PROTEUS

PCA Models of Interest Rates: Potential Chaos Awaits

Why Arbitrage-Free Models are Needed for Risk Management

James P. Norman PhD FIA, Proteus Consulting LLP

27 March 2024

Introduction

Factor-based models such as Principal Components Analysis (PCA) of yield curve movements are used by some financial firms to assess interest rate risk. While appearing statistically reasonable, these approaches have an important and serious flaw which can mislead firms into unknowingly taking risky descisions. In the worst-case, a firm using such an approach could set their capital requirement arbitrarily – even to negative values – while in reality the firm may be exposed to risk. The flaw is that these approaches are not *arbitrage-free*. The importance of using arbitrage-free models is well known in the fields of derivative pricing and valuation, however it seems less well appreciated in risk-management applications such as capital modelling. One reason for this appears to be a general confusion around terminology. Another is a lack of literature and other materials on this topic.

In this article, I aim to clarify the terminology of arbitrage as it relates to interest rate modelling and its relevance to real-world risk management applications. I will then show a simple example of a portfolio of bonds and fixed liabilities where an apparent arbitrage opportunity suggested by a PCA model of yield curves might lead the insurer to adopt a mis-matched asset position. This position appears to be better than risk-free with no downside risk, but shows as risky when viewed with an arbitrage-free model. Finally, I outline a method for testing whether a scenario set of yield curve movements – which may have been produced by any model – contains these simple static arbitrage opportunities.

What is arbitrage?

Put simply, an arbitrage is an opportunity to make a return that is guaranteed in all possible circumstances to be at least the risk-free rate of return, and sometimes exceed it. True arbitrage opportunities, particularly in liquid markets such as government bonds, are rare and fleeting and it would certainly be imprudent to rely on exploiting them in a model used for risk management. This is why it is felt that interest rate models – and indeed financial models in general – should be arbitrage-free. However, an interest rate model which has not been constructed correctly can present an apparent arbitrage opportunity which may be unintentionally exploited to reduce modelled measures of risk.

Mathematically, an arbitrage is present if there is a portfolio of assets which has zero value at time $t = 0$, and has both a zero probability of a loss and positive probability of having a value greater than zero at some future time $T > 0$. That is, an arbitrage exists if a portfolio can be found with value at time t, $V(t)$, such that

$$\begin{aligned}
&V(0) = 0, \\
&P[V(T) < 0] = 0, \text{ and} \\
&P[V(T) > 0] > 0.
\end{aligned}$$

For simple models such as a model of a stock it is relatively straightforward to see whether the model would allow for an arbitrage. If the return on the stock is always at least the risk-free return and sometimes more, then there is an arbitrage opportunity in the model. For interest rate models the situation is more complicated as there are many assets – bonds of different maturities – driven by a (typically) small number of factors. Avoiding arbitrage means ensuring the joint dynamics of the bonds is such that any portfolio of these assets does not earn more than the risk-free rate of return while not having any chance of loss. While the details are beyond the scope of this article, arbitrage-free models are carefully constructed to ensure – and can be mathematically proven to show – this does not happen.

So far, so good. Unfortunately, most of the rest of this section is devoted to clearing up confusion about what arbitrage is *not*.

The concept of arbitrage became crucial in the field of derivatives pricing, and most textbooks treat the problem from that perspective which can give the impression that arbitrage is *only* relevant in that field. However, arbitrage is a more fundamental concept. It is not a concept that is only relevant to derivatives pricing. It is a concept that is relevant to all financial models. Without going into technical details, it turns out that a model is arbitrage-free if there *exists* an abstract quantity called an "equivalent risk-neutral probability measure". All this means is that all asset prices in a model must be able to be expressed as discounted expected values of future cash flows, using a common set of some alternative probabilities, known as risk-neutral probabilities. By "equivalent" we mean that these alternative probabilities and the original real world probabilities agree on what is certain (have probability equal to one), and what is impossible (have probability equal to zero). But probabilities of anything in between can be different, so long as they sum to 1.

