Do Carry-to-Risk Ratios and Risk Reversals Anticipate Unwinds of Yen-Funded Carry Trades?

Alessandro Mazzarini*

12 October 2026

Abstract

Rationale. Galati et al. (2007) proposed the carry-to-risk ratio as a measure of the attractiveness of currency carry trades and pointed to risk reversals as a measure of their directional risk. If either indicator anticipated the losses of carry trades, it would be an early warning for investors and supervisors who follow the build-up of these positions. If both react to losses, their value lies in measuring stress once an unwind has started.

Research question. Does the carry-to-risk ratio or the risk reversal anticipate the losses of yen-funded carry trades?

Data and methods. Both indicators are rebuilt daily for five yen-funded carry trades (the US, Australian, New Zealand and Canadian dollars and the pound sterling against the yen) from Bloomberg option and forward data, from 30 August 2007 to 16 October 2026, with a two-pair basket from 8 October 2004 as a robustness sample. The dependent variable is the excess return of an equally weighted carry basket. The analysis uses predictive regressions with Newey-West standard errors, Granger causality tests and a probit model for tail weeks, specified and checked against the framework of Greene (2003).

Main findings. Basket returns Granger-cause the changes in both indicators at every lag order tested ($p < 0.001$), while the hypothesis that changes in the risk reversal do not Granger-cause returns is never rejected (p between 0.253 and 0.950), which confirms on a new sample earlier evidence that risk reversals follow carry trade losses. Neither indicator has robust predictive power for returns or for tail weeks. The risk reversal remains useful for risk management because it stays high after the first losses. A rule that halves exposure when it is high lowers the maximum drawdown of the basket from −32.6% to −24.5%.

Keywords: Carry Trade, Risk Reversal, Carry-to-Risk Ratio, Japanese Yen, Granger Causality, Newey-West Estimator, Probit

J.E.L. classification: C22, C25, C58, F31, G15, G17

*Luiss "Guido Carli" University, Department of Economics and Finance, Viale Romania 32, 00197 Roma, Italy;

1 Introduction

A carry trade borrows in a low-yielding currency and invests the proceeds in a higher-yielding one. The position earns the interest rate differential as long as the funding currency does not appreciate by more than the differential over the holding period, which requires uncovered interest parity to fail. The Japanese yen has been the main funding currency of these trades for most of the last three decades, and their unwinds have produced some of the sharpest movements in the yen exchange rate, in October 1998, August 2007, October 2008 and August 2024. The episode of August 2024 has renewed the attention of policymakers. The Bank for International Settlements put the size of yen carry positions going into the event at about 40 trillion yen and described it as an example of volatility amplified by procyclical deleveraging and margin increases (Aquilina et al. 2024). A later BIS study finds that large short positions of carry traders in funding currencies amplify the exchange rate response to monetary policy tightening (Banerjee et al. 2026).

Galati et al. (2007) reviewed the data available in 2007 to track carry trade activity and discussed two market indicators. The carry-to-risk ratio divides the three-month interest rate differential by the implied volatility of three-month at-the-money options on the currency pair, and measures the attractiveness of the trade ex ante. The risk reversal, the difference between the implied volatilities of a call and a put that are equally out of the money, measures the price of protection against a move in one direction. Because the risk of a carry trade is concentrated in sudden appreciations of the funding currency, the risk reversal could capture it better than the at-the-money volatility. The same study reports, citing Gagnon and Chaboud (2007), that movements in risk reversals tend to post-date large exchange rate movements in periods of high volatility. Both indicators remain in use. Bhansali (2007) scales interest rate differentials by option-implied volatilities to assess the attractiveness of carry trades. The BIS gauges the incentives to enter yen carry trades with a carry-to-risk ratio computed, as in this report, from forward-implied differentials and three-month implied volatilities (Aquilina et al. 2024), and uses its level to separate states of high and low carry attractiveness (Banerjee et al. 2026).

The research relevant to the question falls into three strands. The first links carry trades to crash risk. Brunnermeier et al. (2009) show that high interest rate differentials predict negatively skewed exchange rate movements, that increases in the VIX coincide with carry trade losses and with the unwinding of speculative positions, and that after losses the price of insurance against crashes rises while the probability of a crash falls. Jurek (2014) finds that crash risk premia account for at most one third of the excess returns of G10 carry trades and that option-implied skewness moves with past carry returns. Hutchison and Sushko (2013) show that macroeconomic surprises move USD/JPY risk reversals and that risk reversals are closely linked to speculative futures positions, and Nirei and Sushko (2011) relate jumps in the yen exchange rate to the stochastic unwinding of carry trades.

The second strand asks which variables predict carry trade returns and losses, and volatility is the best documented answer. High interest rate currencies deliver low returns when global foreign exchange volatility rises unexpectedly (Menkhoff et al. 2012), and carry trades incur losses when realised and implied volatility rise (Clarida et al. 2009). In the time series, market variance predicts large future carry losses in quantile regressions (Cenedese et al. 2014), and changes in currency volatility, a commodity index and liquidity predict carry payoffs at horizons of up to six months (Bakshi and Panayotov 2013). Currency volatility risk premia predict currency returns (Della Corte et al. 2016), and the duration of carry drawdowns depends on the expected carry at their onset, on financial stress and on valuation (Melvin and Shand 2017).

The third strand tests the information content of risk reversals directly, and its results are mostly negative. Earlier studies of major currencies, summarised by Costa Filho (2016), find little predictive power of risk reversals for exchange rate returns. For ten emerging market currencies the same study finds that risk reversals help predict returns in a few cases, while in most cases they are predicted by currency and carry returns. Kütük and van Wijnbergen (2025) find for the Turkish lira that changes in risk reversals are related to expected carry returns measured by survey data and not to realised returns. Guyot et al. (2026) estimate a VAR for USD/JPY in 2018–2024 in which shocks to risk reversals account

for a sizeable share of the forecast error variance of the exchange rate, under an identification that orders risk reversals before the exchange rate.

This report asks whether either indicator anticipates the losses of yen-funded carry trades. If one of them did, it would be an early warning for investors and supervisors who follow the build-up of these positions. If both react to losses, their value lies in measuring stress once an unwind has started. Among the studies reviewed for this report, none tests the timing of both indicators on baskets of yen crosses over a sample that includes the unwinds of 2008, 2015–16, 2020 and August 2024, and none tests the carry-to-risk ratio as a predictor of carry trade returns. The report addresses the question with daily Bloomberg data on five yen crosses (the US, Australian, New Zealand and Canadian dollars and the pound sterling against the yen) from 30 August 2007 to 16 October 2026, and on a two-pair basket (US and Australian dollars) from 8 October 2004. It estimates predictive regressions, Granger causality tests in both directions and a probit model for tail weeks, with inference robust to heteroscedasticity and autocorrelation. Its contribution is an extension of existing evidence to a new sample and to the carry-to-risk ratio, documented so that every result can be reproduced with the code and tests released with it (Mazzarini 2026).

For these two indicators the answer is the same across methods. Basket returns Granger-cause the changes in both indicators at every lag order tested, with $p < 0.001$ in both samples. The hypothesis that changes in the risk reversal do not Granger-cause returns is never rejected (p between 0.253 and 0.950). The daily basket return and the daily change in the risk reversal have a correlation of −0.45 on the same day and −0.30 on the next day, and of at most 0.034 in absolute value at the lags in which the risk reversal moves first. Predictive regressions and a probit model for tail weeks find no robust predictive power. The risk reversal stays high after the first losses, so a rule that halves exposure when it is high lowers the maximum drawdown of the basket from −32.6% to −24.5%. These results confirm on a new sample the evidence that risk reversals follow carry trade losses (Brunnermeier et al. 2009; Jurek 2014). The tests do not cover volatility itself, which the second strand identifies as a predictor of carry losses, so the conclusions are limited to the two indicators.

The econometric framework and terminology are those of Greene (2003). Each estimator and test is tied to the section and equation of the textbook that defines it, the code is tested against those formulas, and Section 8 records every change made to the design during the analysis.

Section 2 sets out the hypotheses, Section 3 describes the data and Section 4 the construction of the variables. Section 5 presents the econometric framework, Section 6 the results and Section 7 the robustness checks. Section 8 is the research log and Section 9 concludes.

2 Background and Hypotheses

Let S_t be the price of one unit of the target currency in yen and i_t^X and i_t^{JPY} the interest rates of the two currencies for the holding period. Over a period of length Δ (in years), a position long the target currency and funded in yen earns the excess return

$$z_{t+1} = \frac{S_{t+1} - S_t}{S_t} + \left(i_t^X - i_t^{JPY}\right)\Delta, \tag{1}$$

up to second-order terms. Uncovered interest parity implies $E_t[z_{t+1}] = 0$. Carry trades have earned positive average excess returns over long samples (Burnside et al. 2006), with large and sudden losses that Brunnermeier et al. (2009) relate to crash risk priced in currency options and to funding liquidity.

The carry-to-risk ratio of Galati et al. (2007) scales the interest differential by the implied volatility of the pair. A high ratio signals that the expected carry is large relative to the risk priced by the option market. The risk reversal is quoted by Bloomberg as the implied volatility of a 25-delta call minus that of a 25-delta put on the first currency of the pair. In this report its sign is reversed, so that a positive value means that protection against a yen appreciation, the event that produces carry trade losses, is more expensive. This is the sign used in Graph 1 of Galati et al. (2007).

The study tests two hypotheses.

- **H1 (anticipation).** Increases in the risk reversal and falls in the carry-to-risk ratio precede the losses of the carry basket. Under H1, lagged changes in the indicators help predict basket returns and raise the probability of a tail week.
- **H2 (reaction).** The indicators respond to basket losses. Under H2, lagged basket returns help predict the changes in the indicators.

The two hypotheses are not mutually exclusive. Granger causality tests in both directions (Section 5.4) can support either, both or neither.

3 Data

All series are daily closing values (field PX_LAST) from Bloomberg, downloaded from the LUISS terminal with generic BGN pricing, so that spot rates and option quotes share the same timestamp. For each of the five pairs (USD/JPY, AUD/JPY, NZD/JPY, GBP/JPY, CAD/JPY) the study uses the spot rate, the three-month at-the-money implied volatility and the three-month 25-delta risk reversal. Interest rate differentials come from three-month forward points of the yen, Australian dollar, New Zealand dollar, pound and Canadian dollar against the US dollar. The VIX index is a control. The Fed funds target rate and the Bank of Japan policy rate are used for a policy-rate version of the carry-to-risk ratio for USD/JPY and in Figure 1. The full list of 27 series, with first date and number of observations, is in Table 18 in Appendix A.

