Systematic investment strategies for sovereign fixed income portfolios

Mike McMorrow[1]

Abstract

This paper constructs and analyses active systematic investment strategies for sovereign fixed income investors based on two signals: carry and term premium. Using generic, fitted zero coupon yields across the five SDR yield curves, we identify sets of monthly time-varying portfolio weights based on each of our signals and an equally weighted combination of the two. The baseline portfolio construction setup is tailored to investors with a relatively low risk tolerance, such as foreign exchange reserve managers, by modelling all assets on an FX-hedged basis and requiring neutral duration at the portfolio level. The backtesting exercise is conducted with realised returns from actual securities, accounts for transaction costs, and uses a systematic rebalancing rule. The results suggest that strategies based on the term premium signal can deliver appealing excess returns, while strategies based on carry are more questionable in terms of return, at least over the limited sample used for this paper. Combining the signals, however, may lead to some diversification benefits.

JEL classification: G10, G11, G12.

1 Bank for International Settlements. The author (mike.mcmorrow@bis.org) is grateful to Ingo Fender for helpful comments on this draft and to various members of the Asset Management and Advisory team at the BIS for many suggestions on earlier iterations of this work. He would also like to recognise the late Vahe Sahakyan for his insights, contributions and warm collaboration. Discussions with participants of the Sixth Public Investors Conference in Ottawa are also gratefully acknowledged. Any remaining errors are the responsibility of the author. The views expressed are those of the author and do not necessarily reflect those of the BIS.

1. Introduction

This paper explores systematic tactical investment strategies for sovereign fixed income portfolios and takes an in-depth look at implementation issues. Systematic investment strategies are part of a rich set of research on empirical factors dating back to the 1980s. This paper contributes to this expansive literature by focusing on sovereign fixed income strategies designed for investors with limited risk tolerance, such as foreign exchange reserve managers, and utilises two signals that are widely followed in the fixed income community: carry and term premium.[2]

The baseline setup reflects an environment in which investors such as reserve managers are likely to operate.[3] Due to the usual reserve management objectives of safety, liquidity and return, such investors may be more comfortable with interest rate risk than with currency risk, and may even be more comfortable with curve positions rather than outright bets on duration. They may also be constrained in their ability to trade certain derivatives for speculative purposes (eg futures), and/or to go outright short a security, ie to sell beyond what is held in the benchmark.

With this environment in mind, we construct portfolio strategies using carry, term premium and an equally weighted combination of the two by finding optimal portfolio weights using a set of generic zero coupon securities across the five SDR currencies. We allow durations from three months to eight years, and constrain the weights to sum to zero, giving rise to a set of fully funded overweight and underweight positions that exploit cross-market and cross-curve dynamics. We design a hypothetical sovereign SDR benchmark to identify constraints on underweight positions at a given yield curve point, given the inability to go short. We model our generic fixed income instruments on a hedged basis, thereby removing currency risk from the strategies, and constrain the optimisation problem such that overall portfolio duration is zero. Our "baseline" setup will thus consist of a set of cross-market, cross-curve positions that incur no net FX or duration risk at the portfolio level.

We build a time series of monthly optimal portfolio weights beginning at the end of September 2016 – which represents the first end-month observation in which the CNY was a constituent of the SDR – through March 2023.[4] We re-optimise on a monthly basis, thereby allowing evolutions in the macro and market environment to be reflected in portfolio weights at a frequency appropriate for tactical asset allocation.

We backtest our strategies with a paper portfolio that uses returns from actual securities and accounts for transaction costs. While futures could be an ideal outlet for implementing such strategies, we avoid any reliance on derivatives given that certain types of investors may be limited in their ability to use such instruments – be it an outright prohibition or a restriction for hedging purposes only. We experiment with different rebalancing rules to account for the trade-off between rebalancing precision and transaction costs, specifically between minimising the duration drift at

2 Carry is defined as the return on a fixed-income security when the yield remains unchanged. The term premium is the excess yield on a term fixed income security over and above the average short rate expected over the life of the security.

3 See Fender et al(2022) for a description of a conventional reserve management framework.

4 See www.imf.org/en/News/Articles/2016/09/29/AM16-NA093016IMF-Adds-Chinese-Renminbi-to-Special-Drawing-Rights-Basket.

the portfolio level due to imperfect rebalancing, and the higher transaction costs that result from frequently switching in and out of securities.

Finally, we experiment with portfolio construction setups that are somewhat less constrained, and may therefore be appropriate for more risk-tolerant investors. Specifically, we explore a version that allows for modestly positive or negative duration positions at the overall portfolio level. In another formulation, we allow FX risk from exposure to CNY, due to the fact that, among other reasons, hedging the CNY gives rise to volatility in portfolio weights over time, thereby generating heavier transaction costs.

We find that the baseline term premium strategy generates encouraging risk-adjusted returns over our sample, while the baseline carry strategy generates slightly negative returns. Combining the two signals can give rise to some diversification benefits, although over the whole sample a combined strategy underperforms one based on the term premium alone. We find that leaving the CNY unhedged generates superior risk-adjusted returns to the fully hedged setup, although with greater maximum drawdown and with minimal diversification between the signals. Finally, we find little evidence that allowing flexible portfolio duration leads to better risk-adjusted performance.

The broader literature on factor-based investing is expansive in the dimensions of both factor types and asset classes. Most of the focus has been on the equity market, where factors such as price-to-book ratios and measures of the equity risk premium have been explored. Exploration of the fixed income market has been a more recent phenomenon, with contributions such as Asness et al (2013) assessing factors such as momentum and value. Value is of particular interest, especially given the lack of clarity of how to define it in a fixed income context. For example, Asness et al (2013) use the five-year change in the yield on 10-year bonds. We find the term premium a more conceptually appealing measure, given that its definition – the premium that can be harvested by holding a term fixed income security in excess of what would be earned through rolling over very short-term assets – contains a forward-looking element. Carry has also been explored extensively in the literature. In addition to Coche et al (2018), Koijen et al (2018), Ahmerkamp and Grant (2013) and Baz et al (2015) all show that carry can be predictive of excess returns in various contexts, including across different asset classes, and may be combined with other signals to exploit diversification benefits.

This paper is closely related to Bjorheim et al (2018) and Coche et al (2018) – both of which were featured at the Sixth Public Investors Conference – and extends their results by offering answers to several portfolio design and implementation questions.

Bjorheim et al (2018) employ a novel macro-based process for modelling yield curves to extract term premium estimates. Specifically, they devise a shadow rate model (see, for example, Lombardi and Zhu (2018)) based on a modified Nelson-Siegel equation to model yield curves at their effective lower bounds – a key feature of monetary policy evolution since the Great Financial Crisis (GFC). Their shadow rate approach allows for the continual use of a modified Taylor Rule, in which macroeconomic variables, such as the output gap and inflation expectations, condition the evolution of the short rate (the first yield curve factor in their approach). This evolution is used to identify term premium estimates, which are subsequently used as signals to predict excess returns across the four currencies that constituted the SDR prior to October 2016. They find that, on an out-of-sample basis, the term premium outperforms expected returns and carry as a predictive signal, whether one assumes perfect foresight of macroeconomic developments or employs a "mean-reverting macro" forecasting approach. Their results showing the superior excess return predictability of the term premium signal are consistent with those found in this study, and their use of a more extended sample – spanning decades rather than years – may help to reassure that the results presented herein are not the spurious product of a limited performance history.

Coche et al (2018) conduct a similar, detailed study of the excess return predictability of a carry signal. Using a history beginning in the 1970s, they backtest strategies from three perspectives: cross-curve, where positions are permitted within a given yield curve; cross-market, where positions at a specific maturity point are taken across yield curves; and cross-curve, cross-market, where flexibility is provided across yield curves and duration points. They find that the cross-market flexibility is key to generating meaningful excess returns. They further find evidence that excess returns may be time-varying and regime-dependent. We make use of the carry signal with these findings in mind.

