the standard error of $\widehat{\Delta}$ is given by

$$s(\widehat{\Delta}) = \sqrt{\frac{\nabla' f(\widehat{v}) \widehat{\Psi} \nabla f(\widehat{v})}{T}}. \tag{OA-7}$$

To test the null hypothesis $\Delta \leq 0$, we bootstrap the returns of the two carry trades that are compared, and consider the studentized random variable $L = \frac{\Delta^* - \widehat{\Delta}}{s(\Delta^*)}$, where Δ^* is a difference in Sharpe ratios computed with bootstrapped returns, and $s(\Delta^*)$ is the corresponding standard error. Even though we bootstrap "under the alternative", this procedure generates meaningful sampling variation under the null of no difference between Sharpe ratios. Given the lack of autocorrelation in the carry trade return series, as noted above, we use an i.i.d. bootstrap (5000 samples, with replacement and pairwise, to preserve a possible cross-sectional correlation between the returns of the two carry trades). A p-value for the null is calculated as the proportion of bootstrapped series for which:

$$\widehat{\Delta} - L = \widehat{\Delta} + \frac{\widehat{\Delta} - \Delta^*}{s(\Delta^*)} s(\widehat{\Delta}) \leq 0, \tag{OA-8}$$

similar to equation (7) in Ledoit and Wolf (2008). These p-values are reported in Tables 2 and 3.

Skewness

To test for a difference in skewness, Annaert, Van Osselaer, and Verstraete (2009, page 277) first ”symmetrize” the compared return series, by appending to them the mirror images of the original observations in terms of distance to the average return. The skewness (as well as any odd central moment) of these modified returns is thus zero, and a bootstrap that resamples from them conforms to the null of no difference in skewness. Given that autocorrelation does not seem to be an issue in our series, we draw pairwise from the modified series of the compared returns, and compute the p-value as the percentage of draws that yield higher improvement on the benchmark skewness than that observed in the data. All bootstraps are performed with 5000 draws.

OA-III Differences in Sharpe ratios - accommodating the selection

The enhancement procedure described in Section 3 introduces a possible selection bias, which is not accounted for by the bootstrap-based test described above, following Ledoit and Wolf (2008). To address this issue, we suggest an alternative approach, and instead of bootstrapping the actual carry returns, we adopt the following randomization procedure:

- at the end of month t keep the interest rate differentials as in the data, but assign to each of them *at random* any of the ten returns for the following month $t+1$.
- to *each* of these ten returns for month $t+1$ add the same constant c_{t+1}. We call the returns obtained in this way ”randomized” returns.

- the constant c_{t+1} can be positive or negative, and is chosen so that a carry trade that uses all ten "randomized" returns would have exactly the same return as the actual SC trade for month $t+1$. Such a carry trade would choose the currencies to be long or short exactly as the SC trade, based on sorting the same interest rate differentials.

- do this for all months in the sample, and repeat 1000 times, to obtain 1000 sets of ten "randomized" return series, that correspond to the actual interest rate differentials. Given the large number permutations of ten numbers, we do not bootstrap in addition the interest rate differentials.

- note that the constants c_{t+1} are different for different months, and that each "randomized" return corresponding to a particular interest rate differential is potentially very different from the actual one. This approach may associate, for example, the JPY returns predominantly with the highest interest rates in some randomization trials. However, the returns for each month, and hence the Sharpe ratios of the carry trades with ten currencies (all 1000 with "randomized" returns and the actual SC trade) are exactly the same.

- on each of the 1000 sets of 10 time series reproduce the enhancement procedure described in Section 3.2. Based on the order of exclusion obtained from this procedure, identify for each of the 1000 sets the currencies that would enter "good" and "bad" carry trades.

- in the full sample period construct trades with the least excluded three, five or seven currencies, corresponding to our G1-G5 trades, and similarly with the most often excluded three, five or seven currencies, corresponding to our B1-B5 trades.

For each of 1000 sets of 10 series of randomized carry trade returns, Δ^* denotes the difference between the annualized Sharpe ratio of a good carry trade (from three, five or seven currencies), constructed from this set following the enhancement procedure, and the SC trade or the corresponding bad trade. As in

Appendix OA-II, $\widehat{\Delta}$ denotes the sample difference between the annualized Sharpe ratio of a good carry trade and the SC trade or the corresponding bad trade. We now show $\widehat{\Delta}$ for each good carry trade, the average of the 1000 Δ^*'s for trades from as many currencies as the good trade on the same line, and the proportion of such Δ^*'s exceeding $\widehat{\Delta}$.

	Good trades vs. SC			Good vs. bad trades		
	$\widehat{\Delta}$	avg. Δ^*	% $\Delta^* > \widehat{\Delta}$	$\widehat{\Delta}$	avg. Δ^*	% $\Delta^* > \widehat{\Delta}$
G1	0.20	0.14	0.21	0.42	0.42	0.50
G2	0.18	0.16	0.43	0.31	0.41	0.71
G3	0.30	0.16	0.09	0.57	0.41	0.19
G4	0.20	0.16	0.37	0.45	0.41	0.41
G5	0.39	0.14	0.02	0.56	0.32	0.05

There is substantial bias in the comparison between the G1-G5 carry trades with the SC trade, with the selection procedure adding 14% (for G1 and G5) or 16% (for G2 to G4) to the annualized Sharpe ratio. Yet, in every case the *observed* increases in the Sharpe ratio (denoted by $\widehat{\Delta}$) are even higher, and for two out of the five good trades the observed Sharpe ratio is in the 10% right tail of the distribution of the Sharpe ratios obtained under the selection procedure using the randomized (scrambled) currency returns. When comparing the G1-G5 carry trades to the corresponding B1-B5 trades, the bias is relatively more important, and in fact at least as large as the observed difference in Sharpe ratios for the G1 and G2 trades. Only the G5 versus B5 comparison yields a Sharpe ratio of a good trade in the right tail (5.3%) of the corresponding distribution under scrambled currency returns.

Of course, these observations alone do not constitute a proper test, since the randomization procedure also can change the variability of the returns, and proper testing requires the use of a pivotal test statistic, such as a t-statistic. To create a proper test statistic, we modify the procedure in Ledoit and Wolf

(2008) by bias-correcting our sample Sharpe ratios, and using t-statistics from the empirical distribution as in Appendix OA-II. The results, which also reproduce the relevant portion from Table 3, to facilitate comparison are as follows:

	Good trades vs. SC						Good vs. bad trades				
	av.ret	std.	SR	bstrp.	rand.		av.ret	std.	SR	bstrp.	rand.
SC	1.02	3.30	0.31								
G1	1.67	3.29	0.51	0.02	0.18	B1	0.68	7.50	0.09	[0.01]	[0.50]
G2	1.70	3.47	0.49	0.13	0.44	B2	0.98	5.54	0.18	[0.07]	[0.72]
G3	2.49	4.09	0.61	0.01	0.06	B3	0.21	4.91	0.04	[0.01]	[0.16]
G4	2.22	4.39	0.51	0.12	0.39	B4	0.28	4.96	0.06	[0.02]	[0.42]
G5	3.97	5.71	0.69	0.03	0.04	B5	0.61	4.66	0.13	[0.01]	[0.09]

Let's first focus on the G1 trade. The t-statistic for its Sharpe ratio (0.51) being different from the benchmark Sharpe ratio (0.31) has a p-value of 0.02. When we do the test using the randomized samples, correcting for selection bias, the p-value increases to 0.18, and the difference is no longer statistically significant. The p-values invariably increase for all carry trades, but remain significant at the 5% level for G5, and at the 10% level for G3. For the good vs. bad carry trade comparison, the p-values increase dramatically and only the G5 trade has a significantly higher Sharpe ratio than B5 (at the 10% level).