Option histories start at different dates. Risk reversals begin on 2 October 2003 for USD/JPY and AUD/JPY and in March 2005 for the other pairs. The NZD/JPY and CAD/JPY option series also contain gaps of several months in 2006 and 2007. Rolling sums and differences require consecutive observations, so each basket is estimated on the sample that starts on the first day after the last gap longer than seven calendar days in any of its series. For the five-pair basket this gives 4,983 trading days from 30 August 2007 to 16 October 2026, with a largest remaining gap of 6 calendar days (a holiday period). A two-pair basket of USD/JPY and AUD/JPY, available from 8 October 2004 (5,744 days, largest gap 4 days), is used as a robustness sample and includes the unwind of August 2007. Isolated missing values of up to five trading days in daily signals are filled with the last available value.

Two checks were made against the terminal on 16 October 2026. The Bloomberg volatility surface (function OVDV) showed a USD/JPY three-month 25-delta risk reversal of $-1.830/-1.645$, with a mid of -1.74; the downloaded value for the same day is -1.7425, which becomes 1.74 with the sign used here. The one-month deposit rates on the same screen (3.985% for the dollar and 1.097% for the yen) imply a yen-minus-dollar differential of about -2.9 percentage points, against -2.98 at three months from the forward points used in this study.

Bloomberg data cannot be redistributed. The repository (Mazzarini 2026) contains the list of tickers, the code and the results, and reproduces every table and figure from a fresh download.

4 Construction of the Variables

4.1 Interest rate differentials

For a currency pair quoted as USD/XXX, with spot S_t and three-month forward F_t, covered interest parity with simple interest gives

$$\frac{F_t}{S_t} = \frac{1 + i_t^{XXX}\,\tau}{1 + i_t^{USD}\,\tau}, \qquad \tau = \frac{91}{360}. \tag{2}$$

The study uses the approximation

$$d_t^{XXX} \equiv i_t^{XXX} - i_t^{USD} \approx \frac{100}{\tau}\left(\frac{F_t}{S_t} - 1\right), \tag{3}$$

in percentage points per year, with the sign reversed for pairs quoted as XXX/USD (AUD, NZD and GBP). The approximation drops the factor $1/(1 + i_t^{USD}\tau)$, which changes the differential by a relative amount of about $i\tau$, below 1.5% for interest rates under 6%. Forward prices are $F_t = S_t + \text{points}_t/m$, with $m = 100$ for USD/JPY and $m = 10{,}000$ for the other pairs. The scale was checked on the Bloomberg forward matrix and against the deposit rates above. The spot rates against the dollar are rebuilt from the yen crosses (for example AUD/USD = AUD/JPY ÷ USD/JPY), so that every input has the same BGN timestamp. The differential of target currency k against the yen is

$$D_t^k = d_t^k - d_t^{JPY}, \qquad d_t^{USD} = 0. \tag{4}$$

A differential implied by forward prices includes the cross-currency basis, the deviation from covered interest parity that has been persistent since 2008. As a comparison, the carry-to-risk ratio of USD/JPY is also computed with the difference between the Fed funds target rate (upper bound since December 2008) and the Bank of Japan policy rate.

4.2 Carry-to-risk ratio and risk reversal

Following note 4 of Graph 1 in Galati et al. (2007), the carry-to-risk ratio is

$$CTR_t^k = \frac{D_t^k}{\sigma_t^k}, \tag{5}$$

where σ_t^k is the three-month at-the-money implied volatility of pair k, in per cent per year. The risk reversal with the sign of Galati et al. (2007) is

$$RR_t^k = -\left(\sigma_t^{k,\,25\Delta\ \text{call}} - \sigma_t^{k,\,25\Delta\ \text{put}}\right), \tag{6}$$

in volatility points, where call and put refer to the first currency of the pair.

4.3 Returns and baskets

The daily excess return of pair k follows equation (1) with the differential known on the previous day and the ACT/360 day count,

$$r_t^k = \frac{S_t^k - S_{t-1}^k}{S_{t-1}^k} + \frac{D_{t-1}^k}{100}\,\frac{n_t}{360}, \tag{7}$$

where n_t is the number of calendar days between two observations. The basket return r_t and the basket indicators CTR_t and RR_t are equally weighted averages over the pairs. Because the position is funded in yen, r_t is an excess return and its Sharpe ratio requires no further deduction of a risk-free rate.

4.4 Implied and realised volatility

Realised volatility σ_t^R is $100\sqrt{252}$ times the standard deviation of the daily log changes of the spot rate over the last 63 trading days. The difference between implied volatility and the realised volatility over the next 63 days, $\sigma_t - \sigma_{t+63}^R$, is the ex post volatility premium.

5 Econometric Framework

Section, equation and page numbers refer to Greene (2003). Table 20 in Appendix C maps every step to the textbook.

5.1 Stationarity

The asymptotic results for least squares with time series data rest on the ergodic theorem (Theorem 12.1), which requires the series to be stationary and ergodic (Section 12.4). The study therefore tests every series used in estimation for a unit root with the Dickey-Fuller regression with a constant (Dickey and Fuller 1979) (Section 20.3.4),

$$y_t = \mu + \gamma\, y_{t-1} + \sum_{j=1}^{p} \phi_j\, \Delta y_{t-j} + \varepsilon_t, \qquad H_0 : \gamma = 1, \tag{8}$$

tested with the t ratio $DF_\tau = (\hat{\gamma} - 1)/\text{Est.Std.Error}(\hat{\gamma})$. The lagged differences make the disturbances serially uncorrelated (the augmented test), and their number p is chosen by the Akaike criterion up to $12(T/100)^{1/4}$. The statistic is compared with the Dickey-Fuller critical values for the model with a constant (Table 20.4, 5% asymptotic value -2.86), and p-values follow MacKinnon (2010). A series that does not reject the unit root enters the regressions in differences (Section 20.3.1).

5.2 Predictive regressions

For horizons of $h = 1$, 5 and 21 trading days the study estimates by least squares

$$Y_{t,h} \equiv \sum_{j=1}^{h} r_{t+j} = \alpha + \beta_1\, \Delta_5 CTR_t + \beta_2\, \Delta_5 RR_t + u_t, \tag{9}$$

with $Y_{t,h}$ in per cent and the two regressors, the changes in the basket indicators over the previous five trading days, standardised to unit variance. Standardisation rescales the coefficients and leaves t ratios and fit unchanged. Model M1 includes $\Delta_5 CTR_t$ only, M2 $\Delta_5 RR_t$ only and M3 both. Under H1 the coefficient on $\Delta_5 RR_t$ is negative and the one on $\Delta_5 CTR_t$ is positive.

When $h > 1$ consecutive dependent variables share $h - 1$ daily returns, so the disturbances follow a moving average process of order $h - 1$ even under the null hypothesis of no predictability. Least squares remains consistent, but $s^2(\mathbf{X}'\mathbf{X})^{-1}$ is an inappropriate estimator of its asymptotic covariance matrix (Section 12.5.2). The study uses the Newey-West robust, consistent estimator for autocorrelated disturbances with an unspecified structure (Newey and West 1987), eq. (12-17),

$$\mathbf{S}_* = \frac{1}{T}\sum_{t=1}^{T} e_t^2\, \mathbf{x}_t\mathbf{x}_t' + \frac{1}{T}\sum_{j=1}^{L}\sum_{t=j+1}^{T} \left(1 - \frac{j}{L+1}\right) e_t e_{t-j} \left(\mathbf{x}_t\mathbf{x}_{t-j}' + \mathbf{x}_{t-j}\mathbf{x}_t'\right), \tag{10}$$

$$\text{Est.Asy.Var}[\mathbf{b}] = \frac{1}{T}\left(\frac{\mathbf{X}'\mathbf{X}}{T}\right)^{-1} \mathbf{S}_* \left(\frac{\mathbf{X}'\mathbf{X}}{T}\right)^{-1}. \tag{11}$$

The maximum lag L must be large enough that autocorrelations at longer lags are negligible, and Greene (p. 267) reports the current practice of the smallest integer greater than or equal to $T^{1/4}$. Because the overlap creates autocorrelation up to lag $h-1$ by construction, the study sets $L = \max(h, \lceil T^{1/4} \rceil)$, which gives $L = 9$ for $h = 1$ and $h = 5$ and $L = 21$ for $h = 21$. The estimator is robust to heteroscedasticity as well, no degrees-of-freedom correction is applied, and the ratios of the coefficients to their standard errors are compared with the standard normal distribution.

Two robustness checks complete the analysis. For $h > 1$ the regression is repeated on each of the h non-overlapping subsamples (one for each starting day), with Newey-West standard errors and $L = \lceil N^{1/4} \rceil$, because the residual tests of Section 5.3 show that non-overlapping weekly and monthly residuals are not free of autocorrelation. The regression is also estimated pair by pair for $h = 5$ and $h = 21$.

5.3 Residual diagnostics

Three tests are applied to the least squares residuals e_t. The Breusch-Godfrey Lagrange multiplier test (Breusch 1978; Godfrey 1978), eq. (12-20), is

$$LM = T\,\frac{\mathbf{e}'\mathbf{X}_0(\mathbf{X}_0'\mathbf{X}_0)^{-1}\mathbf{X}_0'\mathbf{e}}{\mathbf{e}'\mathbf{e}} \;\xrightarrow{d}\; \chi^2[P], \tag{12}$$

where $\mathbf{X}_0$ is the regressor matrix augmented with $P = 5$ lagged residuals. Greene (p. 269) notes that software differs in the treatment of the first P lagged residuals and asks users to state the convention. Here the missing values are filled with zeros. The test remains valid when the regression contains a lagged dependent variable (Section 12.7.4). The Ljung-Box refinement of the Box-Pierce statistic (Ljung and Box 1978), eq. (12-22), is

$$Q' = T(T+2)\sum_{j=1}^{P}\frac{r_j^2}{T-j}, \qquad r_j = \frac{\sum_{t=j+1}^{T} e_t e_{t-j}}{\sum_{t=1}^{T} e_t^2}, \tag{13}$$

with $P = 10$. It uses simple correlations and is less powerful than the LM test when the null hypothesis is false (p. 270). The LM test for ARCH effects (Engle 1982) (Section 11.8.4) is TR^2 in the regression of e_t^2 on a constant and $q = 5$ lags of e_t^2, distributed as $\chi^2[q]$ under the null hypothesis.

The tests are computed on the residuals of the daily regressions ($h = 1$) and of the first non-overlapping subsample for $h = 5$ and $h = 21$. The residuals of overlapping regressions are autocorrelated by construction and are not tested.

5.4 Lead-lag correlations and Granger causality

The first description of timing is the correlation between the daily basket return at t and the daily change in the risk reversal at $t + k$, for $k = -10, \ldots, 10$. Negative values of k are lags at which the risk reversal moves first. The band $\pm 1.96/\sqrt{T}$ assumes independent observations and is only approximate under heteroscedasticity.