The remaining sections of this paper are organised as follows:

2. **Identifying signals:** estimating carry and term premium.
3. **Building portfolios**: identifying weights based on the signals.
4. **Evaluating with a paper portfolio:** simulating a realistic setting to measure historical performance.
5. **Assessing performance:** how did the strategies fare historically?
6. **Considering alternative setups:** do changes in the portfolio construction setup lead to meaningfully different results?

We then conclude with a few reflections going forward.

2. Identifying signals

2.1 Fitting yield curves

To identify signals, we used fitted yields obtained from a modified Nelson-Siegel (NS) model, based on Nelson and Siegel (1987), Nyholm (2015) and Bjorheim et al (2018), and used by Coche et al (2018) and Fender et al (2020), among others. The modelling framework relies on shadow yield curve factors and allows us to retain the linkage between macroeconomic variables and yield curve factors at the effective lower bound. Once projected forward, shadow rates are subsequently transformed into projected yield distributions that – due to the design of the shadow rate procedure – are truncated at the effective lower bound as observed historically.

We rely on spliced data histories for several of the SDR yield curves, allowing us to estimate yield curve factors with long data histories. These long histories will help add richness to the forward-looking simulations that will be used to estimate the distributions that we will associate with our signals (explained in more detail in section 3). Specific data sources and their relevant histories are displayed in Table 1.

Data sources for yield curve estimation — Table 1

Yield curve	Source	Period of use
US	Federal Reserve Board (H.15)	03/1953 - 04/1989
	Bloomberg	05/1989 - 02/2023
Germany	Bundesbank	08/1974 - 12/1994
	Bloomberg	01/1995 - 02/2023
UK	Bank of England	01/1970 - 12/1994
	Bloomberg	01/1995 - 02/2023
Japan	Japan Ministry of Finance	09/1974 - 03/1989
	Bloomberg	04/1989 - 02/2023
China	Bloomberg	03/2003 - 02/2023

Source: Author.

2.2 Estimating carry

Carry is calculated as the return an investor receives over holding period k in an unchanged yield curve environment. Formally, it is described as follows, where P_t is the spot price of a sovereign bond at time t, $F_{t,t+1}$ is the price at time t of a futures contract expiring in period $t+1$, and X_t is the amount of capital that finances the investment into the futures contract.

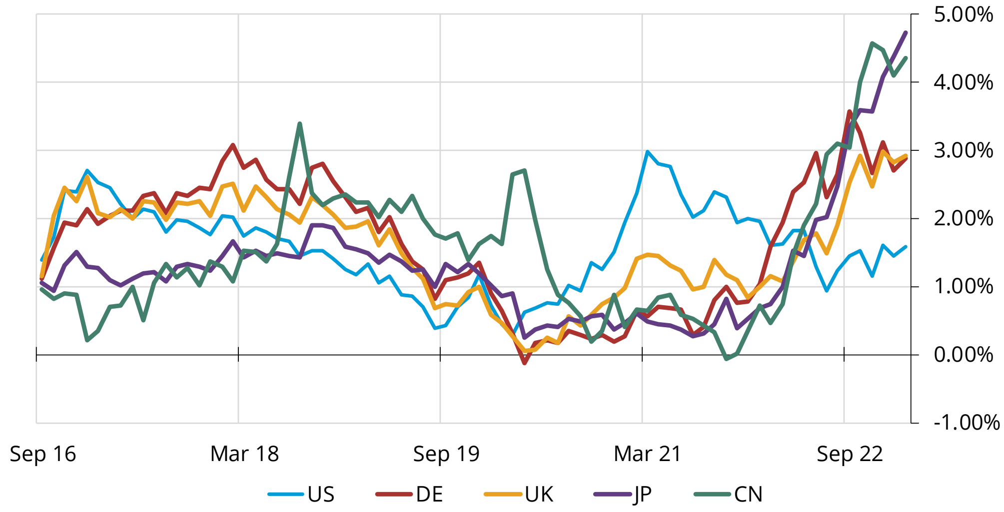
$$r^{C}_{t\to t+1} = \frac{P_t - F_{t,t+1}}{X_t} \quad (1)$$

Note that this description is implicitly in local currency terms. Given that our strategies will involve investment into multiple yield curves, we incorporate an FX dimension to account for hedging. Our carry signal, in final form, can be described with $F^{f}_{t,t+1}$, an FX forward contract on the foreign currency expiring at $t+1$, and S^{f}_{t}, the spot FX rate. Both exchange rates are expressed as the domestic currency per unit of foreign currency.

$$CarrySig^{j}_{t} = (1 + r^{C}_{t\to t+1})\left(\frac{F^{f}_{t,t+1}}{S^{f}_{t}}\right) - 1 \quad (2)$$

We apply these formulas to fitted zero coupon yields to construct our metrics of carry. Given that covered interest parity relationships embedded in FX forward pricing will generally offset short rate differentials in local currency terms, our carry signals will essentially capture the *relative slope and curvature* of yield curves, rather than differences in the levels. Nonetheless, any deviations from covered interest rate parity will also be manifest in the signals. Graph 1 displays the evolution of the carry signal on a 10-year government bond for the five SDR yield curves – the United States (“US”), Germany (“DE”, our proxy for the euro area), the United Kingdom (“UK”), Japan (“JP”) and China (“CN”) – at a monthly frequency over our sample.

10-year carry signals across SDR yield curves | Graph 1
Percentage



Sources: Bloomberg, author's calculations.

2.2 Estimating term premium

The term premium is the excess yield an investor expects to earn on a fixed income security (with no credit risk) of maturity τ over and above the average short rate (SR) expected to prevail over the life of the instrument.

$$TP_t(\tau) = y_t(\tau) - \frac{1}{\tau}\sum_{i=1}^{\tau} \mathrm{E}(SR_i) \qquad (3)$$

Key to measuring the term premium is the estimation of the expected short rate path. With a few extensions, we employ the macro-based approach described in Bjorheim et al (2018), where the first yield curve factor (the shadow short rate) is projected with an autoregressive term along with the output gap and inflation expectations as exogenous variables. Note that the incorporation of the shadow rate approach allows for a coherent estimation with macroeconomic variables at the effective lower bound.

$$\beta_{1,t} = \alpha + \theta\beta_{1,t-1} + \gamma OG_t + \delta\pi_t^e + \varepsilon_t \qquad (4)$$

The output gap is calculated as the ratio of GDP to potential GDP, where potential GDP growth is computed recursively as an exponentially smoothed average of the previous period's realised GDP growth rate and the estimated rate of potential growth; the previous period's estimated output gap is also allowed to influence its evolution.

We extend the approach in a few ways. First, we utilise the filtering approach put forth in Stock and Watson (2007) that incorporates unobserved components with stochastic volatility (UCSV) to estimate inflation expectations. Given that central banks are widely believed to respond to persistent movements in inflation – rather than those perceived to be temporary – this unobserved trend component can help strip out unwanted noise.[5]

We further extend the model setup by using Consensus Forecasts from Consensus Economics as assumptions for future GDP growth rates and CPI inflation. These assumptions are fed into the equations for projecting the output gap and inflation expectations, thereby giving rise to the explanatory variables featured in equation (4).[6]

We estimate Taylor Rules individually for each yield curve with maximum likelihood on the available data histories.[7] Given the limited experience with macroeconomic performance in China, especially under the current – and somewhat more flexible – exchange rate regime, as well as the limited availability of seasonally adjusted macroeconomic data, we treat the projection of the CNY short rate differently. Instead of estimation with a Taylor Rule, we simply take the path implied by mean economist forecasts of the short rate as reported by Bloomberg.[8]

A key step in the shadow rate approach of Bjorheim et al (2018) is the transformation of shadow factors back to yields, with yields constrained by their effective lower bounds. The short rate path used in the calculation of the term premium will embody this feature, ensuring that it never falls below the effective lower bound (which is either 0% or the lowest negative yield observed in the historical data for a given yield curve).

