XAIA INVESTMENT

Empirical evidence for the existence and temporal stability of the negative basis

Jan-Frederik Mai
XAIA Investment GmbH
Sonnenstraße 19, 80331 München, Germany
jan-frederik.mai@xaia.com

Date: August 2, 2017

Abstract In theory, a bond yield may be decomposed into the sum of three parts. One part corresponds to a currency-adjusted, risk-free interest rate, and another part may be viewed as a compensation for exposing oneself to credit risk associated with the bond issuer. While the former part can be estimated from observed market prices of interest- and FX-sensitive standard derivatives (such as interest rate and cross-currency swaps), the latter part can be estimated from observed market prices of credit default swaps (CDS) referring to the bond issuer. The remaining, third part of the yield, which is not explained by these two parts, is called the *negative basis* of the bond. A negative basis fund aims to isolate the negative basis as income source, by investing into bonds while at the same time hedging away interest rate-, FX-, and credit-risk via the aforementioned respective derivatives. According to this theoretical underpinning, the daily PnL of a negative basis fund should result solely from daily consumption, and from daily fluctuation, of the negative basis (averaged over the single positions in the fund). Decomposing the fund into one part consisting of all bond investments (primary assets), and a second part consisting of all other assets (hedging and cash instruments), the resulting two daily PnL time series are both exposed to interest rate-, FX-, and credit-risk in addition to the negative basis risk, while their sum, which by definition equals the fund PnL time series, is only exposed to negative basis risk. The investment idea relies on precisely this decomposition hypothesis, and in addition on the hypothesis that the average negative basis is relatively stable (and sufficiently positive) over time, which renders it an attractive, low-volatile source of income. Under the assumption that both hypotheses hold, the daily PnL time series of primary assets and hedging assets should exhibit an extremely high negative correlation. The present note describes how to compute the latter correlation and thus implement a historical backtest of the two theoretical hypotheses in practice.

1 Methodology

To be slightly more precise regarding the practical implementation of the idea outlined in the abstract, we decompose the negative basis fund into three parts. The first part A contains all bonds (primary assets). The second part B contains all CDS and interest rate swaps. The third part C contains the rest, which comprises cash instruments and FX-hedges. Theoretically, we would like to merge B and C into one part, as indicated in the abstract, but due to technical restrictions[1] cash and FX-hedges need to

[1] Cash and FX instruments are treated differently in the back office data base, which complicates their inclusion in the analysis in practice.

be treated in an idealized manner as described in the sequel.
Cash is ignored completely in the analysis, which is not a severe assumption since the daily PnL of the cash account is extremely stable and small compared to all otherwise considered daily PnLs, hence hardly affects the overall analysis. Moreover, under the absence of significant in- or outflows of the fund, a change in the cash account corresponds to a change in the investment ratio, which affects the absolute values of A and B in a similar order of magnitude, according to the investment idea. Thus, the correlation between the daily PnLs of A and B should remain invariant with respect to a change in the investment ratio, unlike the volatilities of the daily PnLs of A and B which depends on the absolute sizes of A and B.
In order to properly take into account the FX-hedge, we omit all FX instruments in the analysis and instead idealistically assume that the parts A and B exhibit zero FX risk. This is achieved by subtracting the FX-induced parts from the daily PnLs of A and B, as described in Section 1.2 below. Since the portfolio management of a negative basis fund tries to approximate such a perfect FX hedge in its daily business anyway, and since this FX-hedging part of portfolio management is straightforward, accurate, and standard, this idealized treatment of the FX-hedge is a quite realistic and easy-to-implement approximation of reality, hence induces no severe flaws into the analysis. Consequently, after this idealized treatment of FX-hedges, we are left with FX-adjusted parts A (bonds without FX-risk) and B (CDS and interest rate swaps without FX-risk).
Even though FX-risk is already subtracted according to the aforementioned idealization, the daily PnLs of A and B, in the sequel denoted by ΔA and ΔB, are fully exposed to interest rate risk and credit risk and negative basis risk, hence can be highly volatile. According to the theoretical underpinning in the abstract above, however, the sum of the daily PnLs of A and B, which equals by definition the daily PnL of the total fund (modulo the idealized treatment of part C), should be relatively stable and hardly volatile. This idea relies on the theoretical hypothesis that it results solely from daily consumption and daily fluctuation of the negative basis, which is (according to the investment idea) expected to be relatively stable over time. Under this theoretical hypothesis, it is clear that the correlation coefficient between the daily PnL time series of A and B is extremely negative, which is precisely what the method to be described in the sequel seeks to verify. In theory, if this correlation was exactly -100%, this would mean precisely that

