Lee Overlay Partners

CURRENCY OVERLAY MANAGEMENT

Research Report

A discussion on allocation rules, and a recommended approach for use in currency overlay

Lee Overlay Partners

July 1999

Reference Number: lop-0799b

A Discussion on Allocation Rules, and a Recommended Approach for Use in Currency Overlay.

Abstract

In this article we address the problem of allocating amongst alternative currencies in an active management process. We suggest a set of criteria that are designed to avoid many of the pitfalls that can befall an allocation rule, and provide examples of rules used in practice that succumb to such pitfalls. We emphasise how easy it is to inadvertently create poor allocation rules and provide a recommended approach that satisfies our criteria thereby avoiding some of the common allocation problems.

1. Introduction

Making allocation rules to convert investment views into positions is not a trivial exercise. The existence of a broad variety of techniques in the financial industry is testament to this fact. The problem is difficult enough when allocating between a universe of domestic assets, but when the universe is global, so that currencies and perspective (i.e. denomination of domestic currency) enter the equation, the problem becomes even harder. This issue appears in its most distilled form in the discipline of currency overlay where a portfolio of currency exposures is created to beat a benchmark of currency exposures.

In this article we consider the problem of allocating between alternative currencies and provide a recommended approach that satisfies a set of appealing and intuitive criteria. We begin by defining these criteria and give an explanation as to why they are important. We then develop an allocation rule for currency overlay that satisfies these criteria and relate the process to other rules that are commonly used.

We will focus on the portfolio exposures in excess of benchmark exposures, which we term active positions, or just positions. The portfolio exposures and benchmark exposures are fully invested so the positions must sum to zero.

2. Criteria for Allocation Rules

The most important point in thinking about allocation rules is *that no allocation rule, irrespective of its sophistication, can turn a bad forecast into a profit making position.* The focus should instead be in generating sensible positions that are consistent with the forecasts. The emphasis here is on sensible. What follows is a set of criteria that enable us to achieve this objective:

The allocation rule should…

1. yield intuitive positions
2. be consistent from any country's perspective
3. have only limited dependence on required inputs other than the forecasts
4. deal with portfolio constraints in a consistent and systematic way
5. be broadly in line with theoretical concepts
6. exhibit low sensitivity to small input changes

2.1 Yield intuitive positions

It seems obvious, and almost redundant, to include a criterion desiring intuitive positions. Yet it is amazing how quickly intuitiveness of output is lost as algorithms become more complex. A prime example is the use of correlations where an increase in forecast return can lead to a reduction in recommended position because of the correlations with other currencies in the universe.

2.2 Be consistent from any country's perspective

The desire for consistency stems from the fact that the return from a portfolio of active positions is the same from any perspective (see Section 3.1 for proof). Since the portfolio return is independent of perspective then the positions from the allocation rule should also be independent of perspective:

It is surprisingly easy to create an inconsistent allocation rule. We cite an example of a commonly used approach that is inconsistent from different perspectives.

Consider expected return forecasts made relative to some base currency (the Dollar for example). Variances are also provided, and to avoid dependency problems the correlations are set to zero. The positions are generated by the commonly used "mean-variance" optimiser that finds the positions that maximise the expected return of the active position portfolio subject to a targeted level of portfolio variance.

Now a Yen perspective is required so the expected returns are computed from the new base currency (by subtracting the Dollar-Yen forecast from the original forecasts). The variances are also rebased, and again the correlations are set to zero. The positions are generated from the same mean-variance optimiser as the above.

A different set of positions will be generated in each case. This is because if the correlations are zero for the Dollar based forecasts, they cannot logically also be zero for the Yen based forecasts. Consistency is only maintained by deriving the values of the Yen based correlations from the Dollar based data. Non-zero correlation values will be derived and must be used in the optimiser in order to maintain consistency. The problem in doing this is that the intuitive nature of the output by using zero correlations from a Dollar perspective will be lost when working from a Yen perspective.