In derivatives pricing one often does not need to consider the real-world probabilities at all. This leads to the second point of confusion – the term "arbitrage-free" is often interpreted to mean that the model is *only* arbitrage-free when simulated using these risk-neutral probabilities. This is not the case. The mathematical definition of an arbitrage uses real-world probabilities. If an arbitrage does not exist with real-world probabilities, then it is not an arbitrage! For risk management applications such as calculating Value at Risk or use in a capital model, real-world probability distributions are needed but the models should also be arbitrage-free. The model is arbitrage-free if the risk-neutral probabilities *exist* and are *equivalent* to real-world probabilities, but the evolution of the portfolio value over time still happens in the real-world, with real-world probabilities.

Interest rate models which are arbitrage-free must be constructed ensuring that the prices of bonds at any point in time are discounted expected values of their cash flows under the risk-neutral probabilities. They must also construct the relationship between real-world and risk-neutral probabilities in a way which ensures that the two sets of probabilities are equivalent. Examples of interest rate models which are arbitrage-free are the Cox-Ingersoll-Ross model, the Hull and White model, and the affine term structure class of models. [1]

[1] Even more confusingly, there is a similar term in use for interest rate models which can be calibrated to exactly fit an initial yield curve. These are known as "no-arbitrage" models. Being a "no-arbitrage" model is distinct from whether the model is "arbitrage-free". However, most arbitrage-free models can be adjusted to fit the initial yield curve exactly anyway, so we will not divert ourselves further with this subtlety.

Interest rate models which are not constructed in this way may not be arbitrage-free. A naive model based on Principal Components Analysis (PCA) explained below is one such model which is not arbitrage-free. A typical PCA-based model is a statistical approach which models changes in the yield curve as a sum of changes in a number of independent orthogonal factors. Typically, there are 3 factors corresponding to "level", "slope" and "curvature" of the yield curve, but more (or fewer) may be used. Simulations of the change in the yield curve are obtained by drawing a random variable for each factor and multiplying it by the "factor loading", or sensitivity, of that factor for each term of the yield curve. The factor movements are then summed at each term to obtain the modelled change in yield curve. The yield curve change is then added to the starting yield curve to get a simulation of the future yield curve [2]. The success of PCA is primarily due to its ease of implementation — most statistical packages have functionality for performing the required calculations. PCA models are also quite easy to understand and communicate as they reduce the complexity of yield curve movements into a small number of interpretable factors. Further, PCA models typically provide a good match to the statistical properties of yield curve changes.

While this article focuses on PCA as a non-arbitrage free model, there are many other examples of models which are not arbitrage-free. For example the Dynamic Nelson-Siegel model, which although performs well at forecasting yield curve movements in a purely statistical context, allows for arbitrage profits. Models which *may* not be arbitrage-free include empirical bootstrapping approaches, and models based on machine-learning approaches.

[2] There are several variants on this, such as using a spot-curve or forward curve, or allowing for mean-reversion, or using different statistical distributions. Also, some internal models may ignore the passage of time, and apply the shock as-if they happened instantaneously. However, these variants usually still admit arbitrage. See Norman, James P. "Real world interest rate modelling with the BGM model." Available at SSRN 1480174 (2009).Norman, James P. "Real world interest rate modelling with the BGM model." Available at SSRN 1480174 (2009), for an example of how a PCA-based model can be made arbitrage-free.

True vs Statistical Arbitrage

A related concept to arbitrage is *statistical arbitrage*, also known as risk arbitrage. Statistical arbitrage is not a true arbitrage but rather refers to an opportunity for making a profit in excess of the risk free rate which is large in relation to the amount of risk being taken. In reality true arbitrages are rare, and the trades of arbitrageurs in the markets are exposed to risks and are more akin to statistical arbitrage. Statistical arbitrage is sometimes profitable and sometimes (catastrophically) not, and transaction fees, other expenses, and costs of capital may reduce any excess return considerably.

Many statistical models even if technically arbitrage free allow for absurdly large excess returns relative to the risk involved. In extreme cases a statistical arbitrage can be so large that for all practical purposes it becomes a true arbitrage, for example if the probability of loss is so small that a loss would not be expected to occur in the lifetime of the universe. While this would not technically be a true arbitrage, in real applications it is effectively the same. Using a model which is arbitrage-free, and has clearly identified risk premia, can also guard against excessive statistical arbitrage, as these large risk premia will be apparent in the relationship between the real-world and risk-neutral probabilities.

Why does it matter for risk management?