$$\begin{aligned}\Delta A + \Delta B &= \text{a constant} \\ &= \text{daily consumed cash amount resulting} \\ &\qquad \text{from a constant negative basis,}\end{aligned} \tag{1}$$

just as predicted by the aforementioned hypotheses. In practice, we clearly do observe fluctuations of $\Delta A + \Delta B$, which can be due to (i) negative basis fluctuations[2] or (ii) changing investment

[2] The interested readers are referred to Bernhart, Mai (2016) for a list of reasons for the existence of negative basis, which explains which kinds of risk

ratio and fund size (i.e. shifts into/from cash). However, under the asumption of (ii) being relatively stable over time (i.e. relatively stable cash account), our method aims to verify in particular that the negative basis fluctuations (i) are quite small. According to (1), this goes along with an extremely negative correlation coefficient between ΔA and ΔB close to -100%. Thus, the presented analysis substantiates not only a strong fundamental relationship between the returns of primary assets (bonds) and associated hedging assets (CDS and interest rate swaps), but also that the negative basis resulting from this relationship evolves relatively stable over time. Intuitively, the closer the correlation between ΔA and ΔB is to -100%, the smaller are the PnL fluctuations induced by changes in investment ratio and negative basis in comparison to the fluctuations induced by the other involved risk factors (credit-risk and interest rate risk). This intuition can also be understood heuristically by means of the following formula for the Pearson correlation coefficient ρ between ΔA and ΔB, resulting from the formula for the variance of two correlated random variables by rearranging the terms:

$$\rho = \frac{\mathrm{Var}(\Delta A + \Delta B)}{2\sqrt{\mathrm{Var}(\Delta A)\,\mathrm{Var}(\Delta B)}} - \frac{1}{2}\Big(\sqrt{\frac{\mathrm{Var}(\Delta B)}{\mathrm{Var}(\Delta A)}} + \sqrt{\frac{\mathrm{Var}(\Delta A)}{\mathrm{Var}(\Delta B)}}\Big).$$

Under the assumption that the idiosyncratic variances of ΔA and ΔB are identical, which is intuitive by recognizing that the fluctuations of A and B are caused by the same risk factors (credit-risk and interest rate risk), the second summand in this formula equals -1. Consequently, the correlation between ΔA and ΔB is equal to -1 if the fund PnL $\Delta A + \Delta B$ has no variance, i.e. is a constant. Thus, the presented correlation measurement is expected to increase with the historically observed fund PnL.

1.1 Notations

For historical time points $t = 0, 1, \ldots, T$, where T corresponds to the current valuation date, we introduce the following notations:

- **Bonds:** Let N_1 denote the number of distinct bonds that have been in the fund at some time point $t \in \{0, \ldots, T\}$. For $i = 1, \ldots, N_1$ we denote by $X_i(t)$ the value of bond i with unit nominal[3] at time t (arbitrary in case already expired or called, etc.), and by $N_{X_i}(t)$ the nominal of the bond held in the fund at time t (possibly zero, if the bond was not in the fund at time t).

- **Interest rate swaps:** Let N_2 denote the number of distinct interest rate swaps that have been in the fund at some time point $t \in \{0, \ldots, T\}$. For $i = 1, \ldots, N_2$ we denote by $Y_i(t)$ the value of swap i with unit nominal at time t (arbitrary in case already expired, etc.), and by $N_{Y_i}(t)$ the nominal of the swap held in the fund at time t (possibly zero, if the swap was not in the fund at time t).

the negative basis actually compensates for. From this, the readers might retrieve a feeling for how volatile this risk compensation could/should be.

[3] By an asset's "unit nominal" we mean a nominal that equals one unit of the currency the asset is denominated in.