Given that we will be using the term premium signal in a multicurrency, hedged context, we incorporate a foreign exchange component into the signal. Given that the term premium represents the yield *in excess* of the expected short rate path, we judge that the most relevant return is the excess hedged FX return, ie the return harvested when hedging a foreign currency investment with FX forwards above and beyond that expected by covered interest rate parity. Formally, this can be described with $F_{t,t+k}^{f}$, a forward contract on the foreign (ie non-numeraire) currency of expiry $t+k$, S_t^f, the spot FX rate, $y_t^d(k)$, an annualised short-term rate with maturity k on the domestic yield curve, and $y_t^f(k)$, an annualised short-term rate with maturity k on the foreign yield curve. The annualised hedged excess return from the perspective of the domestic currency can be described as follows.

5 Rather than conducting the filtering exercise on each simulation, we take a more computationally efficient approach by first conducting the UCSV procedure on the historical data sample, and then regressing the stochastic trend component on realised inflation. The parameters from this regression are then used to forecast inflation expectations using Consensus Forecasts for headline inflation.

6 As in Bjorheim et al (2018), we transform quarterly GDP growth into a monthly figure by applying the pattern of monthly growth in industrial production to quarterly GDP, while ensuring the average monthly GDP growth rate equals the originally reported quarterly GDP growth rate.

7 When doing so, we also periodically adjust the intercept term to force the long-term equilibrium rate implied by the model to be consistent with long-term expectations for the short rate as reported by Consensus Forecasts.

8 Bloomberg consensus forecasts are available for SHIBOR, the interbank rate in Shanghai. Given that we are interested in a projection of the three-month government bond yield, forecasts are adjusted with the most recently observed difference between SHIBOR and the fitted three-month government bond yield.

$$r^d_{hFX,t} = \frac{1}{k}\left(\frac{F^f_{t,k}}{S^f_t} - 1\right) - \left(\frac{1 + y^d_t(k)}{1 + y^f_t(k)} - 1\right) \quad (5)$$

We note that this expression does not precisely yield the FX basis as conventionally defined, given that the interest rate differential is calculated using government yields rather than LIBOR rates (off of which FX forwards have traditionally been priced). As a result, this will be an imperfect hedge in that it results in exposure to Libor-government spreads.

We project this excess hedged FX return over the life of the relevant instrument with a simple AR(1) process.

$$r^d_{hFX,t} = \varphi + \psi r^d_{hFX,t-1} + \varepsilon_t \quad (6)$$

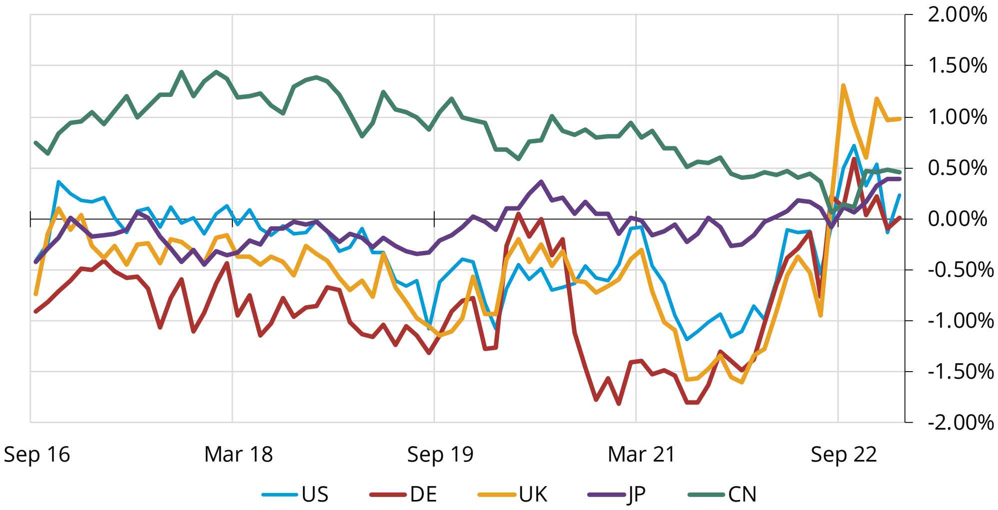
With this FX component, our term premium signal, $TPSig^j_t(\tau)$ at time t for country j and maturity τ is expressed as follows:

$$TPSig^j_t(\tau) = y_t(\tau) - \frac{1}{\tau}\sum_{i=1}^{\tau} \mathrm{E}(SR_i) + \frac{1}{\tau}\sum_{i=1}^{\tau} \mathrm{E}(r^d_{hFX,i}) \quad (7)$$

With this enhancement, we have greater confidence that any deviations from covered interest rate parity (CIRP) are captured by our signal, thereby avoiding – in expectation – bias from the hedging activity. Graph 2 shows the evolution of the 10-year term premium signal in each of the five SDR yield curves over the sample.[9]

10-year term premium signals across SDR yield curves — Graph 2

Percentage



Sources: Bloomberg, author's calculations.

[9] The 10-year term premium in the US is positively correlated with other estimates, including those put forth by Kim and Wright (2005) and Adrian et al (2013).

3. Building portfolios

3.1 Designing a strategy

Equipped with our signals, we design portfolio strategies tailored to our hypothetical reserve manager with a low risk appetite. Given that signals, in this application, are intended to support active management in a tactical asset allocation framework, we construct our portfolios as a set of overweight and underweight positions versus a hypothetical benchmark that sum to zero by construction. The strategies are reoptimised on a monthly basis to ensure that market movements and changes in the outlook are routinely captured.

In defining our universe of eligible assets, we consider five generic securities representing yield curve points across each of the five SDR currencies – specifically, the three-month, two-year, four-year, six-year and eight-year zero coupon yields. For the JPY curve, we restrict the eligible universe to the three-month point in the light of the Bank of Japan's policy of yield curve control over the sample. For EUR-denominated assets, we restrict our interest to the German yield curve, given that the strategy is focused on exploiting signals relating to interest rate risk, not sovereign credit risk. In total, this results in an eligible universe of 21 assets. Each generic instrument is permitted a maximum weight of 50%.

We assume our hypothetical reserve manager holds the five SDR currencies in proportion to their SDR weights.[10] Furthermore, we presume they manage each currency portfolio to a benchmark as determined by the maturity distributions in relevant sovereign bond indices from ICE out to five years – generally equating to currency durations between two and three years (JPY excepted). We further assume that our reserve manager cannot go outright short any sector, and is thus constrained by the weights implied in this benchmark. Table 2 shows the resulting lower-bound constraints that will be applied to our portfolio optimisation problem as implied by the SDR weights and index compositions at the start of our sample (end-September 2016). Given that we expect our reserve manager to hold each currency portfolio for policy purposes (with a small budget for deviation for active management), we presume they select the SDR as the numeraire, thereby viewing currency risk as anything that deviates from these weights.

We design a "baseline" setup that reflects this risk aversion: generally greater comfort with interest rate risk than currency risk, and generally greater comfort with curve risk than with outright duration exposure. We thus require all currency risk to be hedged back to an SDR-neutral position, and that the net duration across active positions is zero.

[10] SDR weights are approximately 41% (USD), 32% (EUR), 8% (GBP), 8% (JPY) and 11% (CNY) and are held constant over the sample for simplicity.