Arbitrage generates money from nothing. It is a free lunch. So using a model which is not arbitrage-free can lead to a firm thinking it can both reduce its risk and make more money, when in fact it is increasing risk. It is often not easy to see that this is happening.

Let us take a simplified example of an insurer with liability cash flows which are fixed, holds an asset portfolio of risk-free zero coupon bonds, and is only exposed to interest rate risk. We assume that the liabilities can be discounted at the risk-free rate. The insurer starts with a portfolio of zero coupon bonds which exactly matches the cash flows of the liabilities. The cash flow profiles are shown in Figure 1:

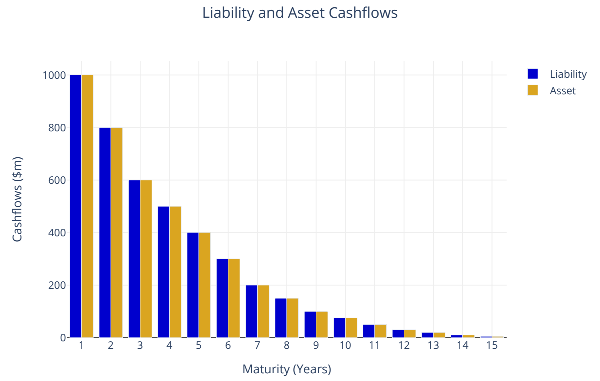

Figure 1: The initial asset and liability cash flows

In this example the liabilities are denominated in US dollars and the assets are also denominated in US dollars. The analysis is performed at year end 2023. The opening zero-coupon yield curve, taken from the US Federal Reserve, is shown in Figure 2.

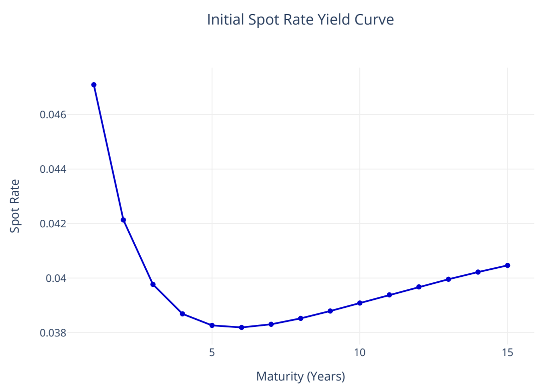

Figure 2: The initial zero coupon yield curve, as at year end 2023. Source: Board of Governors of the Federal Reserve System (US), Nominal Yield Curve https://www.federalreserve.gov/data/nominal-yield-curve.htm

The principal components are obtained from analyzing monthly changes in US zero coupon yield curves since 1990. The first three principal components are found to explain over 99.7% of the variance of the data. The factor loadings of the three principal components shown in Figure 3 display the classic “level”, “slope” and “curvature” shapes.

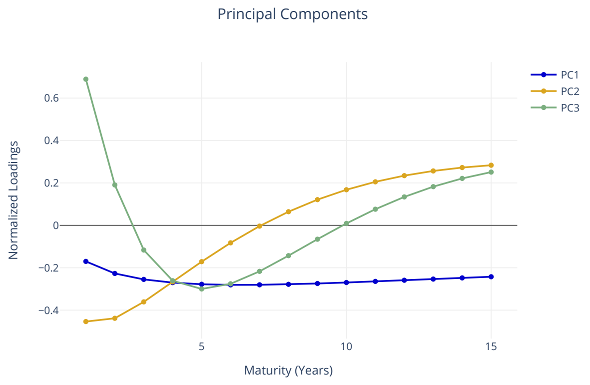

Figure 3: The first three principal components from the principal component analysis of changes in the yield curve, using monthly observations from 1990-2023. Source: Board of Governors of the Federal Reserve System (US), Nominal Yield Curve https://www.federalreserve.gov/data/nominal-yield-curve.htm and authors own calculations

The insurer uses a PCA model to produce simulations of future yield curves one year into the future, and then uses these projections to calculate the present value of the liabilities and the assets at the end of the year.[3]

The initial net asset position is zero. In this case, year end net assets are also zero under every scenario. This is expected as the assets perfectly replicate the liabilities.