2.3 Have only limited dependence on required inputs other than the forecasts

It is tempting to incorporate as much sophistication as possible in the allocation rule under the assumption that a more optimal strategy will result. However, the greater the sophistication, the more inputs are required by the user. Typically the user must estimate these parameters and the output is therefore a function of how good the parameter estimates are. Using variances and correlations, for example, requires that they be estimated. It is important that these estimates do not become the drivers of the output. There is a classic example of this:

Consider forecasting two currencies that are very similar, the Euro and the Swiss Franc for example. The historic correlation of monthly returns of Euro/Dollar and Swiss Franc/Dollar exchange rate is calculated to be 0.95. Assume they have the same variance. Consider the situation where the Euro is forecast to appreciate by 4.5% and the Franc by 4.3%. The fact that the two forecasts are so similar seems consistent with the high correlation, but the mean-variance optimiser will recommend a huge long Euro position versus a huge short Franc position. The optimiser will drive its recommendation off the fact that the correlation is so high. Mathematically, it is possible to create a virtually risk free portfolio by holding two offsetting positions in such highly correlated currencies. Moreover the portfolio will have a positive forecast return because the individual forecasts are different.

This effect is normally undesirable, particularly given the accuracy of the forecasts. More appropriate may be two long positions versus the Dollar, with the Euro position marginally longer than the Franc. This would have been achieved by assuming a zero correlation.

The above example illustrates the problems with allowing the allocation rule to rely on estimated inputs. It also illustrates how important it is to fully understand the nature of the inputs that need to be estimated. We recognised that we needed to include a correlation in our optimisation and so we used the historic data to estimate it. However, the correlation that is required for the optimisation is the correlation of return *conditional on the expected return being known*. The correlation we calculated from historical data is the unconditional correlation. The missing piece is the correlation between the expected return forecasts. It is this correlation that may be the high one, implying that the expected return forecasts are always similar, which would be an appealing result. The unconditional correlation is a function of the conditional correlation and the correlation of forecasts. If the unconditional correlation is high but so is the correlation of forecasts, then the conditional correlation required for the optimiser may actually be close to zero. In any event a much deeper statistical analysis is required to develop the required correlation estimate, and the positions will be just as reliant on the accuracy of this forecast as the return forecast. It is more beneficial to bias the allocation rule to the return forecasts by using simplifying assumptions in place of other parameters.

An additional problem with extra inputs in a process is that as they change the resultant positions will change, even if the return forecasts themselves remain the same. For example, using implied volatilities in a mean-variance optimisation to circumvent the parameter estimation problem can result in frequent and significant position changes with an unchanging return forecast. Simply because the implied volatility itself is a volatile number. It is even possible to have a position reduction recommendation with an increase in return forecast because the volatility input increased more.

2.4 Deal with portfolio constraints in a consistent and systematic way

All of the problem examples cited above involve mean-variance optimisation approaches to allocation. A solution is to use a different approach. A well-used alternative that avoids many of the aforementioned pitfalls is the linear allocation rule. This rule recommends that the positions be proportional to the forecasts. One problem with this approach is how to deal with constraints. As we shall see, constraints are a serious problem and it is worth noting that optimisation approaches deal with this problem with relative ease.

It is difficult to derive a systematic algorithm the reallocate in the event of binding constraints. An obvious solution to upper and lower position bounds is to simply truncate the positions at the bounds if they would otherwise exceed them. It is a deceptively difficult problem though, for what happens if by truncating the positions another limit is broken? This can happen since all the positions must sum to zero, and by truncating one position another must be adjusted to compensate. Additionally, constraints involving combinations of currencies (such as the sum of the European currency positions not exceeding a certain amount) are very difficult to deal with. Algorithms to reallocate in the event of constraints are very hard to write, tend to be heuristic, rarely cover all possible scenarios, and often violate the consistency from any perspective criterion.

2.5 Be broadly in line with theoretical concepts

There exists a large body of literature dealing with portfolio construction. Central to this is modern portfolio theory. This theory is the underpinning to the mean-variance optimisation that is used so widely in practice. However, modern portfolio theory assumes that the positions to be optimised do not change, implying a buy and hold investment strategy. For active processes such as currency overlay, this assumption is not valid as the positions change frequently through time. Mean-variance optimisation on an ongoing basis may not produce the long run optimal results, see for example Lee Overlay Partners (1999). For this reason the mean-variance optimisation may not be the best approach. Dynamic investment theory does is exist but is nothing like as well used in practice, and has a far higher level of complexity than the standard mean-variance approach.

