Optimal investment strategies with investor liabilities

Edwin J. Elton and Martin J. Gruber

New York University, New York, NY 10006, USA

Received June 1991, final version received January 1992

This paper examines the portfolio problem from the viewpoint of an investor who faces a set of liabilities. We show that optimum choices can be made from a set of spanning portfolios which include cash flow matching portfolios, index funds and actively managed portfolios. The optimal composition of each and the optimum mix among the portfolios are presented in terms of criteria which are computationally simple and have an appealing economic interpretation.

1. Introduction

The viewpoint taken in this paper is that of a financial intermediary which has incurred one or more liabilities, wishes to allocate funds among a set of assets, and is concerned with changes in net worth. This is a significant departure from traditional analysis. Traditional analysis has largely proceeded in one of two ways. Much of investment theory and analysis is concerned only with allocation among assets and assumes this allocation can be made without consideration of liabilities. Almost all of the standard references on modern portfolio theory treat the problem in this way. On the other hand, the literature that incorporates liabilities in the investment decision such as the duration literature uses, as an objective function, ending wealth and ignores period by period changes in wealth. We feel that the inclusion of liabilities and the focus on changes in net worth are important.

Correspondence to: Edwin J. Elton or Martin J. Gruber, Nomura Professors of Finance, Stern School of Business, New York University, 90 Trinity Place, New York, NY 10006-1594, USA.

The analysis has implications for an individual investor where the analogy to the liability stream is anticipated consumption expenditures.

An interesting exception to this is Liebowitz and Hendrickson (1988). They recommend a surplus framework for the portfolio problem. Part of our results regarding the riskless assets which are proved later in this paper are arrived at intuitively by Liebowitz and Hendrickson.

See for example, Elton and Gruber (1991). One could view multiperiod consumption investment models as an exception to this statement.

For example, almost all tests of immunization models use as a criteria the ability to match a pure discount instrument and examine the ending value as a way of comparing models. An exception is contained in our prior work [Elton et al. (1988)], where we examine period by period changes in net worth empirically.

0378-4266/92/505.00 © 1992—Elsevier Science Publishers B.V. All rights reserved
modifications of the prior literature. The bulk of this paper is concerned with deriving the new insights that are produced by viewing the problem in this way. For example, this framework allows us to examine in great detail the role of cash flow matching and duration matching as part of an investor's portfolio. As Nelson and Schaefer (1983) point out, while these strategies are widely used, no theory exists on why investors should utilize them.

The paper evolved from extensive observation of the behavior of a number of financial intermediaries. These intermediaries manage an enormous amount of pension assets for companies. Many intermediaries place some clients' assets in actively managed accounts, some in duration matched accounts and some in a cash flow matched account. An intermediary can simply use a quadratic program to solve for an optimum overall portfolio by treating the pension liabilities as negative assets. However, most pension clients require the economic intuition behind the choice and mixture of alternative types of active and passive accounts. This paper is the first to treat analytically the choice between duration matched, cash flow matched portfolios and active portfolios. It is a rigorous attempt to answer questions as to why and when an investor should use various passive techniques versus active techniques to manage funds. In addition, financial intermediaries need to know how to design commingled funds and if they are harming their clients if they do so. This paper answers these questions. The answers we obtain are in many cases quite different from current beliefs.

This paper is divided into three sections. In section 2 we explore why change in net worth is a suitable decision criterion for many investors, and present a general framework of analysis consistent with this decision criterion. In section 3 we examine the efficient frontier facing the investor who limits investments to those which have the same duration as liabilities. While this self-imposed constraint can lead to suboptimization, it is worth examining because it represents the way in which many institutions operate. Section 4, the longest section of this paper, deals with optimum bond portfolio strategies in general. In this section the appropriate role of cash flow matching and duration matching investments as part of an overall bond portfolio are examined and analytical results obtained. Cases where all assets are in equilibrium and cases where some assets are out of equilibrium are studied. In addition, general solutions and solutions where short sales are

\[5\]While that solution will be mathematically equivalent to the solution presented in this paper, solving the problem by QPS will not allow the derivation of the separation properties shown in this paper or allow the user to understand the reasons for the optimum allocation. Using the results of this paper the optimum allocation can be achieved by solving a problem which is much simpler in structure than a QPS problem.