Upper- and lower-bound individual asset constraints — Table 2

In per cent

	US		DE		UK		JP		CN	
	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB
0.25Y	–10%	50%	–7%	50%	–1%	50%	–8%	50%	–2%	50%
2Y	–19%	50%	–14%	50%	–3%	50%	0%	0%	–5%	50%
4Y	–13%	50%	–11%	50%	–4%	50%	0%	0%	–4%	50%
6Y	0%	50%	0%	50%	0%	50%	0%	0%	0%	50%
8Y	0%	50%	0%	50%	0%	50%	0%	0%	0%	50%
Total	–41%		–32%		–8%		–8%		–11%	

Sources: IMF, ICE, author's calculations.

For a given period, we identify separate sets of portfolio weights for each signal by conducting optimisation exercises in the mean-variance space. In the most conventional mean-variance application (Markowitz (1952)), the practitioner would estimate the expected return and portfolio volatility from the observed, historical multivariable return distribution, and minimise volatility subject to a given return objective (among other possible constraints). In our application, the optimisation problem takes a modified set of inputs, in that we minimise portfolio volatility while imposing that the first moment is equal to our signal of interest, $CarrySig_t^j$ or $TPSig_t^j$. For our estimate of the covariance matrix, we do not use the historical return distribution per se; instead, we simulate 1,000 prospective yield curve paths over a horizon of 12-months, and compute monthly returns across simulations for each asset in our universe.[11] We then aggregate monthly returns into a single annual figure for each simulation/asset combination. This leaves us with a 1,000 by 21 matrix of returns from which our covariance matrix is estimated.

We identify optimal portfolio weights for three strategies: carry, term premium and a combination of the two, where the "combined strategy" takes the simple average of the carry and term premium strategy weights. As has been observed in the literature, combining different signals may help enhance risk-adjusted returns with diversification effects. One might also argue that it is another way to reduce exposure to the model risks described above.

We identify portfolio weights for each of the three portfolios at a monthly frequency from a sample beginning end-September 2016 and ending end-February 2023. We are thus left with a 78 by 21 matrix of weights for each of the three portfolios, reflecting our sample size and asset class universe, respectively.

3.2 Historical portfolio weights

Let us now examine the results of these signals-based optimisation exercises. We summarise them using two perspectives, each broken down by the individual currency components: net positions and duration contributions. Net positions reflect the sum of weights within a given currency, while duration contributions reflect the

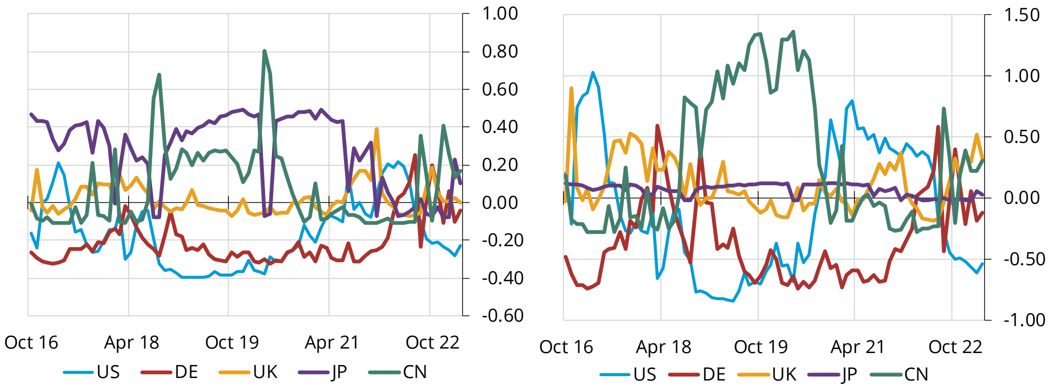
[11] More specifically, time series model errors are drawn at random from the historical distribution of model residuals and applied to the mean projected path.

sum of duration contributions within a given currency. Net positions and duration contributions each sum to zero across currency components, by construction.

Graph 5 shows the resulting portfolio weights for the "carry strategy". A few themes emerge. Over most of the sample's history, long positions in the short-dated JPY security offset a persistently net short EUR position – reflecting in part the well documented attractiveness of the JPY FX basis. A net long CNY duration position tended to offset a net short USD position from roughly mid-2018 to early 2020. There are also at least two instances of sharp reallocations between JPY and CNY assets that are quickly reversed, which upon further examination reflect volatility in the excess hedged FX return for CNY. We do not observe similar volatility in duration contributions, however, reflecting that this "switching" occurs between short-dated assets.

Net positions (lhs) and duration contributions (rhs) of carry strategy — Graph 3

Values (lhs); years(rhs)



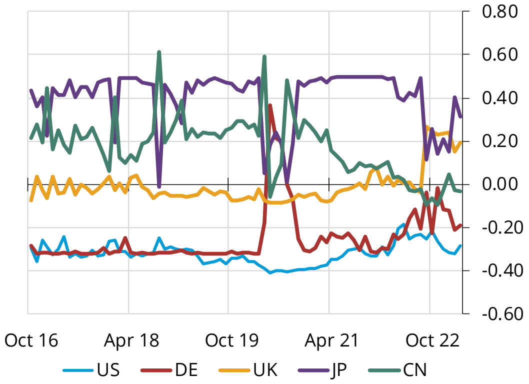
Sources: Bloomberg, author's calculations.

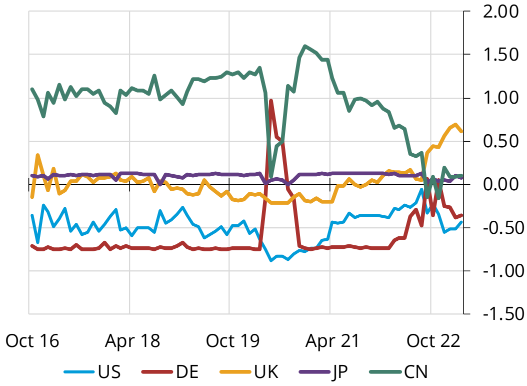
Graph 4 shows the resulting portfolio weights for the "term premium strategy". Themes are not wholly dissimilar from the carry portfolio, though differences are evident. The long positions in CNY and short-dated JPY are again clear themes – though more consistently across the sample – and are funded by persistent short positions in USD and EUR. The contribution from GBP is fairly neutral. We see significant volatility around the onset of the Covid-19 pandemic, when the term premium in EUR rose quickly before falling back to previous levels – a fact that may reflect a lag between changes in the economic outlook as reflected by asset prices and updates to economist forecasts that help drive term premium estimation. We again see some position switching between short-dated CNY and JPY assets, reflecting the same dynamic witnessed in the carry portfolio.

Net positions (lhs) and duration contributions (rhs) of term premium strategy

Graph 4

Values (lhs); years(rhs)





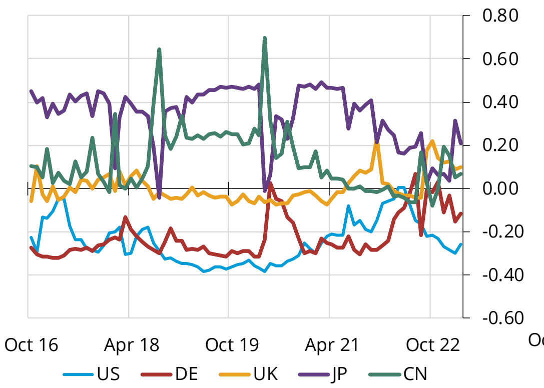
Sources: Bloomberg, author's calculations.

