Credit Market Final Report

Capital Structure Arbitrage Strategies with Price Discovery Augments

Author: Jer-Shen Jeffrey Chen

Introduction

Capital structure arbitrage refers to trading strategies that take advantage of the relationship of mispricing between different security classes issued from the same company's capital structure. Mainly, the arbitrage opportunities appear between equity-linked and debt-linked securities. Structural models treat equity and debt as claims on the value of the firms, thus, the temporary mispricing often arises from equity and debt markets having different participants and market structures that create the discrepancy of price discovery process and speeds. A typical example would be an after-effect of a firm's earnings reports. If a firm surprises the market with disappointing earnings, a company's stock possibly drops 10 percent immediately, yet that same information may not be reflected in the company's bond price until a few days later, and possibly only result in a 2 percent drop in the bond price. However, this scenario can be exploited and make a profit systematically from mispricings and divergent intermarket dynamics.

These capital structure strategies (CSA) are usually implemented by offsetting the positions of an issuer's debt, CDSs, swaps, and equity option securities. The core of making an arbitrage is to go long the undervalued securities linked to one part of the company's capital structure, at the same time hedging the position by going short overvalued securities linked to another part of the capital structure to create a floor or ceiling. It is essentially like a convertible arbitrage strategy where we go long the convertible bond and short the company's stock.

In homework 3, we visited a more sophisticated capital structure arbitrage using the empirical capital structure of Boeing and the 5 year at-market CDS spread. The trading strategies use implied volatilities from the equity options markets against default probabilities implied by the credit spreads, effectively arbitraging the default probability predicted by the CDS market and the equity options market. CDS rate is equivalent to a prox of default probability. Thus, the interrelationship between volatility skew, CDS rate, and credit spreads creates capital structure arbitrage opportunities. More specifically, the implementation of the strategies is based on information coming from two time series of spreads, the 5-year at-market spread and an equity-implied spread which is obtained from CreditGrades, a Merton extension structural credit model. When implementing the CSA strategy, we look at a significant divergence between the CDS spread and its implied spread. Hence, a trader would short (long) a CDS contract if the CDS spread is significantly higher (lower) than the implied spread and short (long) a given number of shares as an equity hedge to offset the CDS position.

The CSA strategy (including hedging) would work well if both markets are equally efficient in the sense that none leads the other one, i.e. any discrepancy between them is random and short lived, and price discovery occurs simultaneously in both of them. The assumptions let structural credit risk models generate good estimates not only on implied spreads but also hedge ratios.

Yet, our homework study showed that the gains of the strategy are not very consistent and profitable across real-world Boeing datasets. We performed data-mining on different parameters (both time-dependent or trigger thresholds) to try to maximize the profits. As the below figure and table shows, the results are not very ideal, and the cumulative returns/sharpe are negative across all data mining.

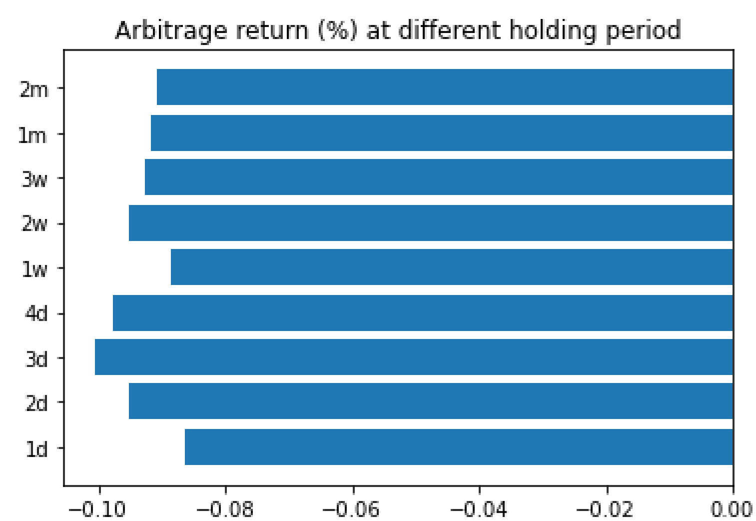


Table. Mean Arbitrage Return (%) of holding periods and thresholds

z-score	1w	2w	3w	1m	3m	6m
3	-0.20	-0.38	-0.38	-0.90	-1.57	-1.07
1	-0.15	-0.47	-0.80	-0.04	-0.10	-0.14
0.5	-0.16	-0.37	-0.57	-0.04	-0.07	-0.09

There are several reasons that account for the loss of expectation of the CSA strategy. The study by Yu (2006) showed that hedging strategies used to offset CDS positions with equities can be ineffective and expensive. Another possible reason, according to Das and Hanouna (2009), is that equity hedges can be very expensive when markets become volatile because the hedge ratio varies very quickly and another determinant factor would be the lack of liquidity of the equity market.