\[6\]Nelson and Schaefer (1983, p. 65) pointed out that 'one of the deficiencies of the immunization literature is the absence of a convincing explanation of why investors should wish to immunize.' The answer of why or how much to immunize has remained unanswered until now.
restricted are examined. Explicit solutions for optimal portfolios are derived and the implications of these solutions for investor behavior are examined. While in this paper we restrict our analysis to bonds, the framework developed and many of our conclusions would hold even with the inclusion of other types of assets.

2. Framework of analysis

Let us start by examining why one might want to view the allocation problem in terms of changes in net worth. Consider three types of financial institutions, where asset-liability planning is common: insurance companies, pension funds, and banks.

Why should insurance companies care about period to period changes in net worth rather than net worth at one particular date in the future? First, most insurance companies hope to exist into the indefinite future and do not plan a finite termination date. While at any point in time the stream of existing liabilities may terminate, as time passes new ones are added and the institution hopes this process will continue indefinitely. Second, for most insurance companies the period by period value of their net worth affects their ability to survive and do business. An insurance company with negative net worth is insolvent even though its investment policy might result in positive net worth subsequently. Third, regulators directly tie an insurance company's ability to engage in new business to its net worth. Thus, changes in net worth are a relevant variable to focus on for an important class of intermediaries.

As a second example of an intermediary where value at the horizon is often suggested as an objective function, consider a pension fund. Once again, at any point in time the liabilities have a finite termination date, but as time passes, new liabilities are added so that for most pension funds the time horizon stretches into the indefinite future. Even pension funds for retired employees where a finite termination date is reasonable should care about changes in net worth. Current accounting standards (FAS 87) call for both assets and liabilities to be recorded at market value and define procedures so that a short fall in market value of assets relative to liabilities results in adverse events for the pension fund. More particularly, failure to have assets with market values at least equal to liabilities lead to differences flowing through to the income statement of the corporation sponsoring the pension fund and results in additional payments to the fund. Thus a focus on net worth and changes in net worth is appropriate for pension funds.

Both the literature on banking and practice of banks support awareness of change in net worth as an appropriate decision criteria for decision making.

7The amount of insurance a company can write is a multiple of net worth.
For example, Grammatikos et al. (1986) specifically use the expected value of and variance of net worth as a decision framework when analyzing issues concerning US banks' foreign currency position. In addition, the Federal Home Loan Banks Board (FHLBB, 1983) has argued for the use of duration gaps as a measure to set insurance premiums for S & Ls. The analysis in this paper is very much in the spirit of both of these themes in the banking literature in that the relationship between return and risk defined in terms of net worth and of duration as a risk measure will be explicatively studied and analytical results obtained.

The argument we have made for insurance companies, pension funds and banks should apply to most financial intermediaries and many individuals. In the rest of the paper we will focus on changes in net worth. With this framework we can now develop the analysis. Let

\[ A = \text{an initial sum of money which can be invested,} \]
\[ L = \text{the present value of the liabilities facing a firm,} \]
\[ R_1 = \text{the one period spot rate,} \]
\[ R_A = \text{the rate of return on assets, a random variable,}^8 \]
\[ R_L = \text{the rate of return on liabilities, a random variable,}^8 \]
\[ W_0 = \text{initial net worth,} \]
\[ W_1 = \text{net worth at the end of period 1,} \]
\[ R_N = \text{rate of return on net worth.} \]

The initial net worth is the asset position minus the present value of the liabilities.\(^9\) Note that the liabilities may be payable many period in the future. In symbols \( W_0 = A - L. \) The rate of return the manager earns on net worth is:\(^{10}\)

\[ R_N = \frac{W_1}{W_0} - 1 = \frac{A(1 + R_A) - L(1 + R_L)}{(A - L)} - 1. \]

Let us examine a riskless strategy from the investor's standpoint.