Causality in the sense of Granger (1969) is tested in single equations (Section 19.6.5),

$$y_t = \mu + \sum_{j=1}^{p}\gamma_j\, y_{t-j} + \sum_{j=1}^{p}\delta_j\, x_{t-j} + \varepsilon_t, \qquad H_0\colon \delta_1 = \cdots = \delta_p = 0, \tag{14}$$

for the four pairs (y, x) given by the basket return and the daily changes in RR_t and CTR_t, in both directions. The hypothesis is tested with the Wald statistic

$$W = \hat{\boldsymbol{\delta}}'\left\{\text{Est.Asy.Var}[\hat{\boldsymbol{\delta}}]\right\}^{-1}\hat{\boldsymbol{\delta}} \;\xrightarrow{d}\; \chi^2[p], \tag{15}$$

with the Newey-West covariance matrix and $L = \lceil T^{1/4} \rceil$. Greene (p. 593) points to the Wald test as the more general alternative when the normality assumption behind the likelihood ratio test is doubtful, and the ARCH tests of Section 6.4 make it doubtful here. The classical F statistic is reported for comparison.

The regression contains lagged values of the dependent variable. Greene (p. 266) shows that least squares is then inconsistent if the disturbances are autocorrelated. The lag order p is therefore chosen as the smallest value in $\{5, 10, 15, 20, 25, 30\}$ for which the Breusch-Godfrey test does not reject at 5%, and results are reported for every order. A finding of Granger causality establishes precedence in time. As Greene notes (p. 593), an omitted variable correlated with both series could produce the same pattern.

5.5 Probit for tail weeks

The probit model addresses H1 directly in the tails. Observations are weekly and do not overlap. From a starting day, every fifth trading day t_s forms observation s. The outcome is

$$y_s = \mathbf{1}\Big\{ \textstyle\sum_{j=1}^{5} r_{t_s+j} \le q_\alpha \Big\}, \tag{16}$$

where q_α is the empirical α quantile of the weekly returns in the same sample, with $\alpha = 5\%$ and 10%. The index function model (Section 21.3.2) is $y_s^* = \boldsymbol{\beta}'\mathbf{x}_s + \varepsilon_s$ with $y_s = 1$ if $y_s^* > 0$, so that

$$\text{Prob}(y_s = 1 \mid \mathbf{x}_s) = \Phi(\boldsymbol{\beta}'\mathbf{x}_s), \tag{17}$$

$$\ln L = \sum_s \Big[y_s \ln \Phi(\boldsymbol{\beta}'\mathbf{x}_s) + (1 - y_s) \ln\big(1 - \Phi(\boldsymbol{\beta}'\mathbf{x}_s)\big) \Big]. \tag{18}$$

The regressors, known at t_s, are a constant and the standardised $\Delta_5 CTR_{t_s}$ and $\Delta_5 RR_{t_s}$; a second specification adds the five-day change in the VIX. Non-overlapping weeks keep the independence assumption behind equation (18) plausible. A weekly grid can start on any of five days, and every specification is estimated on all five grids.

Following Section 21.4, the study reports the maximum likelihood estimates with standard errors from the actual Hessian, the usual choice (p. 673), and with the sandwich estimator of White (1982) (Section 21.4.1) as a sensitivity check, because a robust covariance matrix does not make an inconsistent estimator consistent (p. 674). Marginal effects $\phi(\hat{\boldsymbol{\beta}}'\bar{\mathbf{x}})\hat{\boldsymbol{\beta}}$ are computed at the means of the regressors, with standard errors from the delta method (Section 21.4.2). The joint significance of the slopes is tested with the likelihood ratio statistic $LR = 2(\ln L - \ln L_0)$, where $\ln L_0$ is the log-likelihood with a constant only, eq. (21-28), and the two signals with the Wald statistic of eq. (21-27). Fit is measured with McFadden's likelihood ratio index (McFadden 1974) $LRI = 1 - \ln L / \ln L_0$ (Section 21.4.5).

5.6 Descriptive tools

Unwind episodes. A day belongs to an unwind when the five-day basket return falls in the worst 1% of days. Days less than 60 trading days apart form one episode, dated by its first day. The quantile uses the whole sample, so this definition is descriptive and is never used as a signal.

Event study. The basket return, the risk reversal and the carry-to-risk ratio are followed from 40 trading days before to 20 days after the first day of each episode and set to zero on day −40.

Exposure rule. The exposure to the basket on day t is

$$w_t = 1 - 0.5 \cdot \mathbf{1}\{RR_{t-1} > Q_{0.9}(RR_{t-252}, \ldots, RR_{t-1})\}, \tag{19}$$

where $Q_{0.9}$ is the 90th percentile of the previous 252 observations. The threshold, the window and the reduction were fixed before the rule was evaluated, and only information available on day $t-1$ is used. Annualised return and volatility are $252\,\bar{r}$ and $\sqrt{252}\,\text{sd}(r)$, the Sharpe ratio is their ratio and the maximum drawdown is computed on the compounded index.

5.7 Number of tests

The study reports several dozen tests across horizons, models, pairs, lag orders and weekly grids. At the 5% level some rejections are expected by chance, so the analysis gives weight to patterns that hold across specifications and samples and treats isolated rejections as such.

6 Results

6.1 The indicators over time

Figure 1 shows the divergence of policy rates that has made the yen a funding currency. Figure 2 shows the USD/JPY exchange rate, the carry-to-risk ratio with the two measures of the differential, and

the risk reversal. The ratio reached 0.76 in June 2007, collapsed after the failure of Lehman Brothers, stayed close to zero while both policy rates were near their lower bounds, and returned to 0.73 in November 2023. The two versions of the ratio have a correlation of 0.99 since 2000. The forward-implied differential exceeded the policy-rate differential by 0.15 percentage points on average in 2000–2007 and by 0.29 points in 2008–2026, a gap that combines the term premium of three-month rates and the cross-currency basis. The risk reversal peaked at 10.36 volatility points on 20 November 2008.

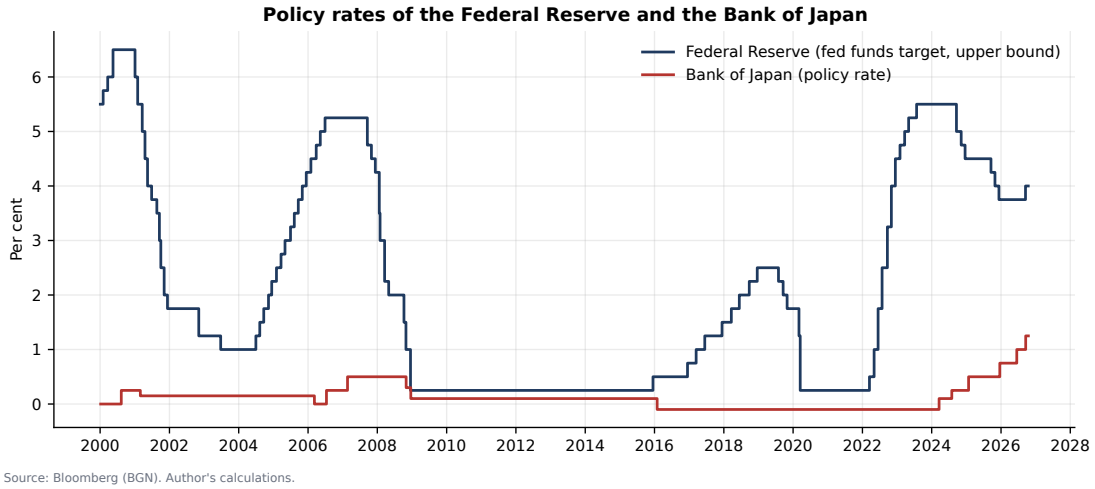


Figure 1: Policy rates of the Federal Reserve and the Bank of Japan. Source: Bloomberg (FDTR, BOJDTR).

Implied volatility exceeded the volatility realised over the following 63 days by 0.58 points on average since 2000, and was higher on 67% of days (Figure 3). A ratio built on a volatility forecast that tracks realised volatility, such as a GARCH(1,1) estimate, is therefore higher on average than the ratio built on implied volatility. Table 1 reports the latest values.

Table 1: USD/JPY indicators on 16 October 2026

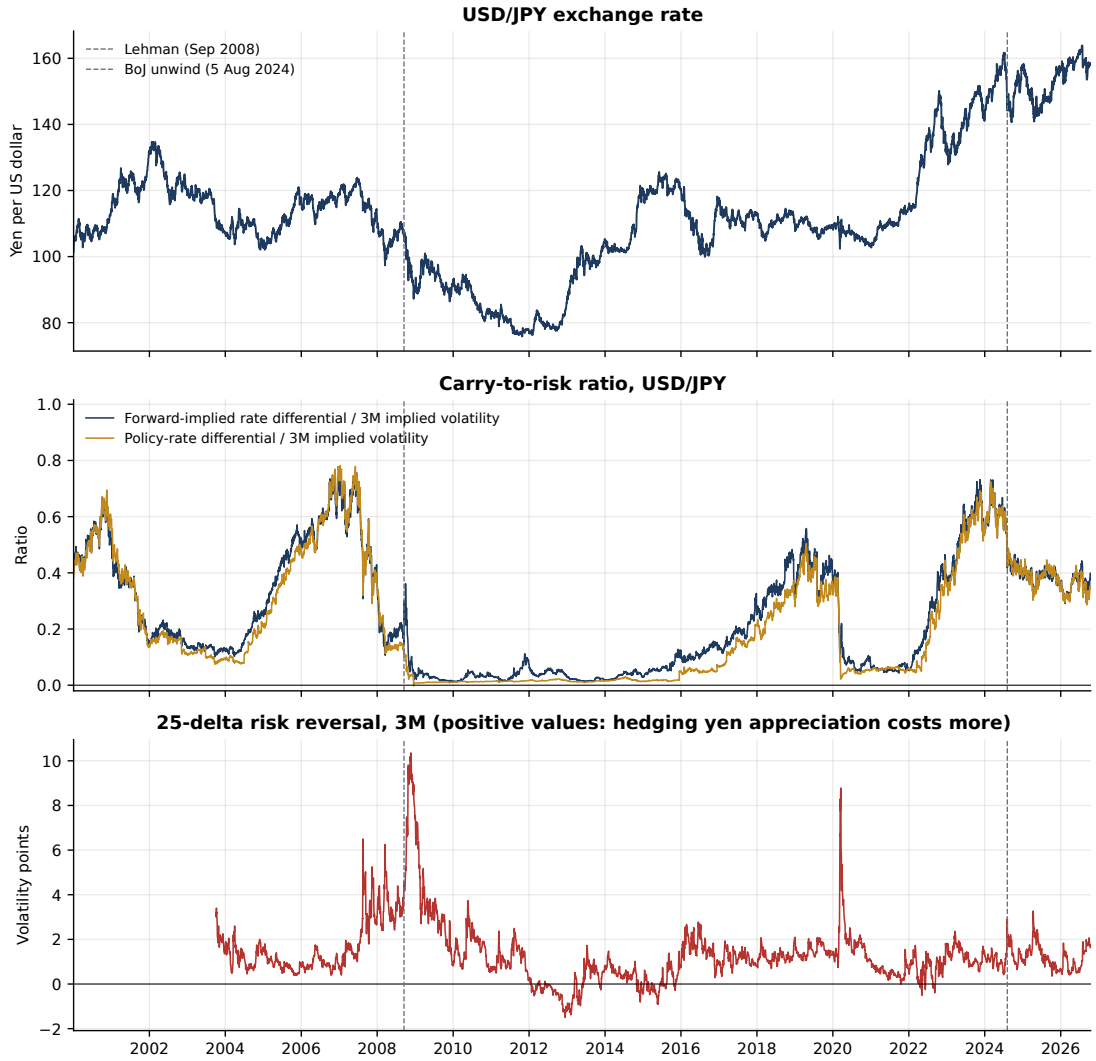
Indicator	Value
Three-month at-the-money implied volatility (%)	7.86
Forward-implied differential, USD minus JPY (percentage points)	2.98
Carry-to-risk ratio, forward-implied differential	0.38
Carry-to-risk ratio, policy-rate differential	0.35
25-delta three-month risk reversal, sign of Galati et al. (2007) (vol points)	1.74

6.2 Unit roots

Table 2 reports the augmented Dickey-Fuller tests on the main sample. The level of the basket carry-to-risk ratio does not reject a unit root (statistic -2.33, $p = 0.162$), and neither do 4 of the five pair-level ratios. The exception, AUD/JPY, rejects at 5% by a small margin (-2.871 against a critical value of -2.862, $p = 0.049$). The ratio moves with policy rates, which change in long regimes, and the series used in estimation behaves as an integrated process in this sample. Its level is therefore used only in charts. The level of the risk reversal rejects the unit root (-4.50), and every change used in the regressions rejects it strongly, the least negative statistic being -11.62.