However, let us now imagine that the company wants to try to make some extra profit, is willing to move away from the perfectly replicating portfolio, and wants to know how much risk and extra capital they would require by doing so. The company determines their capital requirement from the 99.5th percentile of the negative of the net asset value distribution over a one-year time horizon [4] . Amongst others, they decide to trial the following asset portfolio:

[3]In this article I draw the simulations of the factor loadings from a Normal distribution, however similar results are obtained using Student's-t, Log-Normal or other distributions.

[4]Nothing of substance in this example turns on the particular risk metric, tolerance or time horizon. Similar results would be found for TVaR, for example.

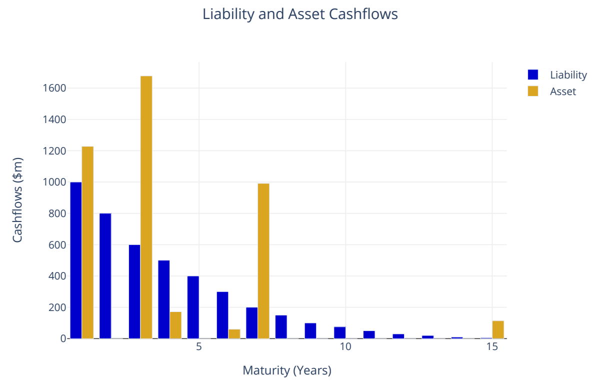

Figure 4: The cash flows of the trial asset portfolio of zero coupon bonds, along with the liability cash flows

The initial net asset position is still zero. The analyst runs the portfolio through the risk model using the same interest rate scenarios as before. The results are shown below.

Interestingly, using the new asset portfolio, the insurer manages to achieve a positive expected profit, while their capital requirement is slightly negative! In fact, all of the simulations from the internal model show a positive profit for the insurer.

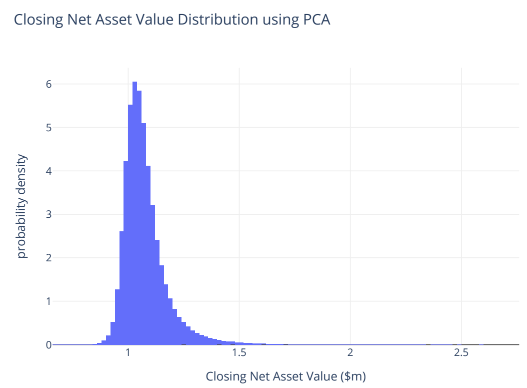

Figure 5: The distribution of the closing net asset value under the alternative asset allocation, using the Principal Component Analysis-based interest rate model

However, alarm bells should be ringing at this point. The insurer appears to have found a free lunch. They have started from a portfolio of assets and liabilities valued at zero, and end up after one year

Table 1: Results from the Principal Component Analysis based interest rate model

Metric	Original Portfolio	New Portfolio
Starting Net Assets ($m)	0	0
Expected Ending Net Assets ($m)	0	1.08
Capital Requirement ($m)	0	-0.91
Minimum Net Assets ($m)	0	0.68

with a portfolio of assets and liabilities which never makes a loss and in fact makes a profit in most cases. This is an arbitrage, and the reason it occurs can be traced back to the PCA model not being arbitrage-free.

What if instead the insurer had used an arbitrage-free model, such as the well known affine term structure model used in the Proteus Scenario Generator? Using a 3-factor affine term structure model in place of the PCA model, it is clear that the insurer would indeed be exposed to interest rate risk with this new asset portfolio, and should be holding capital against the risk. We should emphasize that the affine term structure model is arbitrage-free and also uses real-world probabilities for the evolution of interest rates.

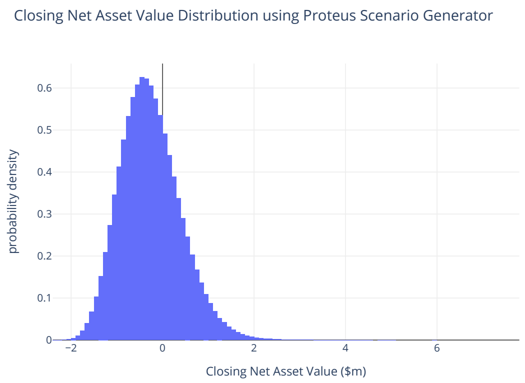

Figure 6: The distribution of the closing net asset value under the alternative asset allocation, using the arbitrage-free affine term structure model from the Proteus Scenario Generator

So while PCA suggests this asset portfolio is better than risk free, the arbitrage-free affine term structure model shows that the insurer is exposed to interest rate risk.