2.a. Riskless portfolio

Assume the investor has designed a portfolio of default free assets with cash flows which exactly match the cash outflows associated with the

\(^8\)This return includes the impact due to shifts in the term structure as well as other economic influences.

\(^9\)In the following discussion we assume that \( A > L. \) If assets are not equal to or larger than the liabilities not only does the investor have negative starting wealth, no riskless strategy is feasible.

\(^{10}\)At this point we have not restricted the form of the return distribution on assets or liabilities. These returns can originate from shifts in the term structure or any other source. Later we will specifically introduce the covariance between these returns and show how they relate to an economic index.
E.J. Elton and M.J. Gruber, Optimal investment strategies

liabilities.\textsuperscript{11} We will assume that liabilities are priced at their market value as if they were tradable assets. Pricing at market value is consistent with regulatory intent and much of regulatory practice. It is also consistent with the way investment advisors analyze pension plans and is consistent with FAS 87 regarding the valuation of pension liabilities.\textsuperscript{12} With this assumption, the law of one price implies that the return on these assets must match the return on the liabilities or \( R_A = R_L \). Furthermore, the cost of these assets must equal the present value of the liabilities and an exact match portfolio will cost \( L \) dollars. If \( A \) is larger than \( L \) and the remainder is put in the one period pure discount bond the return on the portfolio is:

\[
R_N = (A - L)(1 + R_1) + L(1 + R_A) - L(1 + R_L) - 1 = R_1.
\]

Thus the manager is guaranteed the one period spot rate and the risk of the portfolio is zero.

In traditional analysis a one period government bill is considered the riskless investment. When liabilities are considered, the riskless strategy is no longer a single instrument but rather a portfolio of 'risky assets' plus the one period government bill.\textsuperscript{13} The return on the riskless investment has not changed but the construction of the riskless portfolio has. In subsequent sections we will use \( R_F \) to designate the return on the riskless portfolio where \( R_F = R_1 \).

2.6. Combinations of the riskless portfolio and risky assets

It is convenient for purposes of future analysis to show that in this setting all combinations of the riskless portfolio and a risky portfolio lie along a straight line passing through the riskless rate in expected return–standard deviation space. If this statement seems familiar, it is. It is one of the basic tenets of portfolio analysis. What is different in this context is that the definition of the riskless portfolio is not 100\% of assets in the one period.

\textsuperscript{11}Given the large number of debt instruments available, the design of such a portfolio is always possible except in the case where liabilities occur beyond the life of the assets. The portfolio can be designed as a solution to a simple linear programming problem. See Elton and Gruber (1991).

\textsuperscript{12}Actually, under cash flow matching the discrepancy between assets and liabilities cannot occur, not only because of market forces, but also because of the manner in which actuaries function. Once cash flow matching is selected as an investment technique the actuary will use the rate on the bond portfolio as the actuarial rate in all liability calculations.

\textsuperscript{13}We refer to excess funds being placed in the one period government bill. The definition of the maturity on this bill is arbitrary, for example if regulation were to be enforced (capital adequacy examined) on a continuous basis we would use this instantaneous rate of interest as our one period riskless rate. We have purposely kept the definition of one period ambiguous for the analysis in the text follows under any definition of the appropriate measurement interval.
pure discount government bill, and the return on the risky portfolio is different from and more complex than the return on the risky portfolio in standard analysis.

To show the linearity let $R_Y$ be the return on the combination of the risky portfolio and the riskless portfolio. The return on a portfolio consisting of only investment in risky assets is:

$$R_N = \frac{A(1+R_A) - L(1+R_L)}{(A-L)} - 1 = R_A + \frac{L(R_A - R_L)}{(A-L)}. \tag{1}$$

Note the return on the portfolio depends not only on the return earned on the risky assets, it depends on the difference in return between the assets and the liabilities and the amount by which assets exceed liabilities.