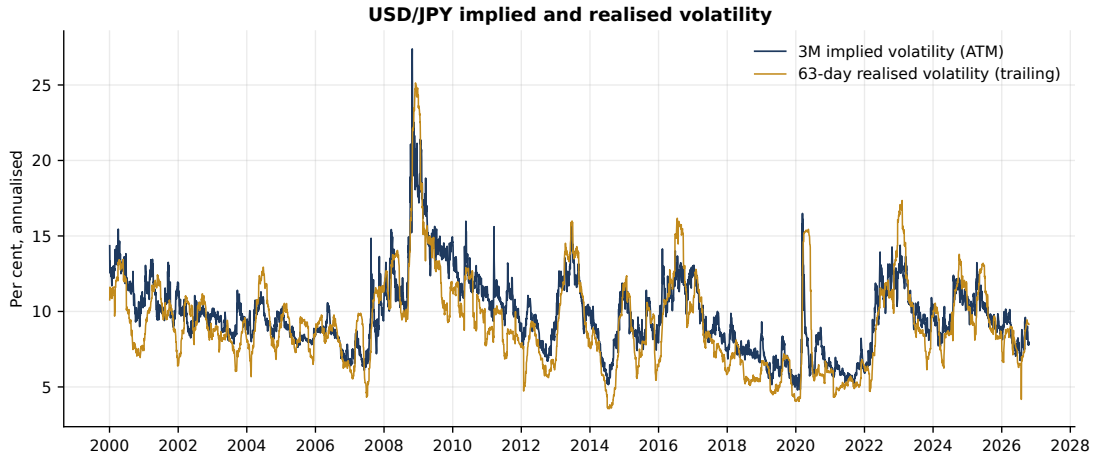
6.3 Predictive regressions

Table 3 reports equation (9). Of the 6 coefficients of model M3, 1 is significant at 5%, the coefficient on $\Delta_5 CTR$ at the horizon of 1 day ($z = 2.08$, $p = 0.037$). Its positive sign means that a fall in the carry-to-risk



Source: Bloomberg (BGN). Author's calculations. Risk reversal sign as in Galati, Heath and McGuire (2007).

Figure 2: USD/JPY exchange rate, carry-to-risk ratio (equation (5)) with forward-implied and policy-rate differentials, and 25-delta three-month risk reversal (equation (6)). Dashed lines mark 15 September 2008 and 5 August 2024. Source: Bloomberg (BGN); author's calculations.



Source: Bloomberg (BGN). Author's calculations.

Figure 3: USD/JPY three-month at-the-money implied volatility and 63-day trailing realised volatility. Source: Bloomberg (BGN); author's calculations.

Table 2: Augmented Dickey-Fuller tests, main sample

Series	ADF statistic	p-value	Lags	N	Unit root rejected (5%)
basket return	−13.21	<0.001	28	4,954	yes
basket carry-to-risk (level)	−2.33	0.162	25	4,957	no
change in basket carry-to-risk (5d)	−11.62	<0.001	30	4,947	yes
daily change in basket carry-to-risk	−14.72	<0.001	24	4,957	yes
basket risk reversal (level)	−4.50	<0.001	16	4,966	yes
change in basket risk reversal (5d)	−12.73	<0.001	32	4,945	yes
daily change in basket risk reversal	−18.09	<0.001	15	4,966	yes
VIX (level)	−5.77	<0.001	11	4,971	yes
change in VIX (5d)	−13.80	<0.001	32	4,945	yes
carry-to-risk USDJPY (level)	−2.00	0.288	32	4,957	no
carry-to-risk AUDJPY (level)	−2.87	0.049	30	4,959	yes
carry-to-risk NZDJPY (level)	−2.74	0.067	32	4,957	no
carry-to-risk GBPJPY (level)	−1.84	0.359	25	4,960	no
carry-to-risk CADJPY (level)	−2.19	0.209	32	4,954	no

Notes: Regression with a constant (equation (8)). Lags chosen by the Akaike criterion. Critical value at 5%: −2.86. p-values from MacKinnon (2010).

ratio over the previous week, which typically occurs when implied volatility rises during losses, is followed by a lower return on the next day. The effect is a short continuation of losses already under way and vanishes at one week and one month. The R^2 never exceeds 0.002. Across all 12 coefficients of models M1 to M3, 1 is significant at 5%.

Table 3: Predictive regressions of future basket returns, main sample

h	Model	Regressor	Coefficient	z	p-value	R^2	N	L
1	M1	$\Delta_5 CTR$	0.015	1.25	0.210	0.0004	4,977	9
1	M2	$\Delta_5 RR$	0.017	0.94	0.347	0.0005	4,977	9
1	M3	$\Delta_5 CTR$	0.029	2.08	0.037	0.0017	4,977	9
1	M3	$\Delta_5 RR$	0.030	1.42	0.154	0.0017	4,977	9
5	M1	$\Delta_5 CTR$	0.011	0.25	0.805	0.0000	4,973	9
5	M2	$\Delta_5 RR$	0.049	0.79	0.432	0.0009	4,973	9
5	M3	$\Delta_5 CTR$	0.042	0.85	0.397	0.0015	4,973	9
5	M3	$\Delta_5 RR$	0.068	0.97	0.333	0.0015	4,973	9
21	M1	$\Delta_5 CTR$	−0.005	−0.06	0.953	0.0000	4,957	21
21	M2	$\Delta_5 RR$	0.019	0.20	0.845	0.0000	4,957	21
21	M3	$\Delta_5 CTR$	0.004	0.04	0.969	0.0000	4,957	21
21	M3	$\Delta_5 RR$	0.021	0.19	0.850	0.0000	4,957	21

Notes: Least squares estimates of equation (9). Dependent variable: basket excess return over the next h trading days, in per cent. Regressors standardised. z ratios from the Newey-West covariance matrix (equations (10) and (11)) with lag L, compared with the standard normal distribution.

The non-overlapping subsamples confirm the result (Table 4). The median z ratio lies between 0.06 and 0.69. At one week no subsample gives $|z| > 1.96$. At one month 2 of the 21 subsamples do so for the risk reversal, with z ratios of opposite sign (−2.77 and 2.33).

Pair by pair (Table 5), 2 of 20 coefficients are significant at 5% (USD/JPY, $\Delta_5 RR$, $h = 5$, $z = 2.50$; CAD/JPY, $\Delta_5 CTR$, $h = 5$, $z = 2.04$), about the number expected by chance.

Table 4: Model M3 on non-overlapping subsamples

h	Regressor	Subsamples	N each	z min	z median	z max	Share $\lvert z\rvert > 1.96$
5	$\Delta_5 CTR$	5	995	0.21	0.45	1.80	0%
5	$\Delta_5 RR$	5	995	0.36	0.69	1.30	0%
21	$\Delta_5 CTR$	21	237	−1.92	0.06	1.88	0%
21	$\Delta_5 RR$	21	237	−2.77	0.24	2.33	10%

Notes: One regression for each starting day. Newey-West standard errors with $L = \lceil N^{1/4} \rceil$.

Table 5: Model M3 by currency pair, z ratios (p-values)

Pair	$h = 5$, $\Delta_5 CTR$	$h = 5$, $\Delta_5 RR$	$h = 21$, $\Delta_5 CTR$	$h = 21$, $\Delta_5 RR$
USD/JPY	1.64 (0.102)	2.50 (0.012)	0.01 (0.992)	0.98 (0.327)
AUD/JPY	−0.67 (0.500)	−0.05 (0.958)	−0.16 (0.873)	0.19 (0.846)
NZD/JPY	−0.11 (0.912)	0.54 (0.591)	−0.36 (0.718)	−0.15 (0.883)
GBP/JPY	1.22 (0.222)	1.47 (0.142)	1.15 (0.248)	0.24 (0.812)
CAD/JPY	2.04 (0.041)	1.83 (0.067)	0.88 (0.380)	1.03 (0.302)

Notes: Pair-level excess returns and indicators, main sample. Newey-West standard errors with $L = \max(h, \lceil T^{1/4} \rceil)$.

6.4 Residual diagnostics

Table 6 supports the choice of covariance estimator. The ARCH test rejects strongly at daily frequency (LM = 918.0) and at monthly frequency ($p = 0.005$), so the disturbances are conditionally heteroscedastic. The daily residuals show mild autocorrelation (Breusch-Godfrey $p = 0.056$, Ljung-Box $p = 0.029$). The non-overlapping weekly and monthly residuals reject the absence of autocorrelation (Breusch-Godfrey $p = 0.001$ and 0.043). For this reason the non-overlapping regressions use the Newey-West estimator instead of the heteroscedasticity-consistent estimator of White (1980), which was the choice of a first version of the study (Section 8).