Given that hedging may be ineffective for the CSA strategy, and several studies have shown that a lead-lag relationship between equity and CDS markets would affect the forecasting power of the Credit models. We found some paper proposed trading CSA with Price Discovery on the market that is being led would generate positive Sharpe ratios and can be more easily to forecast.

We implemented this trading strategy at homework 3, using the empirical capital structure of Boeing and the 5 year at-market CDS spread. USing the CreditGrade implied CDS spread to observe the discrepancy of mispricing opportunity of CDS spread.

The price discovery in the equity market is being widely analyzed and there is a vast literature that can be found, including professor Longstaff (2003) as well. These studies focused on the information flow between CDS and equity markets and most the paper we reviewed shows evidence of time variation in the price discovery of credit-related information. Therefore, we took the similar approach presented in these papers, using both the Vector Error Correction Model (VECM) for changes in spreads, and time-varying price discovery measures to combine with the CSA strategies for trading CDS and equity markets.

In this report, we analyze two different trading strategies and compare them with standard CSA strategy, using the same Boeing dataset to evaluate the performance over the same period of time. The two strategies are based on the flaws of the CSA we observed, namely (1) not sensitive to the information efficiency, meaning that if lead-lag relationship is presented, the leading market information, says market A, makes perfect sense to trade on market B based on information released in market A; (2) the ineffective hedging position might lead to extensive consumption of the profits. We analyzed the risk and returns of these strategies and investigated the different holding periods and trigger moments to maximize the performance of the strategy.

Price Discovery

Information has a fundamental role in the formation of prices in secondary markets. Understanding how prices efficiently incorporate information about fundamentals has been, and remains, one of the main interests of the market microstructure literature. Moreover, increased market fragmentation means that the same asset may be trading simultaneously in multiple markets. Therefore, to measure the relative degree of efficiency of these markets it is important to estimate the speed at which related prices incorporate news. This process is called price discovery, and there are currently two main measures that are widely used in the literature. They both base their analysis on structural models of cointegrated price series that share a common random-walk efficient price. The first metric aims at quantifying how much of the variance of the efficient price can be attributed to the different markets. Hasbrouck (1995) refers to this proportion as the Information Share (IS). The second measure is the Component Share (CS), applied by Booth et al. (2002), Chu et al. (1999) and Harris et al. (2002) on the basis of the econometric work of Gonzalo and Granger (1995). It focuses on the decomposition of the price series into a permanent component, that reflects the contribution of the efficient price, and a transitory component, that represents the deviation from the efficient price due to market microstructure frictions.

The information flow of a given market can be quantified by measures of price discovery. The two most popular measures used in the market microstructure literature are the IS and GG measures, and are defined in Hasbrouck (1995) and Gonzalo and Granger (1995), respectively. In order to compute these measures of contribution to price discovery, we first need to estimate the following VECM of changes in CDS spreads (cds) and equity-implied spreads (eis) for the series of spreads which are non-stationary:

$$\Delta cds_t = \alpha_1 CE_{t-1} + \sum_1^p \beta_{1,j} \Delta cds_{t-j} + \sum_1^p \delta_{1,j} \Delta eis_{t-j} + \varepsilon_{1t}$$
$$\Delta eis_t = \alpha_2 CE_{t-1} + \sum_1^p \beta_{2,j} \Delta cds_{t-j} + \sum_1^p \delta_{2,j} \Delta eis_{t-j} + \varepsilon_{2t}$$

where e1 and e2 are i.i.d. error terms. The cointegrating equation is defined as:

$$CE_t = \alpha\left(cds_t - eis_t \right)$$

We focus on the IS measure because, unlike the GG measure, it takes account of the volatility of the error terms of the VECM. We calculate IS1, IS2, CS1 and CS2 from the error correction parameters and variance–covariance of the error terms, following Baillie et al. (2002). The component shares are obtained from the normalized orthogonal to the vector of error correction coefficients, $\alpha\perp = (\gamma 1, \gamma 2)'$, thus:

$$CS_1 = \gamma_1 = \frac{\alpha_2}{\alpha_2 - \alpha_1}, \quad CS_2 = \gamma_2 = \frac{\alpha_1}{\alpha_1 - \alpha_2}$$