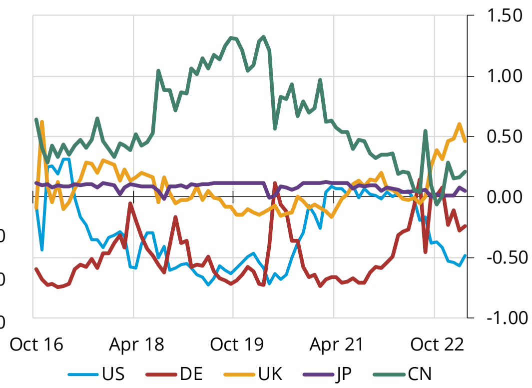
Graph 5 shows the results for the portfolio that takes the simple average of the carry and term premium strategy weights (the "combined strategy"), reflecting a blend of the themes discussed above. Importantly, we are still left with the observation that the weights can be rather volatile.

Net positions (lhs) and duration contributions (rhs) of combined strategy

Values (lhs); years(rhs)

Graph 5





Sources: Bloomberg, author's calculations.

4. Evaluating with a paper portfolio

Our portfolio optimisation exercises make, on an ex ante basis, a number of simplifying assumptions. For example, the use of generic assets implies that each holding is a zero coupon bond with a precise duration. Generic assets are also priced from fitted yield curves, which come with fitting errors. Furthermore, our procedure of monthly re-optimisation implies that portfolio weights will be tweaked on a monthly basis; furthermore, absent any changes in monthly weights, asset and portfolio return projections implicitly assume that the portfolio will be rebalanced back to its assigned weights at a monthly frequency. Finally, the portfolio optimisation problem incorporates no assumed transaction costs.

We seek to build a more realistic setting to perform our backtesting exercise with the use of a paper portfolio represented by actual securities. Given the issues reviewed above, we also need to address the various decisions required to apply model-theoretic weights to a more realistic setting. This includes accounting for transaction costs in some form, and designing a rule for rebalancing a portfolio of actual securities.

To build our paper portfolio, we assign each duration point an actual security holding for the start of our performance period, 1 October 2016, by finding the outstanding government bond with the nearest modified duration to our generic holding (whether the difference is positive or negative). Given that it is unlikely that we will find exact matches for many of our holdings, this will give rise to some error at individual points, though at the overall portfolio level, any drift is unlikely to be significant.

We apply the same approach to each monthly reoptimisation period. However, to economise on transaction costs, we allow positions to roll down one year beyond their generic duration assignment. For example, the US Treasury security corresponding to the four-year US generic asset is permitted to remain in the portfolio until it has a residual modified duration of three or more years. As this rule is applied across securities, it is unlikely to result in systematic duration drifts over time. For three-month assets, we simply allow them to mature. We also assume that JPY holdings are left in a cash account receiving 0% interest.

We apply a static set of transaction cost assumptions unique to each yield curve point. We do this by exercising judgment for each point on each yield curve, based on the observed bid-offer spreads across on- and off-the-run securities over our sample period. As evident in Table 3 below, bid-offer spreads are assumed to be notably higher for Chinese government bonds than for their German, UK and US counterparts. Note that JPY cash is assumed to be held in a cash account bearing a 0% interest rate.

Bid-offer spread assumptions | Table 3

In cents

	US	DE	UK	JP	CN
0.25Y	0.003	0.010	0.015	0.000	0.005
2Y	0.006	0.020	0.015	NA	0.040
4Y	0.010	0.020	0.015	NA	0.080
6Y	0.012	0.025	0.020	NA	0.120
8Y	0.016	0.030	0.020	NA	0.160

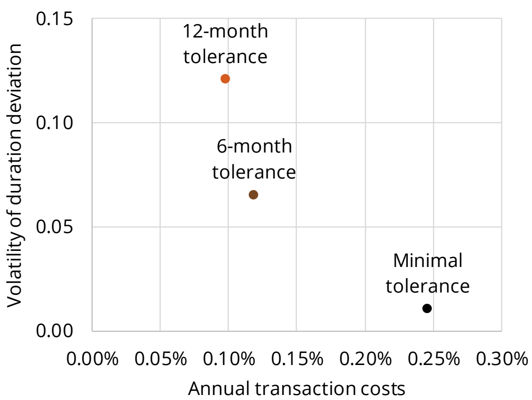
Source: Author.

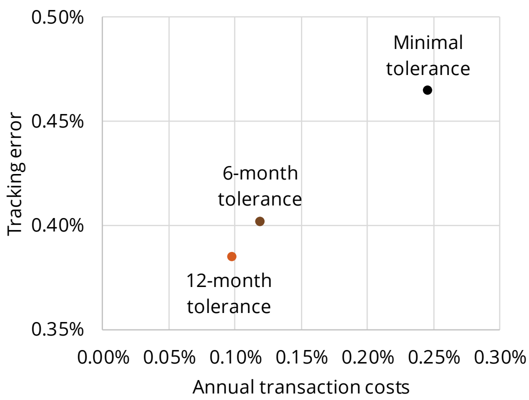
There is a clear trade-off between our tolerance for duration drift at the portfolio level and the level of transaction costs we will incur. The relationship is intuitive: the less tolerance for duration drift, the more tightly duration positions will need to be managed. Such a reduced tolerance will create more frequent rebalancing and result in greater transaction costs.

Graph 6 displays the implications of various rebalancing rules for transaction costs, the volatility of portfolio duration deviation and tracking error. As seen in the left-hand panel, having minimal tolerance for duration deviation implies heavy transaction costs of around 25 basis points per annum, given that one is required to switch into the nearest duration security at each monthly rebalancing cycle. Allowing securities to roll down the curve six months past their target durations cuts transaction costs in half, at the sacrifice of a slight uptick in the volatility of overall duration deviations. Allowing securities to roll down a full year further cuts transaction costs to below 10 basis points per annum, although at modestly higher deviations in duration.

Annual transaction costs versus volatility of duration deviation (lhs) and tracking error (rhs) for the combined strategy

Years and percentage (lhs); percentage (rhs) | Graph 6





Sources: Bloomberg, author's calculations.

One might expect deviations from a net duration position of 0 to result in higher portfolio volatility. The right-hand panel of Graph 6 shows the implications for tracking error with each of the rules. Allowing rolldown actually results in reduced levels of tracking error, despite the fact that duration deviation is more volatile. The desire to minimise transaction costs without taking on more risk motivates the choice of the one-year tolerance rule. Of course, the discussion around the management of duration assumes that that our hypothetical investor has only physical securities at their disposal. The ability to use futures would likely simplify the management of the overall duration position, probably at lower transaction costs.

5. Assessing performance

We examine the performance of each portfolio from multiple vantage points, including the excess return, ex post volatility (or tracking error), information ratio and maximum drawdown. We will do so gross and net of transaction costs, as well as decompose returns into their fixed income, FX and transaction cost components. We report results using an SDR and USD numeraire, recognising that the USD perspective may be more relevant for many investors.

Table 4 displays the performance across the three strategies before transaction costs. Viewed from the vantage point of the SDR, the carry strategy delivered flat returns over the sample before transaction costs, while the term premium strategy has an impressive information ratio of close to 1. The combined strategy has an information ratio of 0.69 and a reduced tracking error, owing to the diversification benefits of combining the signals.

Annual performance statistics gross of transaction costs Table 4
Various units

	SDR numeraire				USD numeraire			
	Excess return	Tracking error	Information ratio	Maximum drawdown	Excess return	Tracking error	Information ratio	Maximum drawdown
Carry	0.05%	0.49%	0.10	2.25%	0.00%	0.49%	0.00	2.17%
Term premium	0.52%	0.52%	0.99	0.84%	0.57%	0.52%	1.10	0.81%
Combined	0.29%	0.41%	0.69	1.12%	0.28%	0.41%	0.69	1.03%

Sources: Bloomberg, author’s calculations. Daily performance data from 1 October 2016 to 31 March 2023.