OA-IV Factor models explaining good and bad carry trades

Tables OA-2 to OA-4 present the results separately for the standard carry trade (SC), the G1-G5 and B1-B5 trades, and the GC and BC trades on average. The first column in Table OA-2 also shows the respective average returns that are to be explained. For the G1-G5 trades these range between 1.7 and 4% (annualized), and are all significantly different from zero at the 1% confidence level (with GMM standard errors); for the GC trades they are on average 2%, and all but two out of 18 are significant at the 5% level. In contrast, the average returns for the bad carry trades never exceed 1%, and are never significant, even at the 10% level.

A. Model with equity volatility

The market factor (denoted MKT) in the model is proxied by the total return of the MSCI-World equity index, in excess of the risk-free rate and expressed in USD. The equity volatility factor (EqVol) reflects innovations in global equity volatilities, as constructed in Lustig, Roussanov, and Verdelhan (2011), and is taken from Verdelhan's website (data until 12/2013). The interaction term (product of MKT and EqVol) is denoted "prod", and exhibits highly negative skewness (-7.6).

The top panel of Table OA-2 reports results from time-series regressions of carry trade returns on the three risk factors. The market betas are significant for both good and bad trades, and of comparable magnitudes. However, the slope coefficient estimates on the product factor are typically negative, albeit rarely significant for good trades, while they are positive, mostly much larger in magnitude, and almost always significant at the 5% significance level for the bad trades. The F-test for no difference between the average slope coefficients across the GC and BC trades rejects only for β_{prod}. Given the high negative skewness of the product factor, the large positive value of β_{prod} implies that the market risk exposure of the bad trades increases substantially in highly volatile times, helping to explain the negative skewness of the bad trades as shown in Table 3.

From the perspective of a time-varying market beta, the large β_{prod} implies, for example, that the effective market beta for bad carry trades ranges between 0.025 and 0.083 for the 10-th and 90-th percentile observations of EqVol (which are -0.67 and 0.59, respectively). This regime dependence is much weaker for good carry trades, due to their smaller β_{prod} estimates. The SC trade resembles the bad trades in this respect, with a β_{prod} that is positive and marginally significant (at the 10% level). Given that increases in volatility tend to characterize periods of market downturns (the correlation between MKT and EqVol is -0.24 in our sample), our findings attribute the under-performance in times of crisis mostly to bad carry trades, while good trades are less affected.

The alpha's obtained in the time-series regressions are difficult to interpret in the presence of non-traded factors. Therefore, we also perform GMM-based cross-sectional tests on the GC and BC return cross sections, and show the results in the last two rows of the table. For the GC trades, the risk price for the MKT factor is significant at the 5% level, while for the BC trades no risk price is significant. However, the joint test does not reject for either of the two cross sections, delivering large p-values.

[Figure 4 about here.]

For further clarification, Panels A and B in Figure 4 plot model-predicted vs. actual average returns for the GC and BC trades, where we see practically no relation for the BC trades, but a much better fit for the GC trades, albeit with a few outliers. When we run a simple OLS regression of actual average returns on a constant and the model-based expected returns, we obtain an R^2 of 0.67 for the GC trades, and 0.29 for the BC trades. The combined evidence suggests that this three-factor model does not adequately describe the returns of the bad carry trades, but still saliently reveals the high exposure of these trades to the equity market during high-volatility periods. In contrast, a significant price of risk for the market factor and Figure 4 show the promise of the model to provide a risk-based interpretation of good carry trades.

B. Model with Up and Down equity market factors

Our interest in such a model is motivated both by the asymmetric patterns in carry trade returns documented above, and the recent work of Lettau, Maggiori, and Weber (2015), who find support for a similar model pricing the joint cross section of several asset classes, including the returns of interest-rate-sorted currency portfolios. Note that their model employs the market factor itself, together with a separate down-market factor, whereas we use uncorrelated down- and up-market factors, which help sharpen the focus on the asymmetric return behavior across good and bad carry trades (see also Ang, Chen, and Xing (2006, Table 2)). Keeping the notation MKT for the total return of the MSCI-World equity index, in excess of the risk-free rate and expressed in USD, the Down factor is taken to be $min(MKT, 0)$, and the Up factor is $max(MKT, 0)$.

Table OA-3 shows that in the time-series regressions the slope coefficient estimates on the Down factor are not statistically significant for about 70% of the good carry trades, but are significant for all but one of the bad carry trades. The pattern is *reversed* for the Up factor, where the estimates are significant for most of the good trades, but are in fact never significant for the bad trades, even at the 10% confidence level. The magnitudes of the respective slope coefficients for good versus bad trades also differ largely, by a factor of three or four, and these differences are highly significant, as evidenced by the reported p-values from GMM tests for the equality of the average β_{Down} or β_{Up} across the 18 GC and BC trades. Additional joint tests for pairwise equality between the corresponding coefficients for the GC and BC trades reject with even smaller p-values. As above, the SC trade exhibits mixed features, with both slope coefficients being significant.

The cross-sectional test results resemble those from Table OA-2, in that both risk prices λ_{Down} and λ_{Up} are statistically significant for good trades, and highly insignificant for bad trades, while the tests for the pricing errors being jointly equal to zero fail to reject, with high p-values. Moreover, the plots of model-based versus actual average returns, similar to those in Figure 4) again reveal a reasonable fit for good trades, but no apparent relation for bad trades, indicating that the model with down- and up-market factors more adequately describes the returns of good carry trades. The important additional insight from this model, however, is the striking dichotomy between the returns of good carry trades, which have relatively high Up-market betas but decouple in bad times, and the returns of bad carry trades, which have relatively high Down-market betas.

C. Fama-French three-factor model

Similar to Table OA-2 and in the same format, Table OA-4 illustrates the ability of the Fama-French three-factor model to explain the returns of good and bad carry trades, and the main finding is that the model does not perform well with respect to good carry trades.

The top panel of this table refers to time-series tests, and shows that the betas on the market factor are economically small for these trades (0.05 on average), albeit often significant, while those on the other two factors typically are not significantly different from zero. The adjusted R^2's in the time-series regressions are relatively low, even sometimes negative, whereas the alphas are only about 5 to 30% lower than the unconditional average carry trade returns, and still statistically significant for all G1-G5 trades and 14 of the GC trades. On the other hand, for bad carry trades the betas on all three factors are higher and statistically significant in most cases, and the R^2's are on average 0.12. A test for no difference between the average slope coefficients across the 18 GC and BC trades rejects for β_{MKT} and β_{SMB}, at the 5% confidence level. Interestingly, the model renders all alphas much lower than the respective average returns for the bad carry trades, so that these trades can be qualified as "negative alpha trades", from the perspective of this model. The model also explains a large part of the SC trade's average returns, with statistically significant factor loadings and a high R^2. The time-series tests therefore suggest that the good carry trades pose a problem for this model, whereas the SC trade and the bad trades at least are meaningfully exposed to standard risk factors. In addition, a test for alphas being jointly equal to zero does not reject for both the GC and BC sets of carry trades, with p-values above 0.30.