Table 6: Residual diagnostics, main sample

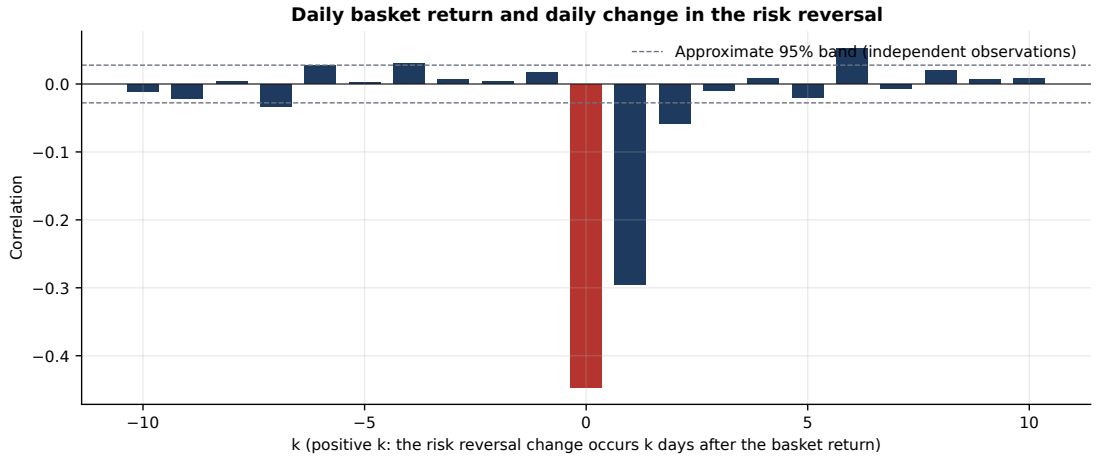
h	Model	Sample	N	LM (BG), $P = 5$	Q' (LB), $P = 10$	LM (ARCH), $q = 5$
1	M1	all days	4,977	16.24 (0.006)	26.57 (0.003)	898.5 (<0.001)
1	M2	all days	4,977	9.11 (0.105)	18.78 (0.043)	917.8 (<0.001)
1	M3	all days	4,977	10.79 (0.056)	20.03 (0.029)	918.0 (<0.001)
5	M1	non-overlapping, first grid	995	25.91 (<0.001)	44.50 (<0.001)	179.5 (<0.001)
5	M2	non-overlapping, first grid	995	19.17 (0.002)	36.38 (<0.001)	159.3 (<0.001)
5	M3	non-overlapping, first grid	995	19.82 (0.001)	36.57 (<0.001)	159.8 (<0.001)
21	M1	non-overlapping, first grid	237	10.94 (0.053)	13.63 (0.191)	17.1 (0.004)
21	M2	non-overlapping, first grid	237	11.49 (0.042)	14.33 (0.159)	17.6 (0.003)
21	M3	non-overlapping, first grid	237	11.46 (0.043)	14.27 (0.161)	16.8 (0.005)

Notes: Statistics with p-values in parentheses. Breusch-Godfrey LM, equation (12), with zero-filled initial lags; Ljung-Box Q', equation (13); ARCH LM, TR^2 with q lags (Greene, Section 11.8.4).

6.5 Lead-lag correlations and Granger causality

Figure 4 shows the correlations between the daily basket return and the daily change in the risk reversal. The correlation is −0.45 on the same day, −0.30 when the risk reversal is measured one day later and −0.06 two days later. At the ten lags in which the risk reversal moves first, 2 correlations exceed the

approximate band of ±0.028 by a small amount, $k = -7$ (−0.034) and $k = -4$ (+0.030), with opposite signs and magnitudes below one tenth of the same-day correlation.



Source: Bloomberg (BGN). Author's calculations.

Figure 4: Correlation between the daily basket return at t and the daily change in the basket risk reversal at $t + k$. Positive k: the risk reversal moves after the return. The band assumes independent observations. Source: Bloomberg (BGN); author's calculations.

The Granger tests in Table 7 formalise the pattern. Lagged basket returns help predict the changes in the risk reversal and in the carry-to-risk ratio, with $p < 0.001$ at every lag order. Lagged changes in the risk reversal do not help predict returns ($p = 0.795$ at the preferred order of 5 lags, between 0.253 and 0.831 across orders). For the carry-to-risk ratio the preferred order gives $p = 0.112$, and one of the six orders gives a p-value below 5% (0.033). The classical F test at the preferred order gives $p = 0.035$, which shows how much the treatment of heteroscedasticity matters for this hypothesis. The evidence supports H2 for both indicators and rejects H1 for the risk reversal.

Table 7: Granger causality tests, main sample

Null hypothesis	p	N	Wald (NW)	F (classical)	LM (BG)	Wald p-value, all p
ΔRR does not Granger-cause r	5	4,977	2.38 (0.795)	0.95 (0.445)	10.78 (0.056)	0.253–0.831
r does not Granger-cause ΔRR	20	4,962	136.64 (<0.001)	20.76 (<0.001)	2.84 (0.724)	<0.001
ΔCTR does not Granger-cause r	5	4,977	8.93 (0.112)	2.40 (0.035)	7.37 (0.194)	0.033–0.296
r does not Granger-cause ΔCTR	10	4,972	56.24 (<0.001)	8.44 (<0.001)	6.27 (0.281)	<0.001

Notes: Single-equation regressions, equation (14). p: preferred lag order, the smallest in $\{5, \ldots, 30\}$ whose residuals pass the Breusch-Godfrey test at 5%. Wald statistic (equation (15)) with Newey-West covariance; classical F test for comparison; LM (BG) on the residuals. Statistics with p-values in parentheses. Last column: range of Wald p-values over the six lag orders.

6.6 Probit for tail weeks

Table 8 summarises the probit estimates over the five weekly grids. With a 5% tail and no controls, the likelihood ratio test rejects the null hypothesis of zero slopes at 5% on 3 of the five grids, with p-values from 0.001 to 0.385 and a mean likelihood ratio index of 0.018. The other specifications reject on at most two grids, and the long sample on one grid (Section 7). A relation that appears or disappears with the starting day of the week, with a likelihood ratio index below 0.02, is not robust. Tables 9 and 10 report fit measures and marginal effects for the first grid, and Table 19 in Appendix B compares the Hessian and sandwich z ratios. Figure 5 shows that the fitted probability rises mostly during and after the stress episodes.

Table 8: Probit for tail weeks across the five weekly grids, main sample

Tail	Controls	N	Events	LR p min	LR p median	LR p max	Grids $p < 0.05$	Mean LRI
5%	none	995	50	0.001	0.045	0.385	3 of 5	0.018
5%	VIX	995	50	0.002	0.082	0.583	2 of 5	0.020
10%	none	995	100	0.014	0.287	0.896	1 of 5	0.006
10%	VIX	995	100	0.036	0.401	0.962	1 of 5	0.006

Notes: Likelihood ratio test of all slopes on each of the five grids. Controls: five-day change in the VIX. LRI: McFadden's likelihood ratio index, averaged over the grids.

Table 9: Probit fit measures and tests, first weekly grid

Tail	Controls	N	Events	Threshold (%)	$\ln L$	$\ln L_0$	LRI	LR [df] (p)	Wald, signals (p)
5%	none	995	50	−2.58	−195.2	−198.3	0.016	6.21 [2] (0.045)	6.35 (0.042)
5%	VIX	995	50	−2.58	−194.9	−198.3	0.017	6.70 [3] (0.082)	3.20 (0.201)
10%	none	995	100	−1.66	−322.4	−324.6	0.007	4.37 [2] (0.112)	4.41 (0.110)
10%	VIX	995	100	−1.66	−321.8	−324.6	0.008	5.49 [3] (0.139)	5.47 (0.065)

Notes: Threshold: the α quantile of weekly basket returns. $\ln L_0$: constant only, eq. (21-28). LR: likelihood ratio test of all slopes, degrees of freedom in brackets. Wald: joint test of the two signals, eq. (21-27).

Table 10: Probit marginal effects at the means, first weekly grid

Specification	Variable	Marginal effect	Std. error (delta)	z	p-value
5%, controls: none	$\Delta_5 CTR$	−0.0064	0.0069	−0.92	0.358
5%, controls: none	$\Delta_5 RR$	0.0107	0.0064	1.67	0.094
5%, controls: VIX	$\Delta_5 CTR$	−0.0058	0.0070	−0.84	0.401
5%, controls: VIX	$\Delta_5 RR$	0.0084	0.0071	1.19	0.236
5%, controls: VIX	$\Delta_5 VIX$	0.0049	0.0070	0.70	0.483
10%, controls: none	$\Delta_5 CTR$	−0.0088	0.0101	−0.86	0.387
10%, controls: none	$\Delta_5 RR$	0.0125	0.0096	1.30	0.192
10%, controls: VIX	$\Delta_5 CTR$	−0.0096	0.0102	−0.95	0.344
10%, controls: VIX	$\Delta_5 RR$	0.0174	0.0107	1.62	0.105
10%, controls: VIX	$\Delta_5 VIX$	−0.0109	0.0104	−1.05	0.293

Notes: Change in the probability of a tail week for a one standard deviation change in the regressor. Standard errors from the delta method (Greene, Section 21.4.2).

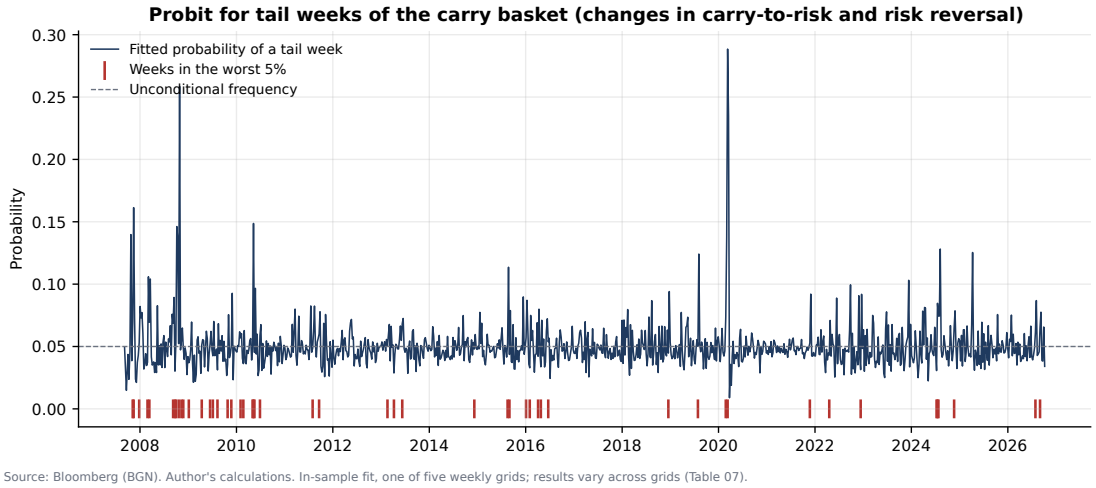


Figure 5: Fitted probability of a tail week (5% tail, no controls) on the first weekly grid. Source: Bloomberg (BGN); author's calculations.

6.7 Unwind episodes and event study

The rule of Section 5.6 identifies 11 unwind episodes in the main sample, with a threshold of −4.50% for the five-day basket return (Table 11). The largest five-day loss, −18.36%, ended on 27 October 2008. Figure 6 shows the basket and its indicators with the episodes marked.