Table 5 shows the net performance once accounting for transaction costs. These push the carry strategy into negative territory, while the term premium strategy retains an impressive information ratio of 0.77 and the combined a respectable 0.47. The maximum drawdown is notably higher for the carry portfolio – over 2% – while those associated with the term premium and combined portfolios are much reduced, a little under or over 1%.

Annual performance statistics net of transaction costs Table 5

Various units

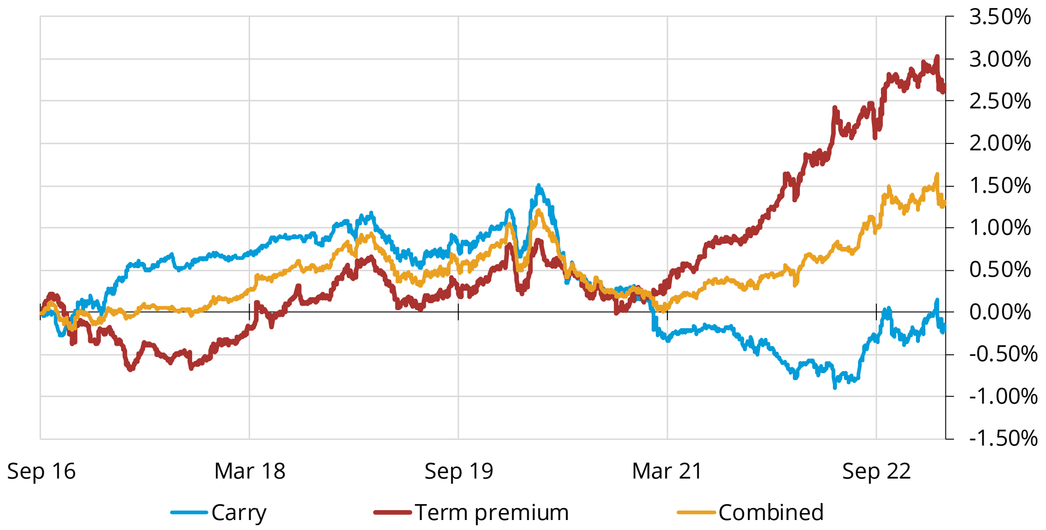
	SDR numeraire				USD numeraire			
	Excess return	Tracking error	Information ratio	Maximum drawdown	Excess return	Tracking error	Information ratio	Maximum drawdown
Carry	−0.03%	0.49%	−0.07	2.40%	−0.08%	0.49%	−0.17	2.33%
Term premium	0.40%	0.53%	0.77	0.90%	0.45%	0.52%	0.87	0.90%
Combined	0.19%	0.42%	0.47	1.22%	0.19%	0.41%	0.47	1.13%

Sources: Bloomberg, author's calculations. Daily performance data from 1 October 2016 to 31 March 2023.

Graph 7 shows a time series of the cumulative returns of each strategy over the period. Several observations stand out. The early and latter parts of the sample show considerable divergence between the performance of the two portfolios, while during the middle years the performance was much more correlated. Comparing Graphs 5 and 7 in Section 3.2 helps explain why: positive duration contributions from CNY holdings are a dominant feature of both portfolios during the middle portion of the sample, while the weights are much less correlated during the earlier and later parts.

Cumulative returns of the carry, term premium and combined strategies

Percentage Graph 7

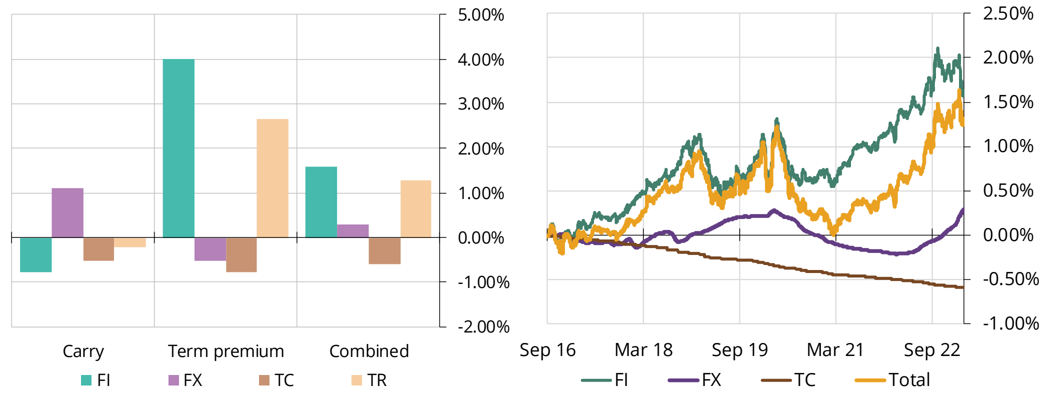


Sources: Bloomberg, author's calculations. Results shown in SDR terms.

Graph 8 shows the decomposition of the combined strategy by return component, specifically the fixed income (FI) return, FX return and transaction costs (TC), which collectively comprise the total return (TR). As evident in the left-hand panel, FI returns were negative on net for the carry strategy, while the FX component actually lifted performance. In contrast, FI returns fully explain the more impressive performance of the term premium strategy, while FX returns were a modest drag. The combined strategy's positive performance is mostly explained by the FI component, while the FX component has been close to neutral, both on net and historically. This decomposition may reflect more favourably on the term premium and combined strategies, given that our signals convey information about carry and value from a fixed income perspective, rather than an FX one.

Decomposition of net returns across strategies (lhs) and cumulative return by component over time of the combined strategy (rhs) Graph 8

Percentage



Source: Bloomberg, author's calculations. Results shown in SDR terms.

6. Considering alternative setups

The constraints put in place reflect our hypothetical reserve manager, who as an investor with a low risk appetite is wary of currency and outright duration risk. In this section, we relax these assumptions by considering two alternative setups, with all other portfolio construction aspects remaining unchanged (unless otherwise noted):

a. Flexible duration
Outright duration positions of +/– 0.5 years are permitted at the portfolio level.

b. Unhedged CNY
The currency risk of CNY positions are not hedged back to the SDR.

We examine the portfolio weights and performance of each.

6.1 Flexible duration

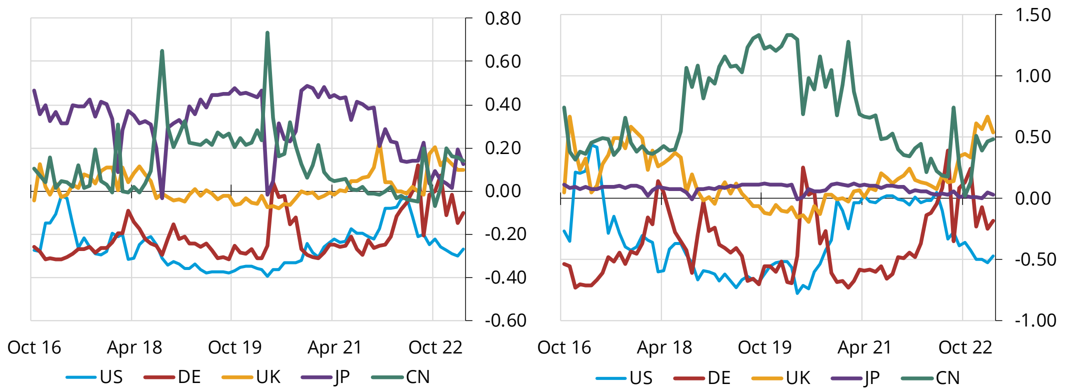
The first alternative setup relaxes the constraint that overall duration must sum to zero. Instead, constraints are put in place to permit portfolio duration positions up to a minimum of –0.5 years and a maximum of 0.5 years.