The last two lines of the table show results from GMM-based cross-sectional tests, using the GC and BC trades as test assets. The estimates of the risk prices λ are all statistically insignificant, except for λ_{MKT} for the GC trades, while the tests for the pricing errors being jointly equal to zero exhibit p-values above 0.70. The results for the risk prices thus cast doubt on the explanatory power of the Fama-French three-factor model for the BC trades as well, whereas the joint test results may reflect power issues.

OA-V Option-hedged good and bad carry trades

Here we re-examine the prior evidence for the crash-risk interpretation of carry returns from currency options bringing in the good-bad carry trade angle. Options have been used to explore the peso problem hypothesis: carry trade returns reflect compensation for small-probability adverse events that are not observed in the researcher's sample. Burnside, Eichenbaum, Kleshchelski, and Rebelo (2011), for example, consider a hedged carry trade that adds currency options and, by construction, is immune to large losses that are potentially associated with a peso event. They find that the hedged carry trade does have lower returns than the unhedged trade, partially due to the cost of the options used for hedging, but the hedged and unhedged carry trades have similar Sharpe ratios. They conclude that the returns in the peso state are not characterized by large negative payoffs, but rather by a large value of the stochastic discount factor (large price of risk) in that peso state.

We calculate returns of hedged trades following Burnside, Eichenbaum, Kleshchelski, and Rebelo (2011), in that we use forward differentials and returns net of transaction costs. However, as in Jurek (2014) we employ option data for all pairs of G-10 currencies, and not only pairs involving the USD, which yields more efficient hedging (see also Daniel, Hodrick, and Lu (2017, Section 7)). We consider hedging with either at-the-money put options or out-of-the-money puts with a delta of -0.25, which are denoted by industry convention as 25 delta puts.

Below we provide details on the construction of the hedged carry trades and the results, presented in Table OA-5, can be summarized as follows: Hedged good trades invariably exhibit lower annualized Sharpe ratios than the corresponding unhedged ones, on average by 15 (12) percentage points when hedged by at-the-money (out-of-the-money) options. In contrast, the respective reduction in Sharpe ratios is only six (five) percentage points for hedged bad carry trades, with Sharpe ratios even increasing in some cases. Moreover, the differences in Sharpe ratios between hedged and unhedged trades are statistically significant in about two thirds of the cases for good carry trades, but only in a few cases for bad carry trades. Thus, the risk-adjusted returns of good trades are affected much more by the costs of hedging, which may indicate that the peso argument may be more applicable to good carry trades. However, hedging results in very attractive risk-return profiles for the G4 and G5 trades, which use only a few currencies, with high Sharpe ratios and significant positive skewness, which counters the peso problem argument.

Construction of hedged carry trade returns

In constructing option-hedged carry trades we borrow elements from two previous approaches. First, as in Burnside, Eichenbaum, Kleshchelski, and Rebelo (2011) we add the hedging component, obtained separately from options to the already calculated unhedged carry trade returns. We do not rely on put-call parity, which is done for example in Daniel, Hodrick, and Lu (2017). Unlike Jurek (2014), we account for transaction costs in both components of the hedged trade, which facilitates comparison with the corresponding unhedged one.

Second, as in Jurek (2014) we use only put options for hedging, and more specifically puts purchased on individual pairs of currencies that enter the trade. These pairs do not need to involve the USD and hedge separately the long and short legs of the trade, thus aiming at a more efficient hedging scheme (see also Daniel, Hodrick, and Lu (2017, Section 7)). We hedge with at-the-money and 25 delta puts.

Another design choice concerns the currency pairs for which puts are bought. The symmetric trades we consider have an equal number of long and short positions, with equal weights, and we first take the pair of the currency with highest forward differential and that with the lowest one, then the pair of the currency with the second highest and that with the second lowest differential, etc. We have also verified that results change little if pairs are assigned via random permutations.

Put option prices are not directly given in the data (JPMorgan's DataQuery), so we invert the provided Garman-Kohlhagen implied volatilities, using our forward prices (mid-quotes) and LIBOR's (also provided). Deltas are converted into strike prices via equations (8a) and (8b) in Jurek (2014). When needed, we also use symmetry relations as in equations (11) and (12) in Jurek (2014), and transform out-of-the-money call prices from provided volatilities into required out-of-the-money put prices.

Regarding transaction costs, the unhedged part of the trade is unchanged. For the hedging part we follow Burnside, Eichenbaum, Kleshchelski, and Rebelo (2011, page 23) and assume a bid-ask spread of 5%, so we add half of this to the price of a put option obtained using mid-quotes.

Table OA-1

Currency volatility vs. good carry trades as currency market pricing factors

This table differs from Table 4 only in one aspect: instead of the HML^{FX} factor, we now use the currency volatility factor as in Menkhoff, Sarno, Schmeling, and Schrimpf (2012) and denote it ”FXVol”.

	avg. ret.	p-val	α	p-val	β_{RX}	p-val	β_{FXVol}	p-val	β_{Good}	p-val	R^2
	2.33	(3/4)	0.04	(5/6)	1.11	(11/11)	0.0000	(6/6)			77.6
G1			0.06	(0/1)	1.11	(11/11)	0.0001	(6/6)	-0.01	(9/9)	80.8
G2			0.06	(1/2)	1.11	(11/11)	0.0004	(5/5)	-0.01	(9/9)	78.9
G3			0.06	(0/1)	1.12	(11/11)	0.0003	(6/7)	-0.01	(8/8)	79.7
G4			0.01	(1/3)	1.11	(11/11)	0.001	(6/6)	0.02	(8/10)	79.4
G5			-0.01	(1/4)	1.11	(11/11)	0.001	(6/6)	0.02	(7/7)	78.3

	λ_{RX}	p-val	λ_{FXVol}	p-val	λ_{Good}	p-val	b_{RX}	p-val	b_{FXVol}	p-val	b_{Good}	p-val	$\chi^2_{pr.err.}$
	2.14	0.10	-0.75	0.06			3.0	0.33	-3.0	0.15			0.00
G1	2.13	0.10	-1.74	0.65	2.43	0.00	4.0	0.17	1.7	0.37	23.6	0.01	0.30
G2	2.14	0.10	-0.04	0.99	4.36	0.00	3.8	0.22	3.1	0.18	37.7	0.00	0.54
G3	2.12	0.10	-1.87	0.63	3.61	0.00	1.6	0.60	2.6	0.21	24.1	0.01	0.08
G4	2.03	0.12	-1.27	0.74	4.34	0.00	-4.5	0.31	0.1	0.94	25.5	0.01	0.00
G5	2.01	0.12	-3.45	0.40	7.44	0.00	-7.4	0.12	-0.9	0.62	26.1	0.00	0.50

Table OA-2

Good and bad carry trades and a three-factor model with market and equity volatility factors

This table reports test results for a three-factor model with a market factor (denoted MKT), an equity volatility factor (EqVol) and the product of MKT and EqVol. MKT is the return of the MSCI-World equity index (total returns in excess of the risk-free rate, in USD). EqVol is the equity volatility factor, used in Lustig, Roussanov, and Verdelhan (2011) and available at Verdelhan's website until 12/2013, and the product of MKT and EqVol is denoted "prod". The test assets are the SC and good and bad carry trades. For the GC and BC trades the "avg. ret." column shows the average of their average returns, and the remaining columns show similarly averages of the regression coefficients and adjusted R^2's (in percent). Standard errors are estimated with GMM and account for one Newey-West lag. The first (second) number in parentheses shows how many of the 18 corresponding individual estimates are significant at the 5% (10%) confidence level. For example, none of the average returns for the 18 BC trades is significant even at the 10% level. In curly brackets are shown p-values for a test of equality of the respective average slope coefficients across the GC vs. BC trades, accounting for heteroskedasticity and one Newey-West lag. The last two lines show, similar to Table 4, results from cross-sectional GMM estimations of the model on the GC and BC carry trades. The sample period is 12/1984 to 12/2013. The reported average returns, α's and λ's are annualized and in percent.