Given the covariance matrix of the reduced form VECM error terms,

$$\Omega = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

and its Cholesky factorization, $\Omega = MM'$, where

$$M = \begin{pmatrix} m_{11} & 0 \\ m_{12} & m_{22} \end{pmatrix} = \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\left(1-\rho^2\right)^{1/2} \end{pmatrix}$$

the IS are calculated then using:

$$IS_1 = \frac{\left(\gamma_1 m_{11} + \gamma_2 m_{12}\right)^2}{\left(\gamma_1 m_{11} + \gamma_2 m_{12}\right)^2 + \left(\gamma_2 m_{22}\right)^2}, \quad IS_2 = \frac{\left(\gamma_2 m_{22}\right)^2}{\left(\gamma_1 m_{11} + \gamma_2 m_{12}\right)^2 + \left(\gamma_2 m_{22}\right)^2}.$$

We introduce the terminology “information leadership” to refer to this metric because it comes from an (unnamed) expression derived by Yan and Zivot (2010) to measure which price series leads the process of adjusting to innovations in the fundamental value. As described earlier, in Yan and Zivot's structural cointegration model CS measures the level of noise in one price series relative to the other, and IS measures a combination of relative level of noise and relative leadership in reflecting innovations in the fundamental value.

Because the IS is defined as a function of the volatility of the error terms in the VECM, a time dependent (daily) IS can be produced by replacing the unconditional error volatilities in (2) with the conditional volatilities obtained with (3). As a result, we can explore the time varying behaviour of the information flow among markets and use it for trading purposes. In order to achieve this aim, for all companies we estimate (1) and (3) by using a rolling window of 1 year of data (250 observations)[11], starting from January 2004. We use the covariance matrix of the error terms (obtained with (3)) at the end of the year to compute the IS measure (the midpoint of the bounds)[12], and we use the latter as an estimate of the price discovery of the CDS market for the following day. The next day, we roll over the 1-year window, we re-estimate (1)[13] and (3) to get a new IS estimate for the following day. We follow this procedure till the end of our sample period, we have a series of estimates of price discovery for the CDS and equity markets for each reference entity. In the next section, we show how to use these estimates to trade both markets.

Trading Strategies

Strategy 1- Standard CSA

The standard CSA is like we mentioned above, it's based on two different time series of data, here would be the market CDS spread and the model spread implied by the equity-based information of a given securities. Capital Structure arbitrage here is implemented on individual securities, also apply on common cases. When these two series of spreads deviate from each other by a threshold value, a trading signal arises. In particular, if the CDS spread is higher than the equity-implied spread by a defined trading trigger θ, we short a CDS position with a notional amount of USD 1 dollar and

$-\delta_{t-1}$ shares of the common stock. Instead, if the equity-implied spread is higher than the CDS spread, we long a CDS position with a notional of USD 1 dollar and at the same time, buy $-\delta_{t-1}$ shares. We also kept the position for a fixed holding period typically unless a convergence occurs between the two time series spreads, then we closed our position and exited the trade.

Strategy 2 - Information Share Augment CSA

With the price discovery signals, x_{lower} and x_{upper} represent, respectively, the lower and upper thresholds of IS price discovery for the CDS market selected by the trade, we could augment the standard CSA. We are filtering CSA trades and executing them only if there is clear evidence of one market leading the other one. However, we still hedge the positions. Hence, if the CDS spread is higher than the equity-implied spread by a defined trading trigger θ and the price discovery of the CDS market is either lower than or higher than , a CDS position with a notional amount of USD 1 dollar and $-\delta_{t-1}$ shares of the common stock are shorted.