Table 11: Unwind episodes of the five-pair basket

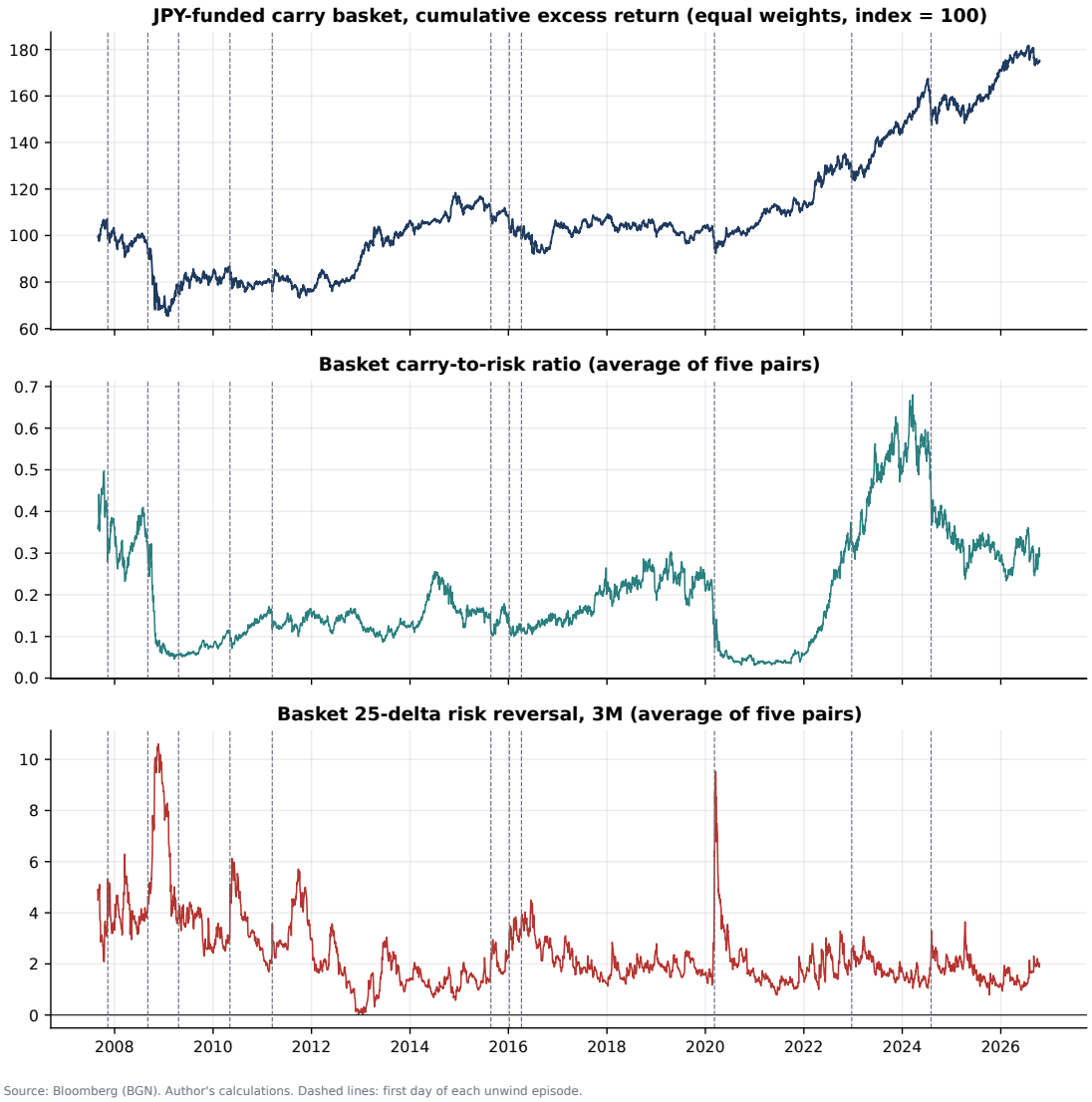
First day	Worst 5-day return, date	Worst 5-day return (%)	Days below threshold
12 November 2007	12 November 2007	−7.05	3
4 September 2008	27 October 2008	−18.36	22
20 April 2009	8 July 2009	−5.87	3
6 May 2010	20 May 2010	−7.77	6
16 March 2011	17 March 2011	−5.95	2
24 August 2015	25 August 2015	−5.81	3
8 January 2016	8 January 2016	−5.38	1
7 April 2016	30 June 2016	−5.74	5
9 March 2020	9 March 2020	−5.16	1
22 December 2022	22 December 2022	−4.51	1
2 August 2024	5 August 2024	−6.49	3

Notes: A day belongs to an episode when the five-day basket return is in the worst 1% of days (threshold −4.50%). Days less than 60 trading days apart form one episode.

The event study (Figure 7) uses all 11 episodes. On average the risk reversal is flat until a few days before the first day of the episode and rises after it, while the carry-to-risk ratio falls in the days of the initial losses. The losses that precede day 0 are partly mechanical, because the episode starts when the five-day return first crosses the threshold. With 11 episodes the figure is descriptive and is not used for inference.

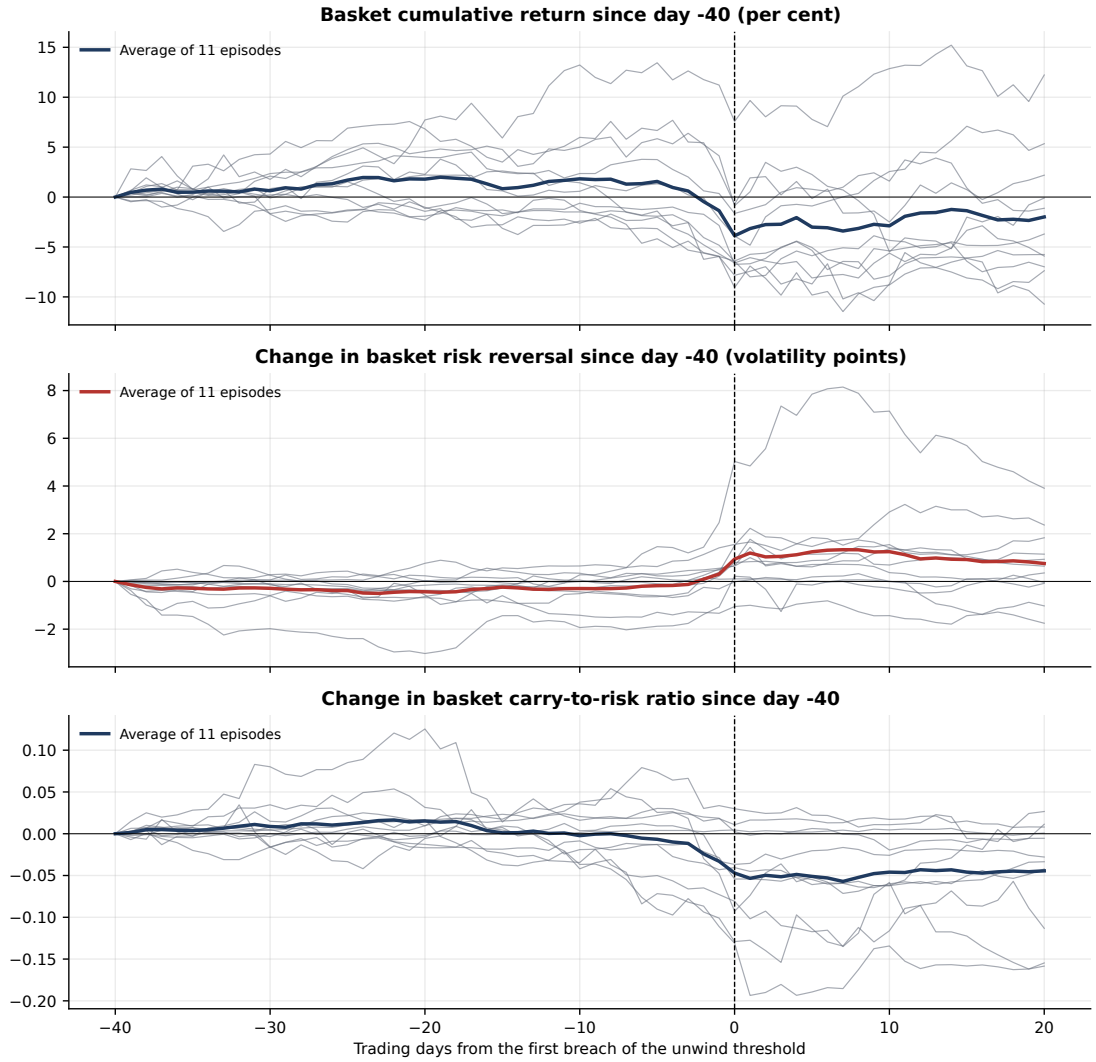
6.8 Exposure rule

The exposure rule of equation (19) reduces exposure on 13.5% of days (Table 12, Figure 8). It lowers the annualised volatility of the basket from 11.52% to 9.79% and the maximum drawdown from −32.6% to −24.5%, and raises the Sharpe ratio from 0.33 to 0.43. These gains are consistent with the Granger



Source: Bloomberg (BGN). Author's calculations. Dashed lines: first day of each unwind episode.

Figure 6: Cumulative excess return of the five-pair basket, basket carry-to-risk ratio and basket risk reversal. Dashed lines mark the first day of each unwind episode. Source: Bloomberg (BGN); author's calculations.



Source: Bloomberg (BGN). Author's calculations. Few independent episodes, the figure is descriptive.

Figure 7: Event study around the first day of each unwind episode. Grey lines: individual episodes; thick lines: averages. Source: Bloomberg (BGN); author's calculations.

results. The risk reversal rises with the first losses and stays high for some time, so the rule cuts exposure during the stress that follows the first shock. In the long sample the rule lowers the drawdown and raises the Sharpe ratio, from 0.43 to 0.47, at the cost of a lower return (Section 7).

Table 12: Risk reversal exposure rule, main sample

Strategy	Return (% p.a.)	Volatility (% p.a.)	Sharpe ratio	Max drawdown (%)	Days reduced (%)
Passive basket	3.82	11.52	0.33	−32.59	0.0
Risk reversal rule	4.21	9.79	0.43	−24.46	13.5

Notes: Exposure halved on day t when the risk reversal of day $t-1$ exceeds its 90th percentile over the previous 252 days (equation (19)). Excess returns; no transaction costs.

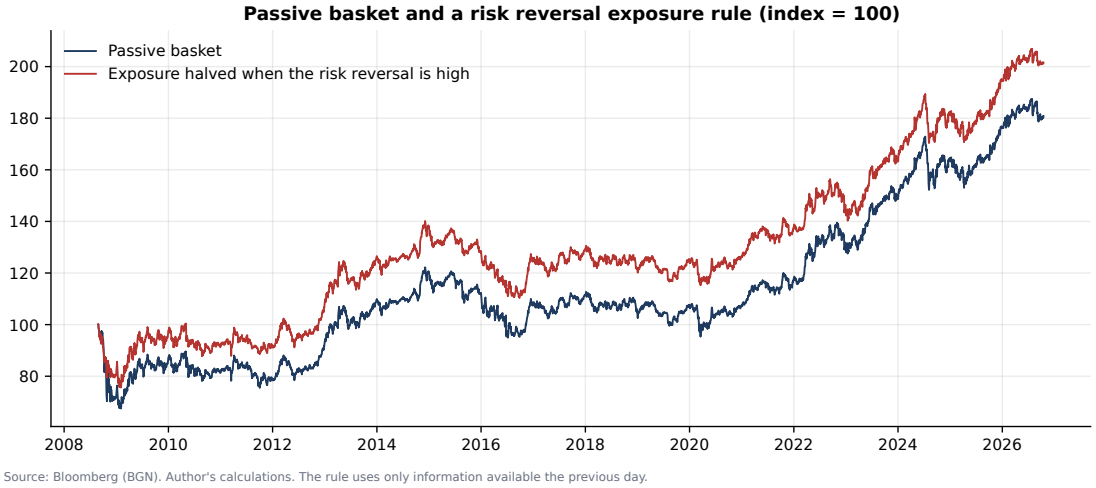


Figure 8: Passive basket and exposure rule, cumulative excess return. Source: Bloomberg (BGN); author's calculations.