	avg. ret.	p-val	α	p-val	β_{MKT}	p-vals	β_{EqVol}	p-val	β_{prod}	p-val	R^2
SC	1.02	0.10	0.81	0.19	0.06	0.00	-0.001	0.35	0.02	0.01	10.2
G1	1.67	0.01	1.46	0.02	0.05	0.00	0.000	0.51	0.01	0.28	5.6
G2	1.70	0.01	1.30	0.05	0.05	0.00	0.000	0.84	-0.01	0.18	3.2
G3	2.49	0.00	1.79	0.02	0.09	0.00	-0.001	0.33	-0.01	0.49	10.4
G4	2.22	0.01	1.47	0.10	0.07	0.00	-0.001	0.37	-0.03	0.15	4.8
G5	3.97	0.00	2.63	0.02	0.11	0.00	0.000	0.99	-0.07	0.00	8.4
B1	0.68	0.65	0.51	0.72	0.07	0.06	0.000	0.94	0.04	0.18	3.8
B2	0.98	0.39	0.88	0.40	0.06	0.01	-0.001	0.29	0.04	0.00	7.2
B3	0.21	0.83	0.40	0.66	0.04	0.07	0.000	0.82	0.05	0.00	8.1
B4	0.28	0.78	0.25	0.78	0.07	0.00	-0.001	0.46	0.05	0.01	12.2
B5	0.61	0.52	0.71	0.40	0.05	0.01	-0.001	0.49	0.05	0.00	12.4
GC	1.96	(16/17)	1.37	(6/10)	0.07	(18/18)	-0.001	(0/1)	-0.02	(3/4)	5.8
BC	0.66	(0/0)	0.69	(0/0)	0.06	(14/17)	-0.001	(0/0)	0.05	(17/17)	9.3
GC vs. BC					{0.49}		{0.42}		{0.000}		
					λ_{MKT}	p-val	λ_{EqVol}	p-val	λ_{prod}	p-val	$\chi^2_{pr.err.}$
GC					31.1	0.04	22.0	0.97	7.93	0.74	0.97
BC					6.54	0.57	-289.0	0.48	1.17	0.93	0.76

Table OA-3

Good and bad carry trades and a model with Down- and Up-market factors

This table differs from Table OA-2 only in one aspect: instead of the global market and equity volatility factors we use a Down and Up equity market factors. If MKT denotes the total return of the MSCI-World equity index, in excess of the risk-free rate, the Down factor is taken to be $min(MKT, 0)$, and the Up factor is taken to be $max(MKT, 0)$. The average returns and their p-values are omitted.

	α	p-val	β_{Down}	p-val	β_{Up}	p-val	R^2
SC	1.21	0.22	0.08	0.00	0.05	0.04	9.4
G1	1.68	0.08	0.06	0.00	0.04	0.06	5.5
G2	0.42	0.67	0.02	0.21	0.07	0.00	3.5
G3	1.49	0.20	0.08	0.00	0.10	0.00	10.4
G4	0.56	0.66	0.03	0.28	0.09	0.00	4.2
G5	-1.91	0.25	-0.05	0.20	0.21	0.00	7.9
B1	1.37	0.58	0.13	0.11	0.06	0.36	3.4
B2	2.47	0.15	0.13	0.00	0.03	0.45	5.8
B3	1.73	0.28	0.12	0.03	0.02	0.64	5.3
B4	1.83	0.26	0.15	0.01	0.04	0.32	9.8
B5	2.76	0.06	0.15	0.00	0.01	0.69	9.3
GC	0.49	(0/1)	0.04	(5/6)	0.09	(15/16)	5.9
BC	2.49	(6/8)	0.14	(17/18)	0.02	(0/0)	7.2
GC vs. BC			{0.02}		{0.02}		

	λ_{Down}	p-val	λ_{Up}	p-val	$\chi^2_{pr.err.}$
GC	14.0	0.05	15.5	0.02	0.96
BC	3.83	0.50	5.35	0.56	0.74

Table OA-4

Good and bad carry trades and the three-factor Fama-French model

This table differs from Table OA-2 only in one aspect: instead of the global market and equity volatility factors we use the three-factor Fama-French factors. The average returns and their p-values are omitted.

	α	p-val	β_{MKT}	p-vals	β_{SMB}	p-val	β_{HML}	p-val	R^2
SC	0.36	0.55	0.07	0.00	0.03	0.10	0.03	0.09	10.4
G1	1.24	0.04	0.05	0.00	0.01	0.58	0.01	0.61	4.5
G2	1.44	0.03	0.03	0.01	0.00	0.86	-0.01	0.81	1.6
G3	1.76	0.02	0.07	0.00	0.01	0.79	0.03	0.16	6.4
G4	1.95	0.02	0.02	0.24	0.00	0.94	0.02	0.49	-0.2
G5	3.79	0.00	0.01	0.59	-0.03	0.36	0.00	0.93	-0.4
B1	-0.97	0.51	0.14	0.00	0.08	0.04	0.13	0.00	9.5
B2	-0.21	0.84	0.11	0.00	0.06	0.02	0.08	0.01	10.9
B3	-0.69	0.48	0.09	0.00	0.07	0.00	0.05	0.08	10.6
B4	-0.93	0.34	0.13	0.00	0.05	0.05	0.05	0.09	15.8
B5	-0.43	0.61	0.11	0.00	0.06	0.01	0.05	0.04	14.6
GC	1.45	(11/14)	0.05	(14/15)	0.00	(0/0)	0.03	(3/4)	2.9
BC	-0.40	(0/0)	0.11	(18/18)	0.06	(13/16)	0.05	(10/14)	12.0
GC vs. BC			{0.02}		{0.02}		{0.25}		
			λ_{MKT}	p-val	λ_{SMB}	p-val	λ_{HML}	p-val	$\chi^2_{pr.err.}$
GC			37.2	0.01	-15.5	0.62	2.81	0.89	0.98
BC			0.43	0.96	8.49	0.55	1.43	0.89	0.73

Table OA-5

Option-hedged carry trade returns

The table shows annualized averages (denoted ”avg. ret.”, in percent) and Sharpe ratios (”SR”), as well as skewness (”skew”), for the monthly returns of unhedged and hedged versions of the SC trade, the G1-G5 and B1-B5 carry trades, and the GC and BC trades as described in Table 3. The columns denoted ”p-val” show bootstrapped two-sided p-values for the null hypothesis that the respective characteristic of a hedged carry trade does not differ from that of the corresponding unhedged trade. For the GC and BC trades the table shows averages of the respective estimates. The first (second) number in parentheses shows how many of the 18 respective individual characteristics are significantly different at the 5% (10%) confidence level from those of the corresponding unhedged trades. Returns of hedged trades are calculated by adding to the unhedged returns a hedging component, obtained separately from at-the-money (ATM) put options and 25 delta puts, and also accounting for transaction costs. Option data for all pairs of G-10 currencies are employed, as in Jurek (2014), and not only for pairs that involve the USD, so we avoid hedging separately the long and short legs of the trade. Appendix B provides details of the construction of hedged carry trades. The sample period is 1/1998 to 6/2014.