Which model is correct?

Neither. “All models are wrong, but some are useful”, said the statistician George Box. The yield curve does not evolve according to a Principal Components Analysis-based model, an affine term structure model, nor any model that we can conceivably write down. Models are always simplifications of reality. They are used because they can be useful; for example to help our understanding, or to help decision making. The usefulness of a particular model depends on the use to which it is being put. Non-arbitrage free models can be useful for forecasting the yield curve, for example, if you view the yield curve as just a set of numbers to forecast. However the yield curve describes asset prices. If the user is making financial decisions aided by the model output then the fact that the model allows for arbitrage opportunities should weigh heavily against using that model, even if otherwise the model is viewed as reasonable.

One of the most important uses of an internal model is to compare the risk/reward trade-off of alternative asset portfolios. Using a model which allows arbitrage opportunities means that the model must be used with extreme caution for this purpose. It will always favour portfolios which contain an element of the arbitrage portfolio. Decisions taken based off risk/reward trade-offs will tend to steer the portfolio towards an arbitrage portfolio. Because the arbitrage portfolio has zero cost, you can add any amount of it to a portfolio, and it will appear to reduce the capital requirement of the portfolio, and increase profitability. Taken to it’s extreme, if the insurer could leverage enough, perhaps by using derivatives such as swaps and forwards, their capital requirement can be made arbi-

Table 2: Results from the arbitrage-free affine term structure model from the Proteus Scenario Generator

Metric	Original Portfolio	New Portfolio
Starting Net Assets ($m)	0	0
Expected Ending Net Assets ($m)	0	-0.26
Capital Requirement ($m)	0	1.67
Minimum Net Assets ($m)	0	-2.39

trarily negative [5]. The danger is that in more realistic situations the arbitrage will not be obvious.

What if the model is not used for decision making, but merely to report on the current risk position? In reality, insurers are rarely perfectly matched, so there is always a question about how much the arbitrage opportunity affects the modelled risk metrics. I would argue that even in these situations, arbitrage-free models should be preferred over models with arbitrage opportunities.

Testing for arbitrage opportunities in interest rate models

Mathematics can prove whether a particular model is arbitrage-free or not. However, sometimes a model is arbitrage-free in theory, but its implementation may not be, for example due to approximations or errors. Sometimes full details about the model or its workings are not available. Alternatively the outputs of an otherwise arbitrage-free model might be adjusted, for example by scaling or imposing a lower bound on interest rates which is not present in the model. Suppose that an analyst receives a scenario set containing simulated yield curves from a third party, such as a vendor or consultancy, or possibly another department in their company, and would like to know if they allow for the arbitrage opportunities discussed above. In this section, a simple test for whether a scenario set contains a static arbitrage over a single time step is presented. This method can be used on both real-world and risk-neutral scenario sets.

[5] Interestingly similar situations can occur when using stress or scenario approaches such as the interest rate risk component of the standard formula SCR. This could be avoided to some extent by basing the stress scenarios on an arbitrage-free model.

Firstly, the user should calculate the returns over the period on a zero coupon bond of each maturity of interest i, in each simulation j in the scenario set. Call the return on the zero-coupon bond of maturity i in simulation j, r_{ij}. The test is then to attempt to maximize the expected return on a portfolio of these zero coupon bonds, subject to constraints that ensure that the initial value of the portfolio is zero, and a second set of constraints to ensure that the final portfolio value is at least zero in each simulation. Mathematically, for a portfolio with initial market value w_i in the zero-coupon bond of term i, the optimization problem is to find w to:

$$\text{maximise } \sum_{i,j} w_i r_{ij}$$
$$\text{subject to}$$
$$\sum_i w_i r_{ij} \geq 0$$
$$\sum_i w_i = 0$$

It is also convenient sometimes to put maximum and minimum bounds on the portfolio weights w.