On the other hand, if the equity-implied spread is higher than the CDS spread and the price discovery of the CDS market is either lower than or higher than , we long CDS position with a notional amount of USD 1 dollar and $-\delta_{t-1}$ shares of the common stock. Thus, the signals are filtered not only on the basis of the deviation between the two spreads, but also according to the informational efficiency of the markets, captured by the IS measure of price discovery.

Strategy 3 - Information Share Augment CSA without hedging

We also observed that hedging CDS positions with equity shares can be ineffective due to the low correlation observed between changes in CDS spreads and stock prices according to the paper by Yu (2006) during the class. A trade could be more efficient if it trades just one market. To be more specific, we would short the equity market only if the CDS spread is higher than the equity-implied spread by a defined trading trigger θ and the price discovery of the CDS market is higher than x_{upper}, so the equity markets are being led. Similarly, a CDS contract with a notional of USD 1 would be bought if the equity implied spread is higher than the CDS spread by a defined trading trigger θ and the price discovery of the CDS market is lower than x_{lower}.

On the other hand, we would sell a CDS contract with a notional of USD 1 if the CDS spread is higher than the CG model implied spread by a defined trading trigger θ and the price discovery of the CDS market is lower than a benchmark, meaning that the CDS market is being led.

Finally, we only long the shares if the CG model implied spread is higher than the CDS spread by the defined trading trigger θ and the CDS market is leading the equity market. Hence, we trade only on one market, the least efficient one (with a low value of price discovery). By not trading the efficient market, we expect to improve capital allocation due to the efficient market being difficult to forecast.

The following table provides a more formulated equation for understanding the trading trigger signal and which asset class are being traded.

Table. Trading rules of strategies

<table>
<tr><th rowspan="2">Type</th><th rowspan="2">Trading Rule (Trigger Condition)</th><th colspan="2">Asset Class</th></tr>
<tr><th>CDS</th><th>Equity</th></tr>
<tr><td rowspan="2">Strategy 1</td><td>$C^o - C^m \geq \theta_t$</td><td>short</td><td>short</td></tr>
<tr><td>$C^o - C^m \leq -\theta_t$</td><td>long</td><td>long</td></tr>
<tr><td rowspan="2">Strategy 2</td><td>$[C^o - C^m \geq \theta_t]$ and
$[(IS_{cds,t-1} \leq X_{lower})\ or\ (IS_{cds,t-1} \geq X_{upper})]$</td><td>short</td><td>short</td></tr>
<tr><td>$[C^o - C^m \leq -\theta_t]$ and
$[(IS_{cds,t-1} \leq X_{lower})\ or\ (IS_{cds,t-1} \geq X_{upper})]$</td><td>long</td><td>long</td></tr>
</table>

Strategy 3	$[C^o - C^m \geq \theta_t]$ and $[IS_{cds,\,t-1} \leq X_{lower}]$	short	-
	$[C^o - C^m \geq \theta_t]$ and $[IS_{cds,\,t-1} \geq X_{wloer}]$	-	short
	$[C^o - C^m \leq -\theta_t]$ and $[IS_{cds,\,t-1} \leq X_{lower}]$	long	-
	$[C^o - C^m \leq -\theta_t]$ and $[IS_{cds,\,t-1} \geq X_{lower}]$	-	long

Data Description

Due to the fact that we couldn't get the access to any raw datasets from CDS markets, we reused the Boeing.csv from the homework3.

Preliminary Result

The below table shows the average return from our trading strategy 2. The use of an additional trigger on price discovery triggers substantially reduces the frequency of trading by the PD thresholds. Our expectation is that the reduced trade could generate a higher average returns percentage by reducing the risk of large hedging costs against volatile markets and having better forecasting power released from the leading market.

Thus, we compared different levels of price discovery triggers. Intuitively, selecting higher triggers (stronger price discovery in one market) should generate higher profits as the second market would follow the first one more closely, so the second market could be predicted more effectively. However, too high triggers would lead to less profit due to the sharp decrease in the number of transactions and due to profitable trades being left out. We initially chose a value of 80% for the price discovery trigger in the CDS market (corresponding to a 20% trigger for the equity market). And due to the adjustment for the crisis period, we also raised the restriction on the trigger threshold to 5% and 95 % respectively to compare.