7 Robustness with a Two-Pair Basket from 2004

The two-pair basket of USD/JPY and AUD/JPY covers 8 October 2004 to 16 October 2026 and includes the unwinds of March 2007 and August 2007 (Table 16). It confirms every conclusion. No coefficient of model M3 is significant at 5% (smallest p-value 0.051, Table 13). Basket returns Granger-cause both indicators with $p < 0.001$ at every lag order, while changes in the risk reversal do not Granger-cause returns (p between 0.816 and 0.950, Table 14). The probit rejects on at most one grid for each specification (Table 15). The same-day correlation between the basket return and the change in the risk reversal is −0.42, and 1 correlation at a lag in which the risk reversal moves first exceeds the band of ±0.026, $k = -1$ (+0.028). The exposure rule raises the Sharpe ratio from 0.43 to 0.47 and lowers the maximum drawdown from −32.8% to −25.6%, with a lower return (4.51% against 4.86%, Table 17).

Table 13: Model M3, two-pair basket

h	Regressor	Coefficient	z	p-value	R^2	N	L
1	$\Delta_5 CTR$	0.022	1.81	0.071	0.0023	5,738	9
1	$\Delta_5 RR$	0.036	1.95	0.051	0.0023	5,738	9
5	$\Delta_5 CTR$	0.004	0.11	0.916	0.0019	5,734	9
5	$\Delta_5 RR$	0.066	1.17	0.241	0.0019	5,734	9
21	$\Delta_5 CTR$	−0.047	−0.55	0.583	0.0003	5,718	21
21	$\Delta_5 RR$	0.001	0.01	0.995	0.0003	5,718	21

Notes: As in Table 3.

Table 14: Granger causality tests, two-pair basket

Null hypothesis	p	N	Wald (NW)	F (classical)	LM (BG)	Wald p-value, all p
ΔRR does not Granger-cause r	10	5,733	5.99 (0.816)	1.16 (0.313)	8.55 (0.128)	0.816–0.950
r does not Granger-cause ΔRR	10	5,733	88.63 (<0.001)	50.32 (<0.001)	8.03 (0.155)	<0.001
ΔCTR does not Granger-cause r	10	5,733	12.92 (0.228)	2.15 (0.018)	10.72 (0.057)	0.161–0.577
r does not Granger-cause ΔCTR	15	5,728	72.06 (<0.001)	7.80 (<0.001)	7.06 (0.216)	<0.001

Notes: As in Table 7.

Table 15: Probit for tail weeks across the five weekly grids, two-pair basket

Tail	Controls	N	Events	LR p min	LR p median	LR p max	Grids $p < 0.05$	Mean LRI
5%	none	1,147	58	0.023	0.547	0.897	1 of 5	0.005
5%	VIX	1,147	58	0.055	0.398	0.854	0 of 5	0.007
10%	none	1,147	115	0.008	0.493	0.838	1 of 5	0.004
10%	VIX	1,147	115	0.010	0.524	0.941	1 of 5	0.005

Notes: As in Table 8.

Table 16: Unwind episodes of the two-pair basket

First day	Worst 5-day return, date	Worst 5-day return (%)	Days below threshold
16 December 2005	16 December 2005	−4.55	1
5 March 2007	5 March 2007	−5.66	1
15 August 2007	16 August 2007	−7.02	7
4 September 2008	27 October 2008	−16.69	18
20 April 2009	8 July 2009	−5.77	4
6 May 2010	20 May 2010	−7.79	6
16 March 2011	17 March 2011	−5.99	2
11 June 2013	11 June 2013	−5.14	1
24 August 2015	24 August 2015	−6.28	3
8 January 2016	3 May 2016	−5.95	9
9 March 2020	9 March 2020	−5.18	2
2 August 2024	5 August 2024	−6.86	3
8 April 2025	8 April 2025	−4.73	1

Notes: As in Table 11, threshold −4.21%.

Table 17: Risk reversal exposure rule, two-pair basket

Strategy	Return (% p.a.)	Volatility (% p.a.)	Sharpe ratio	Max drawdown (%)	Days reduced (%)
Passive basket	4.86	11.21	0.43	−32.81	0.0
Risk reversal rule	4.51	9.53	0.47	−25.58	13.3

Notes: As in Table 12.

8 Research Log

The pipeline was checked against the intended design and against Greene (2003) at several stages. Each check that changed the study is reported here in the order in which it occurred.

1. **First exploratory version.** A first script used daily returns that combined log changes of the spot rate with simple carry accruals, the level of the carry-to-risk ratio as a regressor, a five-pair basket from March 2005, White standard errors for non-overlapping regressions and a probit on a single weekly grid. Returns were changed to the simple excess returns of equation (7), which compound exactly into the basket index.

2. **Unit roots.** The Dickey-Fuller test did not reject a unit root for the level of the carry-to-risk ratio, which violates the stationarity required by Theorem 12.1. The level was replaced by its five-day change, which rejects the unit root strongly.

3. **Robustness of the probit.** On the first weekly grid the probit gave $p = 0.003$ for the joint test of the slopes. On the other grids the p-value reached 0.68. Every probit is now estimated on all five grids. The criterion that a result must hold across grids was set after this sensitivity was observed, and the full set of grid-level results is reported in the repository.

4. **Residual autocorrelation.** Non-overlapping monthly residuals failed the Breusch-Godfrey test, so White standard errors were replaced by Newey-West standard errors. The Granger equations, which contain lagged dependent variables, showed autocorrelated residuals at some lag orders. Following Greene (p. 266), the lag order is now chosen so that the residuals pass the test.

5. **Gaps in the data.** A chart of fitted probabilities showed straight segments in 2006 and 2007. The cause was months of missing option data for NZD/JPY and CAD/JPY, across which rolling windows had been computed. The sample rule of Section 3 removed the problem and moved the start of the main sample to 30 August 2007. The two-pair basket was added to recover the earlier period.

6. **Text against results.** After the change of sample, a statement that no pair-level carry-to-risk ratio rejects a unit root was no longer true for AUD/JPY. The automatic check that builds this report stopped on the discrepancy, and the text was corrected. The same check found a rounding error in a z ratio quoted in the repository, which was corrected.

7. **Formulas.** The test suite of the repository checks that the Newey-West covariance matrix equals equations (10) and (11) computed by hand to ten decimal places, that the Breusch-Godfrey, Ljung-Box and ARCH statistics equal their textbook definitions, that covered interest parity gives the correct sign for both quoting conventions, that h-day sums are aligned with $r_{t+1}, \ldots, r_{t+h}$, that the exposure rule uses only past information and that the probit fit measures match eq. (21-28) and McFadden's index.

8. **Data.** Risk reversal and interest differentials were compared with the Bloomberg terminal screens on 16 October 2026 (Section 3).

Every number in this report is produced by a script that reads the result files of the repository, so no value is transcribed by hand.

9 Conclusions

In 2004–2026 the carry-to-risk ratio and the risk reversal do not anticipate the losses of yen-funded carry trades. Both react to them. Lagged basket returns predict the changes in the two indicators at every lag order and in both samples, while lagged changes in the risk reversal never predict returns. Predictive regressions at horizons from one day to one month, pair-level regressions and a probit model for tail weeks give no robust evidence of anticipation. The result confirms on a new sample the evidence that risk reversals follow carry trade losses (Brunnermeier et al. 2009; Jurek 2014) and the observation of Gagnon and Chaboud (2007) reported by Galati et al. (2007), and extends it to baskets of yen crosses, two decades and formal tests.

The indicators remain useful. The risk reversal measures how much the market pays for protection once an unwind has started and stays high for weeks after the first losses. A rule that halves exposure when it is high lowers volatility and drawdowns. The level of the carry-to-risk ratio describes the attractiveness of the trade and its build-up over policy cycles. Its peaks of June 2007 and November 2023 came two months before the unwind of August 2007 and nine months before that of August 2024, a distance that varies too much for a timing signal.

The study has limits. The number of independent unwind episodes is small (11 in the main sample). Forward-implied differentials include the cross-currency basis. The risk reversal and the carry-to-risk ratio are measured at the close, and daily data cannot separate moves that occur within the same day. The tests do not cover volatility-based predictors, such as the average implied volatility of the basket or the volatility risk premium, which the literature finds informative (Cenedese et al. 2014; Bakshi and Panayotov 2013; Della Corte et al. 2016). Three extensions would address these limits. The first is to test those predictors, including with quantile regressions. The second is positioning data from the Commitments of Traders of the Commodity Futures Trading Commission, which Galati, Heath and McGuire use as a measure of carry trade build-up. The third is intraday option quotes around the unwinds of August 2007 and August 2024.