The reasoning behind relaxing this constraint is straightforward: just as our signals are used to exploit compensation for risk premia at the individual asset level, they can also be used to inform the direction of an overall duration position. For example, higher levels of term premium might motivate longer target durations for tactical or strategic asset allocation purposes, while conversely, negative term premia can motivate shorter target durations. With this alternative formulation, we test whether there is a meaningful performance impact from allowing this flexibility.

Graph 9 summarises the portfolio weights of the combined strategy under the flexible duration ("FD") setup. Thematically, they are quite similar to the baseline setup, generally featuring net long CNY and JPY positions over the history, funded with short USD and EUR. Duration contributions come largely from the CNY curve. Furthermore, the additional degree of freedom reduces portfolio turnover slightly, from 32% in the baseline strategy to about 30%. This will help with transaction costs, even if only slightly.

Net positions (lhs) and duration contributions (rhs) of combined strategy with flexible duration — Graph 9

Values (lhs); years(rhs)



Sources: Bloomberg, author's calculations.

As shown in Table 6 the flexible duration setup exhibits similar though slightly inferior performance to the fully hedged baseline. It is thus hard to conclude that allowing drift in overall duration adds meaningful value, at least over our sample.

Annual performance statistics net of transaction costs with flexible duration

Table 6

Various units

	SDR numeraire				USD numeraire			
	Excess return	Tracking error	Information ratio	Maximum drawdown	Excess return	Tracking error	Information ratio	Maximum drawdown
Carry	–0.08%	0.56%	–0.13	3.05%	–0.10%	0.56%	–0.17	2.99%
Term premium	0.34%	0.53%	0.64	0.99%	0.40%	0.53%	0.74	1.01%
Combined	0.15%	0.43%	0.34	1.36%	0.16%	0.43%	0.38	1.25%

Sources: Bloomberg, author's calculations. Daily performance data from 1 October 2016 to 31 March 2023.

6.2 Unhedged CNY

The second alternative setup relaxes the constraints around currency risk, specifically allowing open FX exposure for any CNY positions.[12] Most other characteristics of the baseline setup remain in place, including the requirement that duration be neutral at the portfolio level. The one exception is the target tracking error, which we allow to be 100 basis points given the additional FX risk. Of course, the tracking error in itself may not be very important, given that the size of an actual portfolio can simply be scaled to give the desired risk level in nominal terms.

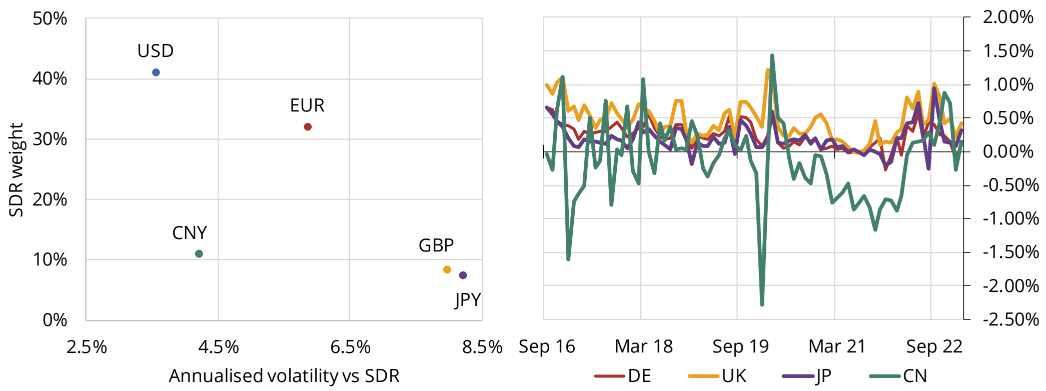
The motivation behind choosing not to hedge some portion of FX risk may be on shakier ground than the previous formulation. Our strategy seeks to exploit information content from signals that reflect interest rate risk, not currency risk. There is certainly empirical evidence that FX carry trades can deliver excess returns (see Fama (1984)), but they reflect a broader set of risks arguably less connected to our signals. For example, Menkhoff et al (2012) show that global FX volatility is an important risk factor in explaining the performance of FX carry trades.

Why do we choose to leave the CNY unhedged as opposed to a different currency or currencies? Several reasons motivate the choice. First, the managed flexible exchange rate exchange rate regime – in place since mid-2015 and paving the way for SDR inclusion – is of a different nature than the free-floating regimes in the euro area, Japan, the United Kingdom and the United States and, potentially helping to mitigate, relatively speaking, unhedged FX risk. The left-hand panel of Graph 10 demonstrates this point, displaying the volatility of each SDR constituent vis-à-vis the SDR's weights. As shown, the CNY-SDR FX rate displays comparable volatility with that of the USD, despite the fact that the CNY's weight in the SDR basket is only about one-quarter in magnitude. In contrast, the volatilities of GBP and JPY – whose weights are of similar magnitude to the CNY – display volatilities around twice that of the CNY.

[12] Given the lack of hedging, the FX component for the CNY term premium signal is set to zero in this formulation.

Annualised daily FX volatility versus the SDR (lhs) and excess hedged FX returns vis-à-vis the USD (rhs) Graph 10

Percentage



Sources: Bloomberg, author's calculations.

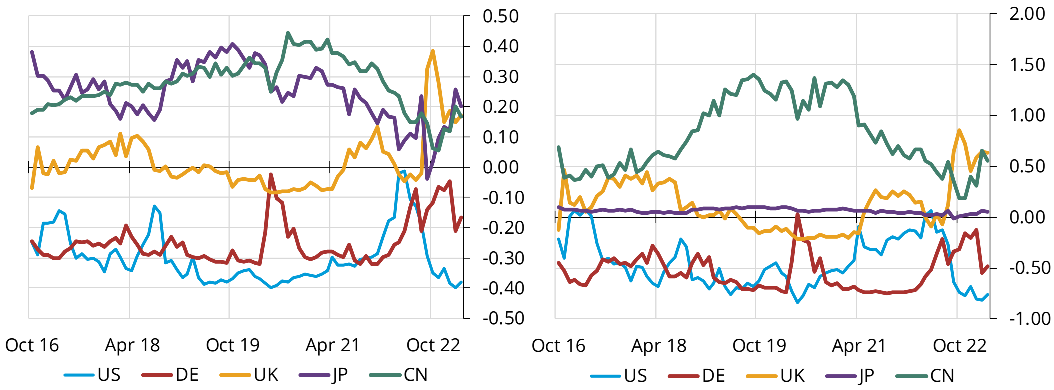
Furthermore, hedging the CNY comes with drawbacks relative to the other currencies. As shown in the right-hand panel of Graph 10, the CNY-USD excess hedged FX return, as measured using three-month government bond yields (the relevant perspective for our strategies), is considerably more volatile than its EUR, GBP and JPY counterparts. As discussed in more detail below, this fact helps explain the volatility in baseline strategy portfolio weights, and by extension the strategy's relatively high turnover.

Graph 11 summarises the portfolio weights of the unhedged CNY formulation. They are thematically similar to the baseline and flexible duration setups, but illustrate more persistent biases towards long CNY duration positions – unsurprising perhaps, given the relatively attractive interest rate levels coupled with the lack of hedging costs. Portfolio weights are also considerably more stable than in the baseline and flexible duration cases, cutting turnover by more than a third to around 20% and helping to reduce transaction costs.

Net positions (lhs) and duration contributions (rhs) of combined strategy with CNY positions unhedged

Graph 11

Values (lhs); years(rhs)



Sources: Bloomberg, author's calculations.