	unhedged			hedged, ATM puts						hedged, 25 delta puts					
	avg. ret.	SR	skew	avg. ret.	p-val	SR	p-val	skew	p-val	avg. ret.	p-val	SR	p-val	skew	p-val
SC	1.10	0.34	-0.56	0.43	0.00	0.17	0.02	-0.01	0.00	0.63	0.00	0.22	0.02	-0.14	0.00
G1	1.18	0.39	-0.11	0.43	0.00	0.17	0.02	0.21	0.00	0.60	0.00	0.21	0.02	0.12	0.00
G2	1.09	0.34	-0.19	0.39	0.00	0.16	0.02	0.22	0.00	0.57	0.00	0.20	0.02	0.04	0.00
G3	2.65	0.69	0.24	1.57	0.00	0.51	0.02	0.76	0.00	1.95	0.00	0.55	0.02	0.48	0.00
G4	1.61	0.39	0.37	1.13	0.02	0.35	0.38	1.04	0.00	1.20	0.00	0.32	0.12	0.71	0.00
G5	3.37	0.63	-0.02	2.28	0.00	0.54	0.12	0.75	0.00	2.78	0.00	0.57	0.22	0.39	0.00
B1	2.37	0.33	-0.77	2.07	0.45	0.39	0.36	0.11	0.00	2.01	0.17	0.32	0.82	-0.16	0.01
B2	2.25	0.37	-0.78	1.42	0.00	0.28	0.14	-0.34	0.00	1.73	0.00	0.31	0.16	-0.44	0.04
B3	0.58	0.11	-1.46	0.15	0.11	0.04	0.20	-0.82	0.02	0.20	0.02	0.04	0.12	-0.86	0.02
B4	1.18	0.24	-1.12	0.86	0.27	0.23	0.86	-0.26	0.02	1.04	0.45	0.25	0.95	-0.43	0.05
B5	0.97	0.22	-1.18	0.34	0.01	0.10	0.08	-0.53	0.00	0.53	0.00	0.14	0.08	-0.62	0.01
GC	1.80	0.45	-0.03	0.99	(18/18)	0.30	(12/12)	0.47	(16/17)	1.25	(18/18)	0.34	(11/13)	0.24	(15/16)
BC	1.23	0.23	-0.84	0.66	(11/12)	0.16	(3/6)	-0.18	(16/16)	0.85	(12/14)	0.18	(2/2)	-0.35	(13/16)

Table OA-6

Economic regressions

The table shows results from multivariate regressions of various carry trade returns on three macro variables: global equity volatility, as in LRV (2011), industrial production growth in the OECD countries, and the residual from regressing the US industrial production growth on that of the OECD. The last three columns show the change in the dependent variable for one standard deviation change in each regressor, all else equal. Intercepts and sensitivities are reported annualized and in percent, and adjusted R^2's are in percent.

	interc.	p-val	β_{EQV}	p-val	β_{IP}	p-val	β_{USres}	p-val	R^2	sensitivity to:		
										EQV	IP	USres
SC	0.71	0.30	-2.23	0.03	0.17	0.08	-0.09	0.51	2.56	-1.72	1.22	-0.55
DC	3.70	0.01	-0.86	0.64	0.27	0.21	0.00	0.99	-0.21	-0.67	1.96	-0.02
G1	1.33	0.05	-1.67	0.08	0.18	0.06	-0.09	0.48	1.84	-1.29	1.34	-0.56
G2	1.28	0.06	-0.50	0.50	0.26	0.00	-0.09	0.47	1.77	-0.39	1.86	-0.54
G3	1.91	0.02	-2.40	0.02	0.32	0.00	-0.25	0.11	4.25	-1.85	2.29	-1.46
G4	1.77	0.05	-1.78	0.16	0.24	0.08	-0.11	0.55	1.26	-1.38	1.71	-0.65
G5	3.69	0.00	0.38	0.78	0.22	0.23	-0.24	0.29	0.30	0.29	1.57	-1.41
B1	-0.05	0.97	-2.85	0.13	0.38	0.13	-0.01	0.98	0.79	-2.20	2.72	-0.04
B2	0.15	0.89	-3.82	0.01	0.40	0.02	-0.14	0.47	3.26	-2.95	2.91	-0.81
B3	-0.21	0.84	-2.69	0.09	0.22	0.17	-0.07	0.74	1.35	-2.07	1.56	-0.42
B4	-0.23	0.83	-3.64	0.03	0.27	0.12	-0.14	0.46	3.07	-2.81	1.92	-0.86
B5	0.04	0.97	-3.45	0.03	0.28	0.04	-0.11	0.52	3.29	-2.66	2.02	-0.68

Table OA-7: **Detailed version of the top panel of Table 4**

Good	port.	avg. ret.	p-val	α	p-val	β_{RX}	p-val	$\beta_{HML^{FX}}$	p-val	β_{Good}	p-val	R^2
	1	-1.63	0.25	-1.74	0.00	1.02	0.00	-0.39	0.00			90.3
	2	-0.19	0.88	-1.15	0.08	0.88	0.00	-0.13	0.00			75.8
	3	0.79	0.54	-0.28	0.65	0.95	0.00	-0.13	0.00			78.4
	4	2.80	0.05	1.12	0.10	1.01	0.00	0.00	0.94			78.4
	5	3.68	0.02	1.62	0.03	1.11	0.00	0.05	0.11			80.0
	6	4.65	0.01	0.43	0.35	1.03	0.00	0.61	0.00			93.8
	7	-0.18	0.92	-0.38	0.67	1.24	0.00	-0.46	0.00			80.5
	8	1.04	0.57	-0.14	0.87	1.27	0.00	-0.23	0.00			77.6
	9	2.76	0.12	1.16	0.14	1.28	0.00	-0.14	0.00			81.7
	10	2.77	0.14	0.39	0.66	1.28	0.00	0.05	0.19			79.7
	11	4.79	0.02	1.55	0.13	1.27	0.00	0.27	0.00			75.2
G1	1			-2.27	0.00	0.99	0.00			-0.61	0.00	76.5
	2			-0.85	0.15	0.89	0.00			-0.50	0.00	78.8
	3			-0.29	0.63	0.94	0.00			-0.30	0.00	77.7
	4			0.87	0.18	1.00	0.00			0.16	0.03	78.9
	5			1.15	0.10	1.10	0.00			0.41	0.00	82.3
	6			1.40	0.16	1.08	0.00			0.85	0.00	71.0
	7			-0.97	0.38	1.21	0.00			-0.75	0.00	69.4
	8			0.34	0.66	1.28	0.00			-0.87	0.00	81.7
	9			0.71	0.38	1.26	0.00			-0.05	0.62	80.1
	10			-0.06	0.94	1.27	0.00			0.41	0.00	81.4
	11			0.78	0.36	1.25	0.00			1.14	0.00	82.4
G2	1			-2.34	0.00	0.99	0.00			-0.56	0.00	76.0
	2			-1.10	0.08	0.88	0.00			-0.34	0.00	76.0
	3			-0.57	0.37	0.93	0.00			-0.13	0.08	76.2
	4			0.72	0.25	0.99	0.00			0.24	0.00	79.6
	5			1.34	0.06	1.10	0.00			0.28	0.00	81.1
	6			1.94	0.07	1.10	0.00			0.50	0.00	65.9
	7			-1.09	0.33	1.20	0.00			-0.66	0.00	68.6
	8			-0.31	0.73	1.25	0.00			-0.45	0.00	75.6
	9			0.59	0.47	1.26	0.00			0.03	0.74	80.1
	10			-0.09	0.92	1.27	0.00			0.42	0.00	81.7
	11			1.67	0.12	1.28	0.00			0.56	0.00	73.5
G3	1			-2.14	0.00	1.04	0.00			-0.50	0.00	76.2
	2			-0.88	0.16	0.92	0.00			-0.34	0.00	76.9
	3			-0.35	0.57	0.96	0.00			-0.19	0.00	76.9
	4			0.97	0.14	0.99	0.00			0.06	0.30	78.5
	5			1.21	0.09	1.07	0.00			0.27	0.00	81.4
	6			1.19	0.23	1.01	0.00			0.70	0.00	70.8
	7			-0.80	0.47	1.27	0.00			-0.61	0.00	69.2
	8			0.11	0.89	1.31	0.00			-0.52	0.00	77.5
	9			0.87	0.28	1.28	0.00			-0.11	0.16	80.3
	10			-0.02	0.98	1.24	0.00			0.28	0.00	80.7
	11			0.59	0.52	1.16	0.00			0.90	0.00	81.2
G4	1			-2.80	0.00	1.05	0.00			-0.26	0.00	71.4
	2			-1.10	0.09	0.98	0.00			-0.33	0.00	76.3
	3			-0.48	0.44	0.99	0.00			-0.18	0.00	76.7
	4			0.91	0.17	0.96	0.00			0.13	0.04	78.8
	5			1.35	0.05	1.03	0.00			0.27	0.00	81.2
	6			2.12	0.06	1.00	0.00			0.38	0.00	64.8
	7			-1.83	0.10	1.23	0.00			-0.19	0.19	63.8
	8			-0.23	0.80	1.39	0.00			-0.49	0.00	76.6