The preliminary result of Strategy 2 although didn't produce very significant improvement, yet still has better performance compared to the original CSA strategy. And if we also take into account the transaction cost in reality, we would have a greater impact on reducing the amount of redundant cost to maximize the profit.

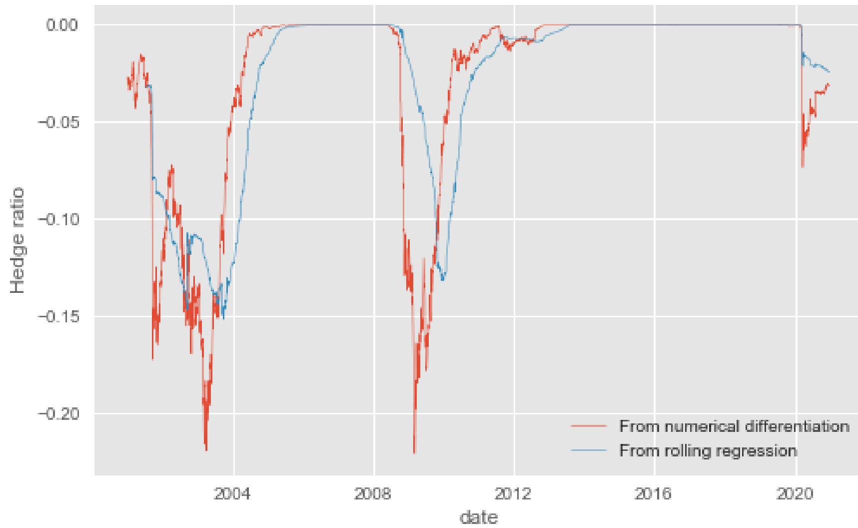
Table. Mean Arbitrage Return (%) of Strategy 2 - PD threshold 0.2

z-score	$x_{lower} = 0.2$; $x_{upper} = 0.8$					
	1w	2w	3w	1m	3m	6m
3	-0.17	-0.26	-0.58	-1.03	-1.93	-1.37
1	-0.21	-0.62	-0.96	-0.06	-0.12	-0.16
0.5	-0.20	-0.37	-0.53	-0.04	-0.10	-0.10

Table. Mean Arbitrage Return (%) of Strategy 2 - PD threshold 0.05

z-score	$x_{lower} = 0.05$; $x_{upper} = 0.95$					
	1w	2w	3w	1m	3m	6m
3	-0.60	-1.43	-1.12	-1.10	-1.86	-1.62
1	0.10	-0.52	-0.58	-0.05	-0.10	0.01
0.5	0.04	-0.58	-0.74	-0.04	-0.08	-0.05

The Strategy 1 and 2 both require hedging positions. For these equity hedges become very expensive, especially during the crisis period, such as 2008 or covids. The figures from homework shows the hedging ratio spike in those periods. Equity hedging costs increase when markets become more volatile and the limits to arbitrage can arise because the liquidity in markets can worsen



From the point of view of implementation, we are not able to perform a complete hedge (as predicted by the hedge ratio calculated with the CreditGrades model) on the days when such an anomaly occurs. Hence, we would need more capital (which becomes a scarce resource) to implement these strategies when volatility in the market is high.

The below table shows the trading strategy 3. The resulting performance is what we're looking for, positive average returns.

Table. Mean Arbitrage Return (%) of Strategy 3 -- PD threshold 0.2

z-score	$x_{lower} = 0.2$; $x_{upper} = 0.8$					
	1w	2w	3w	1m	3m	6m
3	-0.04	-0.04	-0.18	0.01	-0.86	-1.18
1	-0.02	0.11	0.02	-0.02	-0.02	-0.02
0.5	-0.01	0.03	-0.06	-0.01	-0.01	-0.01

Table. Mean Arbitrage Return (%) of Strategy 3 -- PD threshold 0.05

z-score	$x_{lower} = 0.2$; $x_{upper} = 0.8$					
	1w	2w	3w	1m	3m	6m
3	0.10	0.04	0.09	0.03	-0.93	-1.27
1	0.00	0.09	0.02	-0.01	-0.03	0.01
0.5	0.01	0.09	0.03	-0.01	0.00	0.02

Due to the fact that we also don't need to provide hedging, it gives us the dynamics to proxy the market with lower volatility. The below figure also shows the historical arbitrage return across financial crisis and covid periods, we can see that the volatility of strategy compared to CSA strategy decreases significantly especially in these periods. The returns of the strategy were also observed with positive profits more consistently over the time.