A Data Appendix

Table 18: Bloomberg series

Bloomberg ticker	Description	First	Last	Obs.
USDJPY BGN Curncy	USD/JPY spot	1990-01-01	2026-10-16	9,600
AUDJPY BGN Curncy	AUD/JPY spot	1990-01-01	2026-10-16	9,598
NZDJPY BGN Curncy	NZD/JPY spot	1990-01-01	2026-10-16	9,599
GBPJPY BGN Curncy	GBP/JPY spot	1990-01-01	2026-10-16	9,599
CADJPY BGN Curncy	CAD/JPY spot	1990-01-01	2026-10-16	9,599
USDJPYV3M BGN Curncy	USD/JPY 3M ATM implied volatility	1995-12-05	2026-10-16	8,011
AUDJPYV3M BGN Curncy	AUD/JPY 3M ATM implied volatility	1995-12-05	2026-10-16	7,969
NZDJPYV3M BGN Curncy	NZD/JPY 3M ATM implied volatility	1997-08-14	2026-10-16	6,223
GBPJPYV3M BGN Curncy	GBP/JPY 3M ATM implied volatility	2005-03-08	2026-10-16	5,606
CADJPYV3M BGN Curncy	CAD/JPY 3M ATM implied volatility	2005-03-08	2026-10-16	5,243
USDJPY25R3M BGN Curncy	USD/JPY 25-delta 3M risk reversal	2003-10-02	2026-10-16	6,008
AUDJPY25R3M BGN Curncy	AUD/JPY 25-delta 3M risk reversal	2003-10-02	2026-10-16	5,861
NZDJPY25R3M BGN Curncy	NZD/JPY 25-delta 3M risk reversal	2005-03-08	2026-10-16	5,285
GBPJPY25R3M BGN Curncy	GBP/JPY 25-delta 3M risk reversal	2005-03-08	2026-10-16	5,613
CADJPY25R3M BGN Curncy	CAD/JPY 25-delta 3M risk reversal	2005-03-08	2026-10-16	5,236
JPY3M BGN Curncy	USD/JPY 3M forward points (scale 1/100)	1990-01-02	2026-10-16	9,562
AUD3M BGN Curncy	AUD/USD 3M forward points (scale 1/10000)	1990-01-02	2026-10-16	9,394
NZD3M BGN Curncy	NZD/USD 3M forward points (scale 1/10000)	1990-01-02	2026-10-16	9,549
GBP3M BGN Curncy	GBP/USD 3M forward points (scale 1/10000)	1990-01-02	2026-10-16	9,582
CAD3M BGN Curncy	USD/CAD 3M forward points (scale 1/10000)	1990-01-02	2026-10-16	9,552
VIX Index	CBOE VIX	1990-01-02	2026-10-16	9,296
FDTR Index	Fed funds target rate (upper bound from Dec 2008)	1989-12-29	2026-10-15	9,501
BOJDTR Index	Bank of Japan policy rate	1989-12-29	2026-10-16	6,006
MUTKCALM Index	Japan uncollateralised overnight call rate	1992-10-09	2026-10-15	8,345
JPMVXYG7 Index	JPMorgan G7 FX volatility index	1992-06-01	2026-10-16	8,964
USDJPYV1M BGN Curncy	USD/JPY 1M ATM implied volatility	1995-12-04	2026-10-16	8,014
USDJPY25R1M BGN Curncy	USD/JPY 25-delta 1M risk reversal	2003-10-02	2026-10-16	5,983

Notes: Daily PX_LAST. First and last dates and number of observations of the downloaded history. Forward points are divided by 100 for USD/JPY and by 10,000 for the other pairs. The last four series are optional and are not used in the main results.

B Probit Coefficients

Table 19: Probit coefficients, first weekly grid

Specification	Variable	Coefficient	z (Hessian)	z (sandwich)
5%, controls: none	Constant	−1.661	−24.28	−24.07
5%, controls: none	$\Delta_5 CTR$	−0.063	−0.92	−0.79
5%, controls: none	$\Delta_5 RR$	0.106	1.67	1.44
5%, controls: VIX	Constant	−1.663	−24.25	−24.16
5%, controls: VIX	$\Delta_5 CTR$	−0.058	−0.84	−0.73
5%, controls: VIX	$\Delta_5 RR$	0.084	1.19	0.96
5%, controls: VIX	$\Delta_5 VIX$	0.049	0.70	0.61
10%, controls: none	Constant	−1.286	−23.58	−23.50
10%, controls: none	$\Delta_5 CTR$	−0.050	−0.86	−0.80
10%, controls: none	$\Delta_5 RR$	0.072	1.30	1.15
10%, controls: VIX	Constant	−1.287	−23.57	−23.43
10%, controls: VIX	$\Delta_5 CTR$	−0.055	−0.95	−0.87
10%, controls: VIX	$\Delta_5 RR$	0.100	1.62	1.39
10%, controls: VIX	$\Delta_5 VIX$	−0.063	−1.05	−0.95

Notes: Maximum likelihood estimates. z ratios from the inverse of the negative Hessian and from White's sandwich estimator (Greene, Section 21.4.1).

C Correspondence with Greene (2003)

Table 20: Methods and their source in Greene, *Econometric Analysis*, fifth edition

Step	Implementation	Greene
Stationarity and ergodicity	Requirement for least squares asymptotics with time series	Section 12.4, Theorem 12.1
Unit root tests	Dickey-Fuller regression with a constant, augmented with lagged differences	Section 20.3.4, Table 20.4
Differencing	Levels that do not reject a unit root enter in differences	Section 20.3.1
Covariance of least squares	Newey-West, Bartlett weights, $L = \max(h, \lceil T^{1/4} \rceil)$	Section 12.5.2, eq. (12-17), p. 267
Heteroscedasticity	White estimator, first version only	Section 11.2.3
Autocorrelation test	Breusch-Godfrey LM, zero-filled lags	Section 12.7.1, eq. (12-20), p. 269
Autocorrelation test	Ljung-Box Q'	Section 12.7.2, eq. (12-22)
Lagged dependent variable	Inconsistency of least squares with autocorrelation; validity of LM test	p. 266; Section 12.7.4
ARCH effects	LM test, TR^2	Section 11.8.4
Granger causality	Single-equation Wald test, robust covariance	Section 19.6.5, p. 592–593
Binary choice	Probit index function model, maximum likelihood	Sections 21.3.2, 21.4
Covariance of the probit	Hessian; sandwich as a sensitivity check	p. 673; Section 21.4.1
Marginal effects	At the means, delta method	Section 21.4.2
Tests and fit	LR, Wald, likelihood ratio index	Sections 21.4.3, 21.4.5; eqs. (21-27), (21-28)

D Reproducibility

The repository (Mazzarini 2026) contains the code (Python, with pandas, statsmodels and SciPy), the test suite and every table and figure of this report. With the Bloomberg exports listed in Table 18 in one folder, the commands `python run_analysis.py –data-dir <folder>` and `python -m pytest` reproduce the results and run the checks against the formulas of Greene (2003). All parameters of the study (horizons, lag rules, thresholds, windows) are set in a single configuration file before estimation.

References

Aquilina, Matteo, Marco Lombardi, Andreas Schrimpf, and Vladyslav Sushko (Aug. 2024). *The market turbulence and carry trade unwind of August 2024*. BIS Bulletin 90. Bank for International Settlements.

Bakshi, Gurdip and George Panayotov (2013). "Predictability of currency carry trades and asset pricing implications". In: *Journal of Financial Economics* 110.1, pp. 139–163.

Banerjee, Ryan, Lena Boneva, Gabor Pinter, and Vladyslav Sushko (May 2026). *Monetary policy transmission to exchange rates: the role of currency carry trades*. BIS Bulletin 124. Bank for International Settlements.

Bhansali, Vineer (2007). "Volatility and the carry trade". In: *Journal of Fixed Income* 17.3, pp. 72–84.

Breusch, Trevor S. (1978). "Testing for autocorrelation in dynamic linear models". In: *Australian Economic Papers* 17.31, pp. 334–355.

Brunnermeier, Markus K., Stefan Nagel, and Lasse H. Pedersen (2009). "Carry trades and currency crashes". In: *NBER Macroeconomics Annual 2008*. Vol. 23. University of Chicago Press, pp. 313–347.

Burnside, Craig, Martin Eichenbaum, Isaac Kleshchelski, and Sergio Rebelo (2006). *The returns to currency speculation*. NBER Working Paper 12489. National Bureau of Economic Research.

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

Clarida, Richard, Josh Davis, and Niels Pedersen (2009). *Currency carry trade regimes: beyond the Fama regression*. NBER Working Paper 15523. National Bureau of Economic Research.

Costa Filho, Adonias Evaristo da (2016). "The information content of risk reversals in emerging market currencies". In: *Revista Brasileira de Finanças* 14.3, pp. 403–441.

Della Corte, Pasquale, Tarun Ramadorai, and Lucio Sarno (2016). "Volatility risk premia and exchange rate predictability". In: *Journal of Financial Economics* 120.1, pp. 21–40.

Dickey, David A. and Wayne A. Fuller (1979). "Distribution of the estimators for autoregressive time series with a unit root". In: *Journal of the American Statistical Association* 74.366, pp. 427–431.

Engle, Robert F. (1982). "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation". In: *Econometrica* 50.4, pp. 987–1007.

Gagnon, Joseph E. and Alain P. Chaboud (July 2007). *What can the data tell us about carry trades in Japanese yen?* International Finance Discussion Papers 899. Board of Governors of the Federal Reserve System.

Galati, Gabriele, Alexandra Heath, and Patrick McGuire (Sept. 2007). "Evidence of carry trade activity". In: *BIS Quarterly Review*, pp. 27–41.

Godfrey, Leslie G. (1978). "Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables". In: *Econometrica* 46.6, pp. 1293–1301.

Granger, Clive W. J. (1969). "Investigating causal relations by econometric models and cross-spectral methods". In: *Econometrica* 37.3, pp. 424–438.

Greene, William H. (2003). *Econometric Analysis*. Fifth. Upper Saddle River, NJ: Prentice Hall.

Guyot, Opale, Heather A. Montgomery, and Peiqing Yang (2026). "Risk premiums, market volatility, and exchange rate dynamics: evidence from the yen carry trade". In: *Risks* 14.3, p. 46.

Hutchison, Michael and Vladyslav Sushko (2013). "Impact of macro-economic surprises on carry trade activity". In: *Journal of Banking & Finance* 37.4, pp. 1133–1147.

Jurek, Jakub W. (2014). "Crash-neutral currency carry trades". In: *Journal of Financial Economics* 113.3, pp. 325–347.

Kütük, Merve Mavuş and Sweder van Wijnbergen (2025). *Carry trade and currency crash risk*. Tinbergen Institute Discussion Paper 2025-058/IV. Tinbergen Institute.

Ljung, Greta M. and George E. P. Box (1978). "On a measure of lack of fit in time series models". In: *Biometrika* 65.2, pp. 297–303.

MacKinnon, James G. (2010). *Critical values for cointegration tests*. Queen's Economics Department Working Paper 1227. Queen's University, Department of Economics.

Mazzarini, Alessandro (2026). *carry-unwind-signals: Code, Tests and Results*. GitHub repository. URL: https://github.com/AleMazza04-NoScroll/carry-unwind-signals.
McFadden, Daniel (1974). "Conditional logit analysis of qualitative choice behavior". In: *Frontiers in Econometrics*. Ed. by Paul Zarembka. New York: Academic Press, pp. 105–142.
Melvin, Michael and Duncan Shand (2017). "When carry goes bad: the magnitude, causes, and duration of currency carry unwinds". In: *Financial Analysts Journal* 73.1, pp. 121–144.
Menkhoff, Lukas, Lucio Sarno, Maik Schmeling, and Andreas Schrimpf (2012). "Carry trades and global foreign exchange volatility". In: *Journal of Finance* 67.2, pp. 681–718.
Newey, Whitney K. and Kenneth D. West (1987). "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix". In: *Econometrica* 55.3, pp. 703–708.
Nirei, Makoto and Vladyslav Sushko (2011). "Jumps in foreign exchange rates and stochastic unwinding of carry trades". In: *International Review of Economics & Finance* 20.1, pp. 110–127.
White, Halbert (1980). "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity". In: *Econometrica* 48.4, pp. 817–838.
— (1982). "Maximum likelihood estimation of misspecified models". In: *Econometrica* 50.1, pp. 1–25.