	9	0.72	0.37	1.28	0.00			-0.05	0.55	80.1
	10	0.31	0.72	1.23	0.00			0.17	0.08	79.9
	11	1.26	0.19	1.04	0.00			0.80	0.00	78.0
G5	1	-2.68	0.00	1.03	0.00			-0.17	0.00	71.0
	2	-0.93	0.16	0.95	0.00			-0.22	0.00	75.5
	3	-0.83	0.19	0.92	0.00			0.02	0.66	75.8
	4	0.57	0.37	0.94	0.00			0.16	0.00	79.6
	5	1.33	0.06	1.06	0.00			0.14	0.00	80.5
	6	2.53	0.03	1.11	0.00			0.06	0.48	62.8
	7	-1.96	0.07	1.19	0.00			-0.05	0.64	63.3
	8	-0.41	0.65	1.30	0.00			-0.19	0.00	74.0
	9	0.40	0.63	1.23	0.00			0.07	0.22	80.2
	10	-0.03	0.97	1.21	0.00			0.19	0.02	80.4
	11	1.99	0.07	1.24	0.00			0.18	0.05	70.8
G1	1	-1.69	0.00	1.02	0.00	-0.38	0.00	-0.06	0.39	90.3
	2	-0.79	0.19	0.89	0.00	-0.04	0.15	-0.44	0.00	79.0
	3	-0.15	0.81	0.95	0.00	-0.10	0.00	-0.16	0.06	78.7
	4	0.94	0.15	1.00	0.00	-0.05	0.18	0.22	0.01	79.0
	5	1.23	0.08	1.10	0.00	-0.05	0.11	0.48	0.00	82.5
	6	0.47	0.31	1.03	0.00	0.61	0.00	-0.04	0.49	93.8
	7	-0.30	0.74	1.25	0.00	-0.44	0.00	-0.11	0.34	80.6
	8	0.47	0.55	1.28	0.00	-0.09	0.04	-0.75	0.00	82.1
	9	0.99	0.21	1.28	0.00	-0.18	0.00	0.21	0.05	82.1
	10	0.00	1.00	1.27	0.00	-0.04	0.33	0.47	0.00	81.4
	11	0.69	0.42	1.24	0.00	0.06	0.16	1.05	0.00	82.6
G2	1	-1.57	0.00	1.03	0.00	-0.36	0.00	-0.17	0.00	90.8
	2	-0.91	0.15	0.89	0.00	-0.09	0.00	-0.25	0.00	77.0
	3	-0.29	0.64	0.95	0.00	-0.13	0.00	0.01	0.88	78.3
	4	0.83	0.20	1.00	0.00	-0.05	0.16	0.29	0.00	79.8
	5	1.34	0.07	1.10	0.00	0.00	0.98	0.28	0.00	81.1
	6	0.60	0.19	1.03	0.00	0.63	0.00	-0.17	0.00	94.1
	7	-0.18	0.84	1.25	0.00	-0.43	0.00	-0.21	0.02	80.9
	8	0.10	0.90	1.27	0.00	-0.19	0.00	-0.25	0.02	78.2
	9	0.95	0.22	1.28	0.00	-0.17	0.00	0.21	0.03	82.2
	10	-0.05	0.96	1.27	0.00	-0.02	0.65	0.44	0.00	81.6
	11	1.21	0.23	1.26	0.00	0.21	0.00	0.34	0.00	76.1
G3	1	-1.69	0.00	1.03	0.00	-0.38	0.00	-0.04	0.41	90.3
	2	-0.80	0.20	0.92	0.00	-0.07	0.02	-0.27	0.00	77.4
	3	-0.21	0.73	0.96	0.00	-0.12	0.00	-0.05	0.39	78.3
	4	1.00	0.14	0.99	0.00	-0.02	0.51	0.09	0.22	78.5
	5	1.23	0.09	1.07	0.00	-0.02	0.42	0.30	0.00	81.3
	6	0.47	0.31	1.03	0.00	0.61	0.00	-0.03	0.52	93.8
	7	-0.27	0.76	1.26	0.00	-0.44	0.00	-0.08	0.44	80.6
	8	0.30	0.72	1.31	0.00	-0.15	0.00	-0.33	0.00	78.9
	9	1.05	0.18	1.27	0.00	-0.16	0.00	0.08	0.34	81.8
	10	0.00	1.00	1.24	0.00	-0.02	0.69	0.30	0.00	80.7
	11	0.49	0.59	1.16	0.00	0.08	0.09	0.80	0.00	81.5
G4	1	-1.72	0.00	1.03	0.00	-0.39	0.00	-0.02	0.75	90.3
	2	-0.83	0.18	0.97	0.00	-0.10	0.00	-0.27	0.00	77.8
	3	-0.15	0.80	0.98	0.00	-0.12	0.00	-0.10	0.06	78.6
	4	0.95	0.15	0.96	0.00	-0.02	0.57	0.14	0.04	78.8
	5	1.31	0.07	1.03	0.00	0.02	0.54	0.26	0.00	81.2
	6	0.44	0.34	1.03	0.00	0.61	0.00	-0.01	0.90	93.8
	7	-0.52	0.55	1.21	0.00	-0.47	0.00	0.11	0.38	80.6
	8	0.30	0.72	1.38	0.00	-0.19	0.00	-0.37	0.00	79.4
	9	1.11	0.15	1.27	0.00	-0.14	0.00	0.04	0.65	81.7