Figure. Arbitrage Return of different strategy over periods 2018 - 2021

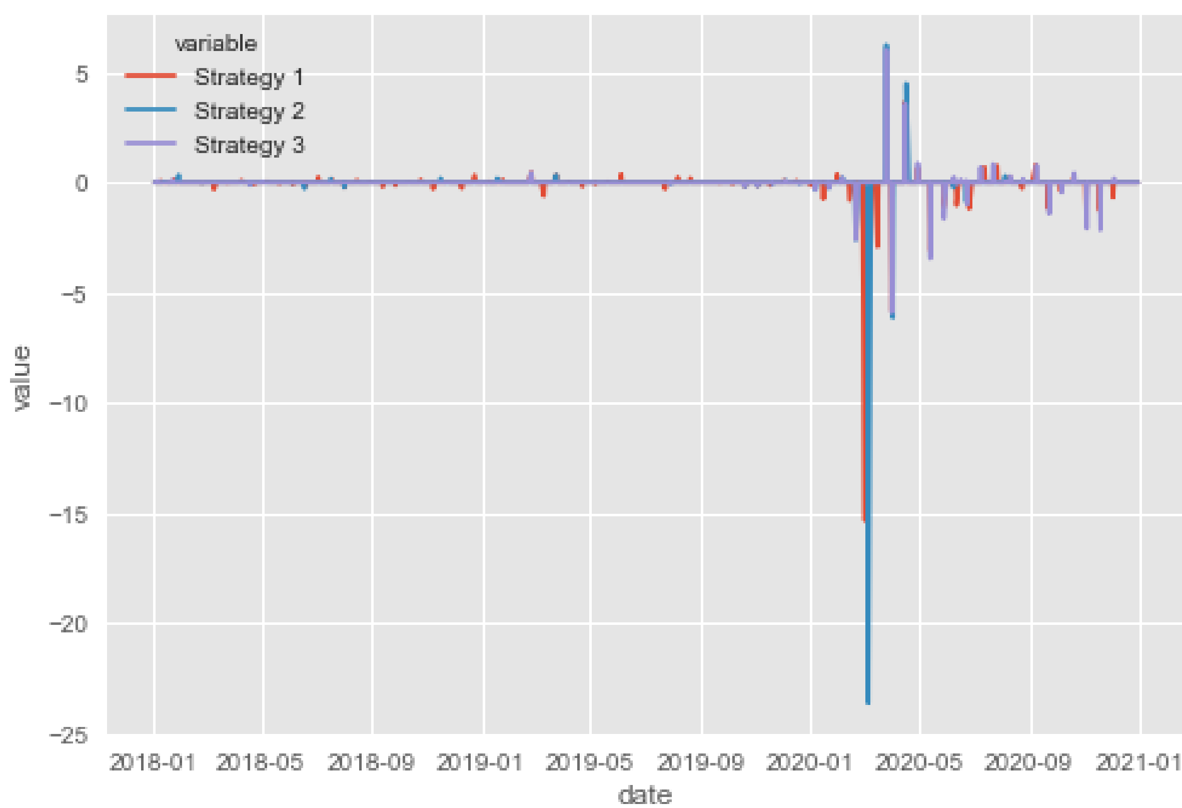
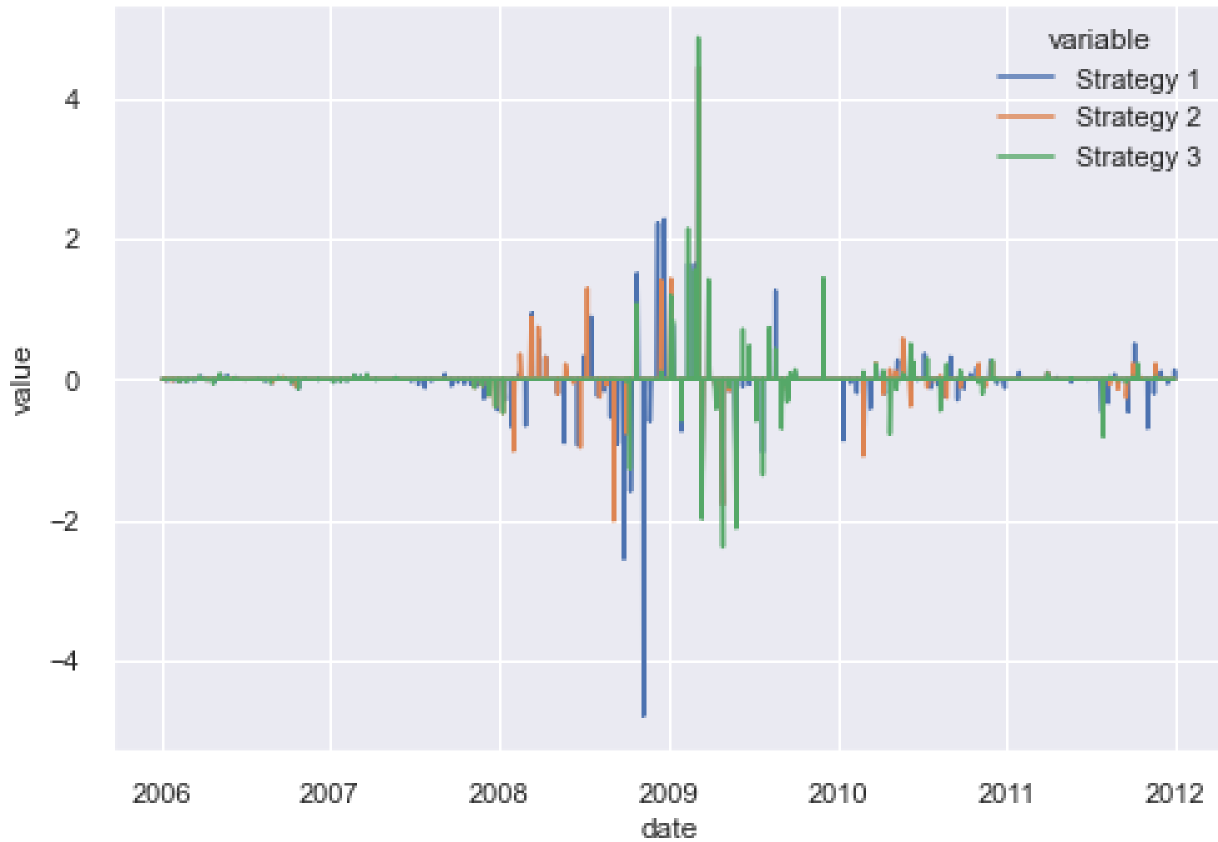