- **CDS:** Let N_3 denote the number of distinct CDS that have been in the fund at some time point $t \in \{0, \dots, T\}$. For $i = 1, \dots, N_3$ we denote by $Z_i(t)$ the value of CDS i with unit nominal at time t (arbitrary in case already expired, etc.), and by $N_{Z_i}(t)$ the nominal of the CDS held in the fund at time t (possibly zero, if the CDS was not in the fund at time t).

For each of the enumerated assets $x \in \{X_i, Y_j, Z_k \,:\, i = 1, \dots, N_1,\, j = 1, \dots, N_2,\, k = 1, \dots, N_3\}$ we denote by $\mathsf{FX}_x(t)$ the EUR-value of one currency unit of the asset x. For instance, if $x = Y_2$ is a USD-interest rate swap, then $\mathsf{FX}_{Y_2}(t)$ is the value of one USD in EUR at time t.

1.2 Computation

We let $\Delta A(t)$ (resp. $\Delta B(t)$) denote the PnL of A (resp. B) in the period $[t, t+1]$ from t to $t+1$. We compute these quantities by aggregating all single asset PnLs in the fund A (resp. B) as

$$\begin{aligned}
\Delta A(t) &= \sum_{i=1}^{N_1} \text{PnL of asset } X_i \text{ in period } [t, t+1], \\
\Delta B(t) &= \sum_{i=1}^{N_2} \text{PnL of asset } Y_i \text{ in period } [t, t+1] \\
&\quad + \sum_{i=1}^{N_3} \text{PnL of asset } Z_i \text{ in period } [t, t+1].
\end{aligned}$$

The PnL of a single asset x in period $[t, t+1]$ is computed as follows. If the asset was not FX-hedged, the PnL was given as

$$\big(x(t+1)\,\mathsf{FX}_x(t+1) - x(t)\,\mathsf{FX}_x(t)\big)\, f\big(N_x(t), N_x(t+1)\big), \quad (2)$$

where the function f is explained in brevity. Clearly, the first factor in (2) equals simply the difference between asset value at $t+1$ and asset value at t (for one unit nominal). The second factor in (2), involving the asset nominal at times t and $t+1$, guarantees that a potential purchase or sale of the asset in period $[t, t+1]$ is not mistaken for PnL, since in that case there is an offsetting entry on the cash account in part C (which is ignored in the analysis). There are two possibilities for the function f, which are both equally feasible for the analysis. Either, one assumes that a nominal change in $[t, t+1]$ becomes PnL-effective already in the current period $[t, t+1]$ (i.e. the nominal change is effective immediately at t), or it only becomes effective the next period $[t+1, t+2]$ (i.e. the nominal change is effective at the end of the period at $t+1$):

$$\begin{aligned}
f\big(N_x(t), N_x(t+1)\big) &= N_x(t+1) && \text{(PnL-effective at } t\text{)}, \\
f\big(N_x(t), N_x(t+1)\big) &= N_x(t) && \text{(PnL-effective at } t+1\text{)}.
\end{aligned}$$

Now let us assume additionally that asset x was (perfectly) FX-hedged in period $[t, t+1]$. This means that at time t we entered into an FX-forward contract with maturity $t+1$ and nominal $x(t)\, f\big(N_x(t), N_x(t+1)\big)$. This contract would have contributed the following additional PnL:

$$-x(t)\, f\big(N_x(t), N_x(t+1)\big)\,\big(\mathsf{FX}_x(t+1) - \mathsf{FX}_x(t)\big).$$

The sum of this PnL and (2) equals the PnL of asset x in period $[t, t+1]$ under the assumption of a perfect FX-hedge:

$$\big(x(t+1) - x(t)\big)\, \mathsf{FX}_x(t+1)\, f\big(N_x(t), N_x(t+1)\big). \tag{3}$$

In the sequel, for the sake of a more compact notation we introduce the notation $\Delta x(t) = x(t+1) - x(t)$. Summarizing, under the asumption of a perfect FX-hedge we obtain a total PnL of A (resp. B) in period $[t, t+1]$ by

$$\begin{aligned}
\Delta A(t) &= \sum_{i=1}^{N_1} \Delta X_i(t)\, \mathsf{FX}_{X_i}(t+1)\, f\big(N_{X_i}(t), N_{X_i}(t+1)\big),\\
\Delta B(t) &= \sum_{i=1}^{N_2} \Delta Y_i(t)\, \mathsf{FX}_{Y_i}(t+1)\, f\big(N_{Y_i}(t), N_{Y_i}(t+1)\big)\\
&\quad + \sum_{i=1}^{N_3} \Delta Z_i(t)\, \mathsf{FX}_{Z_i}(t+1)\, f\big(N_{Z_i}(t), N_{Z_i}(t+1)\big).
\end{aligned}$$