	10	0.21	0.81	1.23	0.00	0.03	0.41	0.15	0.15	79.9
	11	0.73	0.43	1.05	0.00	0.19	0.00	0.68	0.00	80.3
G5	1	-1.49	0.00	1.06	0.00	-0.39	0.00	-0.08	0.01	90.5
	2	-0.58	0.36	0.96	0.00	-0.11	0.00	-0.19	0.00	77.6
	3	-0.42	0.49	0.93	0.00	-0.13	0.00	0.05	0.19	78.4
	4	0.62	0.35	0.94	0.00	-0.01	0.64	0.17	0.00	79.5
	5	1.22	0.09	1.06	0.00	0.04	0.18	0.13	0.01	80.6
	6	0.65	0.16	1.06	0.00	0.61	0.00	-0.07	0.02	93.9
	7	-0.53	0.53	1.22	0.00	-0.46	0.00	0.05	0.53	80.5
	8	0.27	0.74	1.32	0.00	-0.22	0.00	-0.14	0.03	78.0
	9	0.85	0.29	1.24	0.00	-0.14	0.00	0.10	0.11	82.0
	10	-0.15	0.86	1.21	0.00	0.04	0.34	0.18	0.02	80.5
	11	1.19	0.25	1.22	0.00	0.26	0.00	0.12	0.12	75.4

Table OA-8: **Detailed version of the top panel of Table 10**

Good	port.	avg. ret.	p-val	α	p-val	β_{DC}	p-val	β_{Good}	p-val	R^2
	1	-1.63	0.25	-3.23	0.02	0.38	0.00			12.6
	2	-0.19	0.88	-1.37	0.27	0.28	0.00			8.7
	3	0.79	0.54	-0.98	0.41	0.42	0.00			17.6
	4	2.80	0.05	0.18	0.88	0.63	0.00			34.3
	5	3.68	0.02	0.81	0.54	0.69	0.00			33.8
	6	4.65	0.01	2.27	0.18	0.57	0.00			17.7
	7	-0.18	0.92	-2.64	0.12	0.59	0.00			18.1
	8	1.04	0.57	-1.24	0.45	0.54	0.00			16.1
	9	2.76	0.12	0.17	0.91	0.62	0.00			21.7
	10	2.77	0.14	-0.18	0.91	0.70	0.00			26.5
	11	4.79	0.02	1.53	0.35	0.78	0.00			27.5
G1	1	-1.63	0.25	-1.02	0.48			-0.36	0.04	2.1
	2	-0.19	0.88	0.28	0.83			-0.28	0.04	1.6
	3	0.79	0.54	0.90	0.50			-0.07	0.66	-0.2
	4	2.80	0.05	2.13	0.15			0.40	0.01	2.7
	5	3.68	0.02	2.54	0.11			0.69	0.00	6.8
	6	4.65	0.01	2.77	0.12			1.12	0.00	14.4
	7	-0.18	0.92	0.56	0.76			-0.45	0.03	1.9
	8	1.04	0.57	1.96	0.30			-0.55	0.00	3.2
	9	2.76	0.12	2.31	0.20			0.27	0.26	0.6
	10	2.77	0.14	1.55	0.41			0.73	0.00	5.7
	11	4.79	0.02	2.36	0.20			1.45	0.00	20.0
G2	1	-1.63	0.25	-1.07	0.46			-0.33	0.02	1.9
	2	-0.19	0.88	0.04	0.98			-0.13	0.33	0.2
	3	0.79	0.54	0.64	0.63			0.09	0.51	-0.1
	4	2.80	0.05	2.00	0.17			0.47	0.00	4.3
	5	3.68	0.02	2.76	0.08			0.54	0.00	4.6
	6	4.65	0.01	3.36	0.07			0.76	0.00	7.2
	7	-0.18	0.92	0.46	0.80			-0.38	0.05	1.5
	8	1.04	0.57	1.31	0.49			-0.16	0.41	0.0
	9	2.76	0.12	2.21	0.22			0.32	0.11	1.1
	10	2.77	0.14	1.54	0.39			0.72	0.00	6.2
	11	4.79	0.02	3.32	0.09			0.86	0.00	7.7
G3	1	-1.63	0.25	-1.61	0.28			-0.01	0.95	-0.3
	2	-0.19	0.88	-0.41	0.76			0.09	0.37	0.0
	3	0.79	0.54	0.14	0.92			0.26	0.02	2.0
	4	2.80	0.05	1.48	0.31			0.53	0.00	7.7
	5	3.68	0.02	1.76	0.26			0.77	0.00	13.7
	6	4.65	0.01	1.71	0.29			1.18	0.00	24.7
	7	-0.18	0.92	-0.15	0.94			-0.01	0.93	-0.3
	8	1.04	0.57	0.79	0.68			0.10	0.50	-0.1
	9	2.76	0.12	1.52	0.40			0.50	0.00	4.3

	10	2.77	0.14	0.61	0.73			0.86	0.00	12.8
	11	4.79	0.02	1.18	0.49			1.45	0.00	30.9
G4	1	-1.63	0.25	-2.94	0.03			0.59	0.00	11.1
	2	-0.19	0.88	-1.22	0.33			0.47	0.00	8.8
	3	0.79	0.54	-0.61	0.62			0.63	0.00	14.6
	4	2.80	0.05	0.78	0.53			0.91	0.00	27.1
	5	3.68	0.02	1.22	0.38			1.11	0.00	33.0
	6	4.65	0.01	1.99	0.21			1.20	0.00	29.6
	7	-0.18	0.92	-1.99	0.25			0.82	0.00	12.9
	8	1.04	0.57	-0.41	0.82			0.65	0.00	8.6
	9	2.76	0.12	0.55	0.73			1.00	0.00	20.9
	10	2.77	0.14	0.15	0.93			1.18	0.00	27.9
	11	4.79	0.02	1.13	0.44			1.65	0.00	46.5
G5	1	-1.63	0.25	-3.33	0.01			0.43	0.00	9.9
	2	-0.19	0.88	-1.54	0.23			0.34	0.00	7.9
	3	0.79	0.54	-1.41	0.23			0.56	0.00	19.3
	4	2.80	0.05	-0.02	0.98			0.71	0.00	28.0
	5	3.68	0.02	0.65	0.65			0.76	0.00	26.2
	6	4.65	0.01	1.83	0.27			0.71	0.00	17.4
	7	-0.18	0.92	-2.72	0.12			0.64	0.00	13.5
	8	1.04	0.57	-1.24	0.47			0.57	0.00	11.3
	9	2.76	0.12	-0.38	0.81			0.79	0.00	22.4
	10	2.77	0.14	-0.81	0.60			0.90	0.00	27.4
	11	4.79	0.02	1.20	0.49			0.90	0.00	23.4
G1	1	-1.63	0.25	-2.48	0.05	0.54	0.00	-0.83	0.00	23.0
	2	-0.19	0.88	-0.81	0.50	0.40	0.00	-0.63	0.00	16.2
	3	0.79	0.54	-0.51	0.66	0.52	0.00	-0.52	0.00	22.1
	4	2.80	0.05	0.34	0.78	0.66	0.00	-0.17	0.19	34.6
	5	3.68	0.02	0.72	0.59	0.67	0.00	0.10	0.54	33.7
	6	4.65	0.01	1.59	0.34	0.43	0.00	0.75	0.00	22.9
	7	-0.18	0.92	-1.61	0.31	0.80	0.00	-1.14	0.00	30.0
	8	1.04	0.57	-0.14	0.93	0.77	0.00	-1.22	0.00	30.5
	9	2.76	0.12	0.47	0.77	0.68	0.00	-0.33	0.13	22.5
	10	2.77	0.14	-0.31	0.85	0.68	0.00	0.14	0.51	26.5
	11	4.79	0.02	0.70	0.66	0.61	0.00	0.92	0.00	34.2
G2	1	-1.63	0.25	-2.55	0.04	0.55	0.00	-0.80	0.00	23.1
	2	-0.19	0.88	-0.98	0.43	0.38	0.00	-0.46	0.00	13.0
	3	0.79	0.54	-0.69	0.56	0.49	0.00	-0.33	0.01	19.5
	4	2.80	0.05	0.25	0.84	0.64	0.00	-0.08	0.46	34.2
	5	3.68	0.02	0.86	0.51	0.70	0.00	-0.07	0.64	33.7
	6	4.65	0.01	1.99	0.24	0.50	0.00	0.32	0.06	18.6
	7	-0.18	0.92	-1.72	0.29	0.81	0.00	-1.08	0.00	29.7
	8	1.04	0.57	-0.59	0.72	0.70	0.00	-0.76	0.00	22.1
	9	2.76	0.12	0.39	0.80	0.67	0.00	-0.26	0.16	22.2
	10	2.77	0.14	-0.29	0.85	0.68	0.00	0.13	0.48	26.5
	11	4.79	0.02	1.34	0.42	0.73	0.00	0.23	0.20	27.8