Given the lack of a simple, well-known, theoretical approach to optimal investing through time, many variations to the mean-variance approach have been created. Given other criteria of concern, using alternative allocation rules may better suit an individual need. However, there is a safety net in remaining close in spirit to the academic theories in that the allocation rule will be built on proven principles. Straying too far from the concept of maximising return subject to risk carries with it its own conceptual risks.

2.6 Exhibit low sensitivity to small input changes

One of the well-known drawbacks to the mean-variance approach is its sensitivity to changes to the inputs. Consider the Euro and Swiss Franc example cited in Section 2.3. A change in the Swiss Franc forecast by +0.2% and in the Euro forecast by -0.2% will result in a complete reversal of the position recommendation. The huge long Euro position will become a huge long Swiss Franc position, and correspondingly the huge short Swiss Franc position will become a huge short Euro position.

To some extent the sensitivity of the optimisation can be removed by careful choice of inputs (zero correlations for example). However, it is probably fair to say that the primary reason why an alternative to mean-variance optimisation is used is the input sensitivity problem.

For all its constraint-related shortcomings the linear rule is the superior approach, of the standard techniques, if stability of positions is an important issue.

3. Creating an allocation rule

In this section we create an allocation rule that satisfies the criteria we developed earlier. We outline the statistical framework and provide some mathematical results central to our problem. We develop an objective function and recommend some key features to aid criteria satisfaction. We discuss methods of reducing the unwanted degrees of freedom and provide suggestions for dealing with parameter estimation.

3.1 A statistical view of the basic notions

Suppose we have a vector of future returns. Let **R** denote their joint distribution. The portfolio return distribution, denoted by P, is given by $P = \mathbf{w}'\mathbf{R}$ where w denotes the vector of positions. Recall that these are active positions, so $\mathbf{w}'\mathbf{1}=0$ where **1** is a vector of ones.

To consider the above example from a different perspective we must subtract the new perspective versus old perspective return, denoted by the scalar r_0, from each of the returns. The portfolio return distribution from the new perspective, denoted by P_0, is therefore given by $P_0=\mathbf{w}'(\mathbf{R}-\mathbf{1}.r_0)$. Simple algebra gives us that $P_0=\mathbf{w}'\mathbf{R}-\mathbf{w}'\mathbf{1}.r_0$, but $\mathbf{w}'\mathbf{1}=0$ and therefore $P_0=P$.

The conclusion is that the portfolio return is independent of perspective and therefore the positions must be consistent from any country's perspective.

This result invalidates approaches based around ideas such as taking positions in proportion to the forecasts converted to the appropriate perspective.

The theory that the portfolio return is the same from any perspective stems from the fact that a scalar can be subtracted from the vector of returns without affecting the portfolio return. We take advantage of this property and subtract the average of the expected returns from the vector of returns. As before, the portfolio return is unaffected by this adjustment and this vector of adjusted returns has the property of being identical from any perspective. The obvious benefit is that a linear rule based on these adjusted forecasts is indeed consistent from any country's perspective.

3.2 Refining the process

Dealing with constraints in a consistent and systematic way is a tremendously difficult exercise. For this reason we recommend using an optimisation platform where constraints are dealt with more easily. In choosing optimisation, above a linear rule, we must be cognisant of the issues surrounding stability of output.

Our ability to satisfy our criteria depends on what objective function is used in the optimisation. To ensure we are broadly in line with theoretical concepts we will assume a quadratic, or mean-variance, objective function. Using the result in Lee Overlay Partners (1999) for investing through time, we will replace the variance term with the second moment to be more optimal in maximising the information ratio. By choosing optimisation above linear rules we are also ensuring a closer link with theoretical concepts.