Figure. Arbitrage Return of different strategy over periods 2006 - 2012



Conclusion and Future Application

From the above results, we verify that by combining with the IS factor from price discovery could extend the original limitations of Capital Structure Arbitrage Strategy, such as hedging ratio, anomaly position and trading frequency. Specifically, triggers based on daily price discovery estimates for the two markets are introduced, which allow a filtration framework to be built for more profitability.

Further robustness testing could be investigated through the implementation of applying strategies on different yield and rating grade securities. We could also look into the economy factor, in particular, the TED spread or VIX index to see if our strategy related to dealer funding costs or market volatility.

In sum, the trading strategy that we introduced is based on the information flow and cointegration of the market movement. It can be used to diversify risk in investment fund portfolios at times. Further data mining and variation on the prediction model could also advance the profitability of these strategies. Further research could also focus on innovative strategies based on more advanced price discovery measures.

Reference

1. Tālis J. What do price discovery metrics really measure? Journal of Empirical Finance
2. Longstaff et al. (2003) Corporate Yield Spreads: Default Risk or Liquidity? New Evidence from the Credit-Default Swap Market, previously titled: "The Credit-Default Swap Market: Is Credit Protection Priced Correctly
3. Baillie, R.T., Booth, G.G., Tse, Y., Zabotina, T., 2002. Price discovery and common factor models. Journal of Financial Markets 5, 309-321.
4. Blanco, F., Brennan, S., Marsh, I.W., 2005. An empirical analysis of the dynamic relationship between investment grade bonds and credit default swaps. Journal of Finance 60, 2255-2281.