G3	1	-1.63	0.25	-2.66	0.05	0.50	0.00	-0.43	0.00	16.3
	2	-0.19	0.88	-1.11	0.38	0.33	0.00	-0.19	0.07	9.4
	3	0.79	0.54	-0.82	0.49	0.45	0.00	-0.12	0.30	17.7
	4	2.80	0.05	0.17	0.89	0.62	0.00	0.01	0.95	34.1
	5	3.68	0.02	0.46	0.73	0.62	0.00	0.26	0.04	34.8
	6	4.65	0.01	1.04	0.51	0.32	0.01	0.91	0.00	29.0
	7	-0.18	0.92	-1.75	0.30	0.76	0.00	-0.65	0.00	23.5
	8	1.04	0.57	-0.62	0.70	0.67	0.00	-0.46	0.00	18.8
	9	2.76	0.12	0.21	0.90	0.63	0.00	-0.03	0.87	21.4
	10	2.77	0.14	-0.66	0.68	0.61	0.00	0.35	0.04	28.0
	11	4.79	0.02	0.14	0.93	0.50	0.00	1.03	0.00	39.6
G4	1	-1.63	0.25	-3.45	0.01	0.26	0.03	0.33	0.02	14.6
	2	-0.19	0.88	-1.57	0.21	0.17	0.11	0.29	0.01	10.7
	3	0.79	0.54	-1.20	0.31	0.30	0.01	0.33	0.01	20.0
	4	2.80	0.05	-0.12	0.92	0.46	0.00	0.46	0.00	38.5
	5	3.68	0.02	0.36	0.78	0.43	0.00	0.68	0.00	41.4
	6	4.65	0.01	1.60	0.30	0.19	0.13	1.01	0.00	30.7
	7	-0.18	0.92	-2.88	0.09	0.45	0.00	0.37	0.03	19.5
	8	1.04	0.57	-1.36	0.41	0.48	0.00	0.18	0.31	16.3
	9	2.76	0.12	-0.23	0.88	0.39	0.01	0.60	0.00	26.3
	10	2.77	0.14	-0.68	0.64	0.42	0.00	0.76	0.00	33.7
	11	4.79	0.02	0.62	0.66	0.26	0.08	1.39	0.00	48.3
G5	1	-1.63	0.25	-3.55	0.01	0.28	0.02	0.19	0.13	13.4
	2	-0.19	0.88	-1.68	0.18	0.18	0.08	0.18	0.08	9.7
	3	0.79	0.54	-1.59	0.17	0.23	0.03	0.36	0.00	21.8
	4	2.80	0.05	-0.37	0.75	0.45	0.00	0.33	0.00	37.3
	5	3.68	0.02	0.26	0.84	0.51	0.00	0.32	0.00	36.2
	6	4.65	0.01	1.56	0.33	0.34	0.01	0.42	0.00	20.7
	7	-0.18	0.92	-3.07	0.07	0.45	0.00	0.25	0.06	19.0
	8	1.04	0.57	-1.58	0.34	0.44	0.00	0.20	0.15	16.7
	9	2.76	0.12	-0.66	0.67	0.35	0.01	0.49	0.00	26.0
	10	2.77	0.14	-1.12	0.44	0.40	0.00	0.55	0.00	31.9
	11	4.79	0.02	0.78	0.63	0.54	0.00	0.44	0.00	30.3

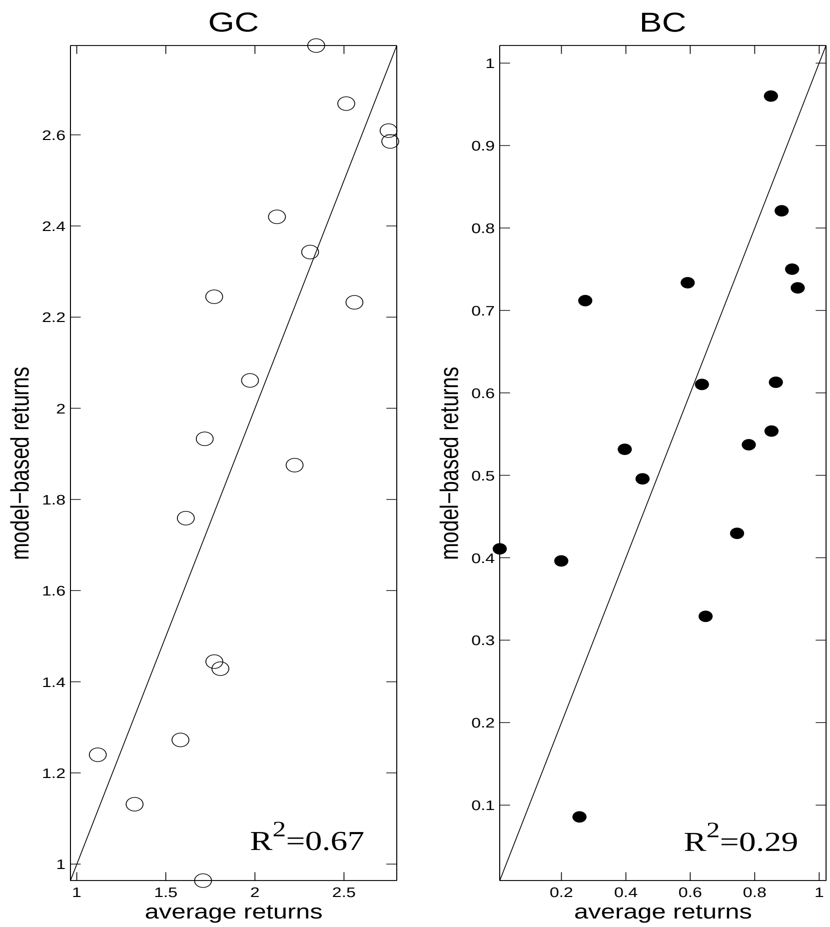

Figure 4: **Average vs. model-based expected returns**

Circles with no fill (large black dots) plot model-based expected monthly returns versus average monthly returns (annualized and in percent) for the GC (BC) set of 18 carry trades, as described in Table 3. The model based returns refer to the three-factor model with a market factor (MKT), an equity volatility factor (EqVol) and the product of MKT and EqVol, and are estimated, for each trade, as the product of its time-series slope estimates (β) with respect to the factors in the model, and the corresponding estimates of the factor risk prices λ, as shown in Table OA-2. The bottom right corner of each plot shows the R^2 obtained in regressing average returns on model-based returns (with a constant). The sample period is 12/1984 to 12/2013.