Using a mathematical notation, we are attempting to maximise

$$f(P)= E[\mathbf{w}'\mathbf{R}] - \tfrac{1}{2} A\, E[(\mathbf{w}'\mathbf{R})^2]$$

which can be rewritten as

$$f(P)=\mathbf{w}'\mathbf{m} - \tfrac{1}{2} A\, \mathbf{w}'\mathbf{V}\mathbf{w}$$

where **m** denotes the vector of forecasts and **V** denotes the second moment of the returns. The parameter A denotes the aggressiveness of the active management. This parameter is estimated outside of the allocation framework and represents the risk profile of an individual investor.

At this point the remaining criteria (stability, intuition, parameter dependence) are satisfied by careful choice of **V**. Central to this effort is the fact that many properties of **V** when viewed from one perspective are lost when converted to another perspective.

Assuming **V** to be diagonal is a classic example of this, since it will not be diagonal from another perspective. Forcing it to be diagonal by ignoring the off diagonal elements (i.e. ignoring correlations in the analysis, from any perspective) results in different positions from different perspectives using the same forecasts.

The adjustment to the forecasts by the average forecasts, described above, is the solution to this issue. Using the theory outline in Section 3.1, if w maximises f(P) then **i** also maximises

$$g(P)=E[\mathbf{w}'(\mathbf{R}-\underline{\mathbf{m}})]- \tfrac{1}{2} A\, E[(\mathbf{w}'(\mathbf{R}-\underline{\mathbf{m}}))^2]$$

where $\underline{\mathbf{m}}=(\mathbf{m}'\mathbf{1})/(\mathbf{1}'\mathbf{1})$ (this is the matrix algebraic form of the standard equation for an average).

The objective function can be rewritten as

$$g(P)=\mathbf{w}'(\mathbf{m}-\underline{\mathbf{m}}) - \tfrac{1}{2} A\, \mathbf{w}'\mathbf{S}\,\mathbf{w}$$

where **S** is the second moment of **R-m**.

The key feature of **S** is that it is unchanged by a change in perspective. Therefore whatever properties are imposed on its structure they will remain in place from any perspective. It is therefore useful to satisfy the remaining criteria by structuring **S** rather than **V**.

3.3 Finalising the structure

It is important to make a decision on what constitutes intuitive. Given the difficulty in forecasting direction, let alone magnitude of return, correlation of return, or volatility of return, our preference is not to get too complex. The following rule of thumb reflects this preference: If the forecast is positive then the position should be positive. If the forecast increases then the position should increase. The opposite should also hold. To satisfy this rule we take **S** to be diagonal.

In making **S** diagonal we have also gone part way to ensuring stability and limiting the dependence to other parameters. How much further we go to achieve these criteria depends on two decisions. The first is to what extend do we allow the diagonal elements to change through time, and the second is do we want to reduce the number of parameters to one by assuming all the diagonal elements are equal.

Given the information in the system is contained in the forecasts we believe it not to be a bad decision to keep the elements of **S** constant. The only major objection to this is that if the currency markets operate under a new volatility environment the allocation rule will not reflect that, although it could be argued that it is the forecasts that should be accounting for such a change, not the allocation rule. Certainly having very volatile numbers in **S** will make the positions volatile, which is undesirable. A good middle ground is to use estimates that vary very gradually through time, but do eventually account for any big changes in the market environment. Estimates using five or ten years of historical data are typically smooth enough. Again, we caution against ignoring the conditional nature of **S** when using statistical techniques to estimate its elements.

Assuming that **S** has the same number in every element of its diagonal has both positive and negative implications. The negative implication is that it may be an unrealistic assumption. This may not be such a serious concern as it seems. By assuming a single number we are making the statement that we do not wish to bias the allocation toward any particular currency based on our estimates of the elements of **S**. Unless we have a high conviction level that we can accurately estimate these parameters, assuming the same number is the neutral stance.