Finally, we compute the linear correlation coefficient between the two time series $\big(\Delta A(t)\big)$ and $\big(\Delta B(t)\big)$, with $t \in \{0, \ldots, T-1\}$.

Remark 1.1 (Aggregated vs. single-position view)
In theory, the described methodology could also be performed on every single negative basis position, i.e. one could assign one correlation measurement to every single(-name) negative basis position in the fund. However, the described aggregated view has several advantages compared to this single-position approach:

- There is no need to map the bonds to their associated CDS contracts in the described methodology, which would be technically slightly more inconvenient.
- An efficient portfolio management implements the interest rate hedge on a fund level, i.e. interest rate swaps are used to hedge away interest rate deltas that are observed on aggregated fund level. Consequently, in the single-position approach this global hedge would have to be decomposed into many parts, each part corresponding to a single position, which is technically burdensome.
- If one was interested in a single correlation figure for the fund also in the single-position approach, one would have to average the single-position correlation figures somehow. A position weighting in this averaging process would naturally be linked to the position nominal somehow. But bond nominals can differ from CDS nominals due to jump-to-default exposures etc., so the aggregated view allows to conveniently circumvent non-trivial deliberations as to how define a position weight appropriately.
- If a position has been built up only recently, its history in the fund is very short, possibly resulting in a less meaningful correlation measurement. Thus, in the single-name approach only those positions with sufficiently long history could be assigned a meaningful correlation measurement as described above.

2 Application

For a representative negative basis fund[4], the described analysis has been carried out on 23 June 2017, with the last data point corresponding to 22 June 2017 (corresponding to time point $t = T$). In the sequel, we describe the practical implementation of the described methodology in some detail, in order to highlight that it is strikingly easy to implement. The historical data ranges back to 22 June 2016 (corresponding to $t = 0$), which amounts to 250 business days, i.e. $T = 249$. For every asset (bond, interest rate swap, or CDS) that was in the fund on at least one day during this time frame, our back office provides three time series:

(i) the nominal of the asset in the fund (i.e. $N_x(t)$),

(ii) the value of the asset in its currency,

(iii) the value of the asset in EUR.

This data is provided in an xls-file, where for each asset class and for each of (i), (ii), and (iii) there is one xls-sheet, i.e. $9 = 3 \times 3$ sheets in total. Columns in the respective sheets correspond to assets, and rows correspond to business days. All 9 sheets have the precisely same number of rows, namely $T + 1$, and for each asset class the three sheets (i), (ii), and (iii) have the precisely same number of columns. All required input data for the analysis is comprised within these 9 xls-sheets, except for the required FX rates, which we retrieve from Bloomberg on an extra xls-sheet. This leaves us with 10 input data xls-sheets in total.

For the implementation of the formulas derived in the main body of this note, we create one additional xls-sheet for each asset class. A particular column of this sheet corresponds to a particular asset of the respective class, i.e. the sheet has precisely as many columns as its associated 3 input data sheets. Also, it has one row less[5] than the 10 input data sheets, i.e. T rows. In a particular row, say t, and in a particular column, say x, of this sheet, we compute the PnL of the asset x in the period $[t, t+1]$ according to formula (3), using input data from the three associated input xls-sheets. Having done this, we have 13 xls-sheets in total, 10 input data sheets, and 3 sheets with daily PnLs for each asset. Finally, we create a summarizing xls-sheet that performs the aggregation of single assets to the two PnLs ΔA and ΔB, and computes the desired correlation coefficient. To this end, for each asset class, we create one column giving the time series of daily aggregated PnLs of all assets of the respective asset class. This simply amounts to a column of row sums computed from the sheet containing the daily single-asset PnLs. Finally, we add together the two columns associated with CDS and interest rate swaps to obtain the time series $\big(\Delta B(t)\big)$, the bond column is already equal to $\big(\Delta A(t)\big)$, $t = 0, \ldots, T-1$. Via the xls-built-in function `CORREL()` we ultimately compute the desired correlation coefficient.