As shown in Table 7, the unhedged CNY formulation dominates the other two in terms of performance, delivering an information ratio – after transaction costs – of over 1 when using an SDR numeraire. The greater allocation to CNY and lack of hedging costs each play important roles in boosting the performance. More recently, the appreciation of the CNY has also contributed heavily, accounting on net for about 1 percentage point of the sample-wide performance. The smaller transaction costs – about 6 basis points per year, down from about 10 in the baseline setup – also help. Importantly, the higher tracking error helps reduce their *relative impact*, with transaction costs as a percentage of total performance dramatically lower when the CNY is left unhedged than in the baseline or flexible duration formulations.

Annual performance statistics net of transaction costs with CNY positions unhedged

Table 7

Various units

	SDR numeraire				USD numeraire			
	Excess return	Tracking error	Information ratio	Maximum drawdown	Excess return	Tracking error	Information ratio	Maximum drawdown
Carry	1.00%	1.31%	0.76	2.51%	0.55%	1.03%	0.53	2.52%
Term premium	1.15%	1.19%	0.97	1.76%	0.78%	1.08%	0.72	2.76%
Combined	1.08%	1.21%	0.89	1.91%	0.67%	1.02%	0.66	2.30%

Sources: Bloomberg, author's calculations. Daily performance data from 1 October 2016 to 31 March 2023.

Performance is modestly less impressive when viewing results from the USD numeraire perspective. Indeed, the choice of numeraire is more consequential when FX risk is left unhedged.

As noted above, tracking error may not be the most relevant risk metric for certain types of investor. For example, reserve managers, who can be quite sensitive to the near-term performance of their tactical positions, may see maximum drawdown as a more relevant consideration. Strategies that have periods of sharp drawdowns may come under pressure from a board or investment committee with limited risk appetite, even if medium-term expectations are consistent with a recovery in performance. As evident in Table 7, the maximum drawdown of the unhedged CNY formulation, in *percentage terms*, is quite elevated relative to the baseline and flexible duration versions, reflecting in large part the FX risk stemming from CNY positions. While reserve managers could scale down the size of their portfolios in nominal terms to account for the higher risk (as noted above with the TE), investment committees and risk departments may still be troubled by the larger drawdown risk per unit invested.

Concluding remarks

This paper has drawn from previous research at the Public Investors Conference to formulate and test systematic investment strategies based on two well known concepts in the fixed income space: carry and term premium. Our main conclusion – that term premium strategies deliver promising excess returns – is consistent with the earlier papers, lending additional confidence to the finding.

What can we say about the lukewarm results delivered by the carry strategy? Coche et al (2018) document that positive excess returns associated with the carry strategy began to take hold persistently only in the 1980s – the beginning of a period marked by a secular decline in interest rates, when high-carry assets would have naturally outperformed. Our sample, however, is taken after this period, with most SDR yield curves having reached their historical lows in the early to mid-2010s. We may thus be suspicious of the value of the carry signal outside such environments.

Why might an investor want to nonetheless consider a combined strategy, when the carry strategy looks less compelling and the information ratio of the combined strategy appears inferior, at least over our sample? One reason may be the management of model risk, as estimating a term premium requires an unobservable element to be estimated – the expected short rate path. While our macro-based approach is consistent with the general understanding of the drivers of central bank policy, it may not always represent the market's physical expectations for the evolution of the policy rate. The combined strategy helps to reduce this model risk, given that carry is a model-free measure. It also helps to diversify the risk – as evident in the lower tracking error, and a maximum drawdown figure that is much closer to that of the term premium strategy than the carry strategy. Finally, it's worth noting that the combined strategy has spent nearly the entire sample above water, save for a few days here and there in slightly negative territory.

There are nonetheless many reasons to be cautious with strategies, despite the promising results reviewed here and in earlier research. Any backtesting exercise is by virtue a reflection of past market dynamics, and provides no guarantee of a future relationship. Systematic strategies also remove discretion by design, and may be slow to recognise regime changes. For example, estimation of term premia may lag important market developments, such as the onset of a pandemic, implying that signals may have less meaning during times of rapid changes in the financial environment.

These cautionary points lead to a few simple final remarks. First, investors in such strategies cannot necessarily expect immediate results, and should be prepared for periods – perhaps sustained periods – of underperformance. This can be a challenge, particularly in the presence of investment boards with a low tolerance for P&L volatility. Second, risk should be appropriately calibrated, such that periods of underperformance and incidents of sharp drawdowns can be sustained without abandoning the strategy. Third, it is useful to keep in mind why we would expect signal-based strategies to work in the first place. As Bjorheim et al (2018) point out, strategies such as these may be powerful tools not due to the superior quality of the models and their predictive power, but because they give rise to reasonable estimates of risk premia. The investor is thus merely being compensated for exposure to risk. A key element to investigate further may be how correlated such risk factors are with other elements of active management, or with the underlying strategic asset allocation. At the end of the day, systematic strategies such as these may prove most useful as a way of diversifying elements of a broader strategic and active management approach.

References

Adrian, T, R Crump and E Moench (2013): "Pricing the term structure with linear regressions", *Staff Reports*, Federal Reserve Bank of New York, no 340.

Ahmerkamp, J and J Grant (2013): "The returns to carry and momentum strategies", Imperial College Business School Working Paper.

Asness, C, T Moskowitz and L Pedersen (2013): "Value and momentum everywhere", *The Journal of Finance*, vol 68, no 3, pp 929–85.

Baz, J, N Granger, C Harvey, N Le Roux and S Rattray (2015): "Dissecting investment strategies in the cross section and time series", working paper.

Bjorheim, J, J Coche, A Joia and V Sahakyan (2018): "A macro-based process for actively managing sovereign bond exposures", *Advances in the Practice of Public Investment Management*, pp 103–29, Palgrave Macmillan.

Coche, J, Knezevic, M and V Sahakyan (2018): "Carry on", *Advances in the Practice of Public Investment Management*, pp 131–59.

Fama, E (1984): "Forward and spot exchange rates", *Journal of Monetary Economics*, no 14, pp 319–38.

Fender, I, M McMorrow, V Sahakyan and O Zulaica (2020): "Reserve management and sustainability: the case for green bonds?", *BIS Working Papers*, no 849.

Fender, I, M McMorrow and O Zulaica (2022): "Sustainable management of central banks' foreign exchange (FX) reserves", *The Inspire Sustainable Central Banking Toolbox*, Policy Briefing Paper, no 6.

International Monetary Fund (2016): "IMF adds Chinese Renmimbi to Special Drawing Rights basket", press release, 30 September.

Kim, D and J Wright (2005): "An arbitrage-free three-factor term structure model and the recent behavior of long-term yields and distant-horizon forward rates", *Finance and Economics Discussion Series*, Federal Reserve Board, no 2005-33.

Koijen, R, Moskowitz, T, Pedersen, L and Vrugt, E (2017): "Carry", *Journal of Financial Economics*, pp197–225.

Lombardi, M and F Zhu (2018): "A shadow policy rate to calibrate US monetary policy at the zero-lower bound", *International Journal of Central Banking*, vol 14, no 5, pp 305–46.

Markowitz, H (1952): "Portfolio selection", *The Journal of Finance*, vol 7, no 1, pp 77–91.

Menkhoff, L, L Sarno, M Schmeling and A Schrimpf (2012): "Currency momentum strategies", *Journal of Financial Economics*, vol 106, issue 3, pp 660–84.

Nelson, C and A Siegel (1987): "Parsimonious modelling of yield curves", *The Journal of Business*, vol 60, issue 4, pp 473–89.

Nyholm, K (2015): "A rotated dynamic Nelson-Siegel model with macro-financial applications", *ECB Working Paper Series*, no 1851.

Stock, M and M Watson (2007): "Why has US inflation become harder to forecast?", *Journal of Money, Credit and Banking*, supplement to vol 39, no 1.