There is a big positive implication from assuming the same number. We do not have to estimate *any* parameters, not even the one element of **S**. As previously noted, because we are only using one number in **S**, the allocation rule is impacted equally. That is to say it only effects the portfolio in terms of the aggressiveness level, which we denoted by A. Our estimation of **S** is then merged with the estimation of aggressiveness level. The estimation of either the level alone, or the merged estimation of level and **S**, is considered identical for many estimation procedures. Therefore our positions are now driven by the forecasts alone and we fully satisfy the parameter dependence criterion. In addition, the optimisation approach with the same number in all diagonal elements of **S** produces identical positions to the linear approach in the absence of constraints. Such a relationship provides intuitive appeal, stability, and ties together the linear approach with more mainstream academically accepted approaches.

3.4 The mathematics of the process

3.4.1 Unconstrained optimisation

To substantiate our claims we derive the solution to the optimisation problem in the absence of constraints, other than that the positions must sum to zero.

Using Lagrangian optimisation we maximise

$$g = \mathbf{w}'(\mathbf{m}-\underline{\mathbf{m}}) - \tfrac{1}{2} A\, \mathbf{w}'\mathbf{S}\mathbf{w} - q\mathbf{w}'\mathbf{1}$$

where q is the Lagrange multiplier for the position constraint.

Differentiating g with respect to w gives us that

$$dg/dw = \mathbf{m}\text{-}\underline{\mathbf{m}} - A\,\mathbf{S}\,\mathbf{w} - q\mathbf{1}$$

and setting this expression equal to zero and solving for w gives us that

$$A\,\mathbf{w} = \mathbf{S}^{-1}\,(\mathbf{m}\text{-}\underline{\mathbf{m}}) + q\ \mathbf{S}^{-1}\ \mathbf{1}.$$

Assuming that **S**=s**I**, where s is a scalar and **I** is the identity matrix, we have that

$$\mathbf{w}=K((\mathbf{m}\text{-}\underline{\mathbf{m}}) + q\ \mathbf{1})$$

where K is a constant equal to $(As)^{-1}$.

We can solve for q by multiplying the equation by **1'** and noting that **1'w**=0, we also denote $\underline{\mathbf{m}}$ more formally as **1**.(**m'1**)/(**1'1**), thus

$$0=(\mathbf{1'm}) - (\mathbf{1'1}).(\mathbf{m'1})/(\mathbf{1'1}) + q\,(\mathbf{1'1})$$

so q=0 and the optimal weights are given by

$$\mathbf{w} = K\,(\mathbf{m}\text{-}\underline{\mathbf{m}}).$$

3.4.2 Constraining a currency to be zero

Suppose now that the i'th element of the weight matrix is set to zero, indicating that the i'th currency is not relevant for a specific analysis. To account for this we add an additional term and Lagrange multiplier to the objective function, and hence maximise the function h given as

$$h = \mathbf{w}'(\mathbf{m}\text{-}\underline{\mathbf{m}}) - \tfrac{1}{2} A\ \mathbf{w}'\mathbf{S}\mathbf{w} - q\mathbf{w}'\mathbf{1} - p\mathbf{w}'\mathbf{e}$$

where p is the Lagrange multiplier for the constraint that sets the i'th currency exposure equal to zero. The vector **e** denotes a vector of zeros except the i'th element which is a one.

Differentiating g with respect to **w** gives us that

$$dg/dw = \mathbf{m}\text{-}\underline{\mathbf{m}} - A\,\mathbf{S}\,\mathbf{w} - q\mathbf{1} - p\mathbf{e}$$

and setting this expression equal to zero and solving for **w** gives us that

$$A\,\mathbf{w} = \mathbf{S}^{-1}\,(\mathbf{m}\text{-}\underline{\mathbf{m}}) + q\ \mathbf{S}^{-1}\,\mathbf{1} + p\,\mathbf{S}^{-1}\,\mathbf{e}.$$

Assuming that **S**=s**I**, where s is a scalar and **I** is the identity matrix, we have that

$$\mathbf{w}=K((\mathbf{m}\text{-}\underline{\mathbf{m}}) + q\ \mathbf{1} + p\ \mathbf{e})$$

where K is a constant equal to $(As)^{-1}$.