In the particular analysis the resulting correlation measurement was -97.55%, which provides satisfying evidence for the aforementioned hypotheses, i.e. that there exists a negative basis that

[4] A certain average of our own funds.

[5] Recall that the assset value time series runs through $t = 0, \ldots, T$, and the PnLs computed from them runs through $t = 0, \ldots, T-1$.

can be consumed via the negative basis fund and behaves relatively stable over time.

Finally, Figure 1 depicts two time series that provide a feeling for the quality of the underlying data and for the approximations made due to the technical restrictions. On the one hand, the red line in Figure 1 is derived from the official NAV time series of the representative negative basis fund. The retrieved time series from 22 June 2016 to 22 June 2017 is normalized such that it starts at 100. On the other hand, the blue line in Figure 1 is derived from the data underlying the described analysis. It basically shows the value process of the sum of the two partial funds A and B, normalized such that it also starts at 100.

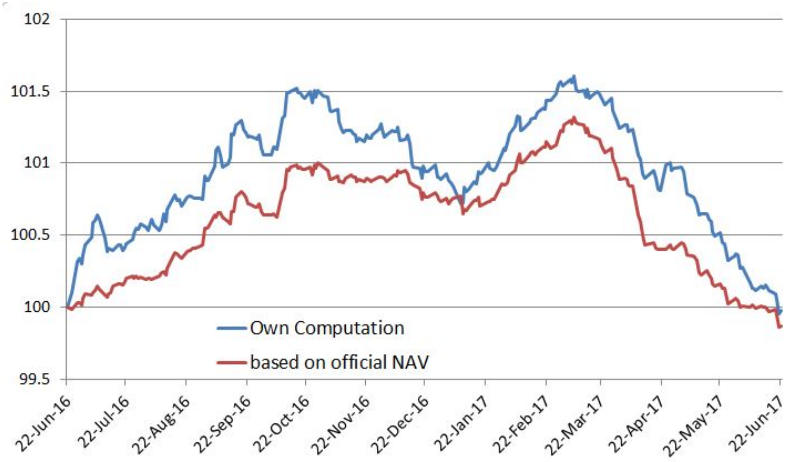

Fig. 1: Comparison of official NAV and sum of A and B as described in the present note for the representative negative basis fund, both time series normalized such that they start at 100.

The two lines are apparently quite similar, the slight difference between them is due to some abstractions from reality taken for granted in the described analysis, due to technical restrictions. However, we believe that this difference has no essential impact on the computed correlation. In particular, the following potential reasons for the difference of red and blue lines come to one's mind:

- The red line contains a cash account, the blue line does not. The actual fund had an investment ratio fluctuating between 73% and 95% in the considered time period. As argued before, the contribution of this difference to the computed correlation is expected to be minimal.
- The red line contains fund administration fees, the blue line does not. This implies a constant downward drift of the red line that is not present in the blue line. On the level of daily PnLs, this constant drift contributes a daily amount with extremely low volatility, hence with minimal effect on the computed correlation.
- The red line contains the actual FX hedge performed by portfolio management, the blue line is based on the described

idealized FX hedge. As argued before, since the portfolio management tries to approximate the ideal FX-hedge in practice, the FX-contribution to the difference of red and blue lines should be negligible.

3 Conclusion

Based on historical data of the assets in a negative basis fund, it has been shown how to implement a backtest for the hypothesis that a negative basis exists and is relatively stable over time. This hypothesis is equivalent to saying that the daily PnLs of primary assets (bonds) and hedging assets (CDS, interest rate- and FX-hedges) exhibit almost perfect negative correlation. Furthermore, this backtest has been implemented for a representative negative basis fund, with the aforementioned correlation estimated at -97.55%, providing significant evidence for the hypothesis to hold.

References

G. Bernhart, J.-F. Mai, Negative basis measurement: finding the holy scale, in *Innovations in Derivatives Markets - Fixed Income modeling, valuation adjustments, risk management, and regulation, edited by K. Glau et al., Springer-Verlag* (2016) pp. 385–403.