This is a constrained linear optimization, and can be solved quickly and robustly on a computer using linear programming routines, even for scenario sets with millions of simulations. If a solution to the optimization can be found, then that suggests an arbitrage opportunity may be present in the model, and the solution w gives a portfolio which produces that arbitrage. Of course if the scenario set contains only a small number of simulations, this could just have been bad luck and a larger simulation set may have produced scenarios which do in fact cause a loss. However for large numbers of simulations, the presence of a solution means that the model effectively contains an arbitrage opportunity.

If the optimization routine does not find a solution, then the scenario set does not allow this type of static arbitrage. However, one should keep in mind that this test will not identify more complicated dynamic trading strategies which might present other arbitrage opportunities.

The arbitrage solutions found will typically contain negative positions in some bonds and positive positions in others. In the context of a financial institution, liabilities provide the negative positions. A further refinement of the above is to restrict the portfolio weights so that the asset position is positive at each duration[6]

The interest rate scenarios sets produced by the affine term structure model in the Proteus Scenario Generator pass this test.

Summary

I have shown above some of the dangers of using models which are not arbitrage-free in real-world modelling applications, such as an internal model used for risk management and capital assessment. While models such as PCA may appear statistically reasonable, and in some circumstances — like a stopped clock — can give reasonable results, they can encourage the firm to take incorrect and potentially dangerous decisions because they contain arbitrage opportunities.

Whether a model is arbitrage-free or not is important not just for derivatives pricing and valuation, but also for real-world risk management applications. Our example shows that arbitrages in interest rate models can occur in even the simplest setting of a static portfolio of zero-coupon bonds and fixed liabilities. I have not needed to use complicated dynamic trading strategies or non-linear derivatives for them to appear.

Interest rate models which are arbitrage-free have been available for decades. With modern approaches to affine term structure models, as used in the Proteus Scenario Generator, they are able to provide a good statistical description of yield curves as well as not producing absurd results due to arbitrages. While they may be more mathematically involved, the importance of using a theoretically rigorous model becomes apparent when using the model for decision making.

[6] This is, of course, how I found the example portfolio in the previous section.

About Us

Proteus Consulting is a cloud provider of risk models for use across the financial services community. We generate distributions of economic, asset and other risk types for subsequent upload or linkage into your capital model, pricing application or other risk aggregation systems.

The Proteus Scenario Generator is an arbitrage-free financial market and economic scenario generator, designed for both real-world risk management and pricing/valuation applications. The PSG can be used to generate arbitrage-free market risk scenarios with corresponding portfolio valuations.

If you are a user of scenarios from a model and you are not sure whether the model is arbitrage-free, Proteus Consulting can help you by testing the scenarios for various types of arbitrage. We can also provide you with alternative scenario sets from arbitrage-free models.

To find out more, please contact us at:

james.norman@proteusllp.com

Disclaimer

The information and results presented here are based on models and analyses conducted by Proteus Consulting LLP. While every effort has been made to ensure the accuracy and reliability of the data and insights provided, it is important to acknowledge the inherent limitations and uncertainties associated with statistical and financial modeling.

1. Interpretation: This article and the results generated by the models are intended for informational purposes only and should not be construed as definitive conclusions or predictions. They represent estimates based on available data and assumptions at the time of analysis.
2. Limitations: This article and the model may not account for all relevant variables or factors that could influence the outcomes. Additionally, it is subject to the quality and completeness of the input data, as well as the assumptions and methodologies used in the modeling process.
3. Uncertainty: There is inherent uncertainty associated with any model, and the results may vary depending on different modeling approaches or data inputs. Therefore, the findings should be interpreted with caution and considered in conjunction with other sources of information and expert judgment.
4. Not Financial or Legal Advice: This article and/or model results do not constitute financial, investment, legal, or professional advice. Users are encouraged to consult with qualified professionals for specific guidance or decisions related to their individual circumstances.
5. Liability Waiver: Proteus Consulting LLP disclaims any liability for errors, omissions, or inaccuracies in this article and/or model results or any decisions made based on them. Users assume full responsibility for their use of the information provided herein.
6. Modification and Updates: Proteus Consulting LLP reserves the right to modify, update, or discontinue the model or its results at any time without prior notice.

By accessing or using this article and/or model results, you agree to the terms and conditions outlined in this disclaimer. If you do not agree with these terms, please refrain from using the information provided.