We can solve for q and p by multiplying the equation by **1**' and **e**' and noting that **1'w**=0 and **e'w**=0, we also denote $\underline{\mathbf{m}}$ more formally as **1**.(**m'1**)/(**1'1**), thus

$$0=(\mathbf{1'm}) - (\mathbf{1'1}).(\mathbf{m'1})/(\mathbf{1'1}) + q\,(\mathbf{1'1}) + p\,(\mathbf{1'e})$$

$$\text{and } 0=(\mathbf{e'm})\text{-}(\mathbf{e'1}).(\mathbf{m'1})/(\mathbf{1'1}) + q(\mathbf{e'1}) + p\,(\mathbf{e'e})$$

so

$$0 = q\,n + p$$

$$\text{and } 0 = m_i - \underline{m} + q + p$$

where mi is the i'th element of the vector m and n is the size of vector m, which gives us that $q=(m_i\text{-}m)/(n\text{-}1)$ and $p=(m\text{-}m_i)n/(n\text{-}1)$ and the optimal weights are given by

$$\mathbf{w}=K(\mathbf{m} - ((n-1)\underline{\mathbf{m}} + \underline{\mathbf{m}} - m_i)/(n\text{-}1)) + (\underline{\mathbf{m}}\text{-}m_i)n/(n\text{-}1)\ \mathbf{e}$$

which simplifies to

$$\mathbf{w}=K(\mathbf{m}\text{-}\underline{\mathbf{m}}^*)$$

for all elements of **w** except the i'th element which is zero, where $\underline{\mathbf{m}}^*$ is the average of the elements of **m** except for the i'th element ($\underline{\mathbf{m}}^*$=**1**.$(n\underline{m}\text{-}mi)/(n\text{-}1)$).

This result is identical to discarding the i'th element from the system and computing the optimal positions in an unconstrained context.

4 Unattractive features of the approach

Our recommended approach is designed to be appealing from a number of viewpoints. However, it too has some features that may not be liked by all investors.

The positions are a function of the average of the forecasts. It is true that if any element of the vector $(\mathbf{m}-\underline{\mathbf{m}})$ does not change then the optimal position will not change, but if one of the raw forecasts (i.e. an element of $\mathbf{m}$) changes, then all of the positions will change.

The positions will satisfy our intuition criterion relative to $(\mathbf{m}-\underline{\mathbf{m}})$. As $(\mathbf{m}-\underline{\mathbf{m}})$ increases so the positions increase, and vice versa. This property does not hold for the raw forecasts.

Our diagonal matrix $\mathbf{S}$ is invariant to perspective. However, the covariance matrix $\mathbf{V}$ will implicitly have a different structure. This structure will be a function of the forecasts and therefore will change through time.

An erroneous criticism of this approach is that if only a subset of the currencies are required for a particular application, then we must choose between using the full currency universe and setting undesired currency exposures to zero in the optimisation, or using a subset of the full universe. This is a redundant decision since, as we have shown in Section 3.4, the two processes produce identical mathematical solutions.

5 Conclusion

The allocation rule which best satisfies our criteria is to use the weights, $\mathbf{w}$, which maximise

$$\mathbf{w}'(\mathbf{m}-\underline{\mathbf{m}}) - \tfrac{1}{2} K \, \mathbf{w}'\mathbf{w}$$

subject to **w'1**=0 and any other desired constraints, where K denotes the aggressiveness level of the active management.

Of course alternatives are plentiful and we do not promote this approach as the single best solution. It may be that our criteria our not appropriate for a specific purpose. It may also be that the importance rankings of the criteria differ for different investment approaches and that may lead to a different solution. The point we would make is that careful consideration must go into the allocation rule. Not because it will create significantly superior performance (that is the role of the forecaster), but because a process can be undermined by an allocation rule that yields inconsistent or unusual positions. It is surprisingly easy to inadvertently incorporate such an allocation rule.

We have proposed a set of criteria that help with avoiding unwanted allocation rule properties and result in at least a sensible position recommendation. Our allocation rule satisfies these criteria and therefore provides some protection against inevitable unforeseen problems.

6 Reference

Lee Overlay Partners, (1999) Optimal and naïve investment strategies: A paradox, *Lee Overlay Partners Technical Report. Number lop-0799a.*