If this portfolio is combined with the riskless portfolio then

$$R_Y = XR_F + (1-X)\left[ R_A + \frac{L(R_A - R_L)}{(A-L)} \right],$$

where $X$ equals the fraction of the funds invested in the riskless portfolio, which is a combination of liabilities, the one period spot rate, and the cash flow matched portfolio. Employing this equation to solve for the expected value and variance of the return on portfolio $Y$ (arguing that $R_F$ has zero variance and zero covariance with all other investments) and using the equations for variance to eliminate $X$ from the equation we find:

$$\bar{R}_Y = R_F - \left[ \frac{\bar{R}_N - R_F}{\sigma_N} \right] \sigma_Y, \tag{2}$$

where $\sigma$ indicates standard deviation.

Eq. (2) is a straight line and has the same functional form as the capital market line in standard portfolio theory. What is different is the $\bar{R}_N$ and $\sigma_N$ depend on both the return on assets and the return on liabilities as well as the ratio of assets to liabilities held by any investor. [See eq. (1).] In particular $\bar{R}_N$ has a standard deviation which depends on the standard deviation of the return on the assets, the return on the liabilities and the relative size of the two. Even in a world of homogeneous expectations, eq. (2) implies that alternative investors will face different efficient frontiers. If

14The riskless portfolio contains, in addition to the discount bond, an exact match portfolio of bonds and the liabilities. The exact match portfolio and the liabilities are obviously correlated with the liabilities in the risky portfolio and probably correlated with the assets in the risky portfolio. However, they net out of the riskless asset and the remainder is the pure discount instrument which was zero correlation with the risky portfolio.
combinations of the riskless portfolio and risky portfolio lie on a straight line in expected return–standard deviation space, then the efficient frontier becomes the straight line passing through the riskless portfolio which is tangent to the efficient frontier of risky assets (see fig. 1). It is important to note that different investors can face different efficient frontiers not only because of heterogeneous beliefs but also (unlike traditional analysis) because of different liabilities.

It is possible that point $N$ is optimal. However, most investors will want to combine some exact matching and one period instruments with the optimal risky portfolio. Thus, almost all investors should place some money in a cash flow matched portfolio. We will now turn to a more detailed examination of the efficient frontier of risky assets depicted in fig. 1. We shall first examine its composition when investments are restricted to duration matched strategies. Subsequently we will examine the more general case.

3. The efficient frontier of duration matched strategies

In this section we will examine the efficient frontier when the investor restricts the choices to immunized strategies. Although this restriction may lead to suboptimal decisions, many financial intermediaries operate under this self imposed constraint. For example, many pension funds hire insurance
companies to construct a set of assets to immunize a fixed set of liabilities. In the next section we will relax this assumption and explore more general solutions. For the time being we will assume that management has decided to duration match (immunize) and to restrict themselves to holding some combinations of:

1. a riskless portfolio, as defined in the previous section, with return $R_F$,
2. a one period pure discount instruments with return $R_1$,
3. a portfolio of risky assets with return $R_A$.

By considering both the one period pure discount instrument and the riskless portfolio we allow any combination of the cash flow matched portfolio and one period pure discount instrument to be held. Thus adding a cash flow matched portfolio would be redundant. Returns are assumed to follow a one factor linear return-generating process of the form:

$$ R_i = R_1 + D_i l + e_i. \quad (3) $$

To be completely general we are allowing the mean of $l$ to be nonzero. This allows for the possibility of a term premium on bond returns. Elsewhere [Elton and Gruber (1989)] we have shown that while this type of return generating process is a very reasonable representation of returns due to shifts in the term structure as well as other sources of return, in general, a multi-index model is necessary. In this paper we will, for simplicity, use a single factor model. The generalization to a multifactor model is trivial analytically and should be obvious to the reader. Furthermore, expected returns are assumed to be given by either

$$ \bar{R}_i = R_1 + D_i \bar{l}, \quad (4a) $$

or

$$ \bar{R}_i = \alpha_i + R_1 + D_i \bar{l}. \quad (4b) $$

The interpretation of these equations is that the investor chooses to measure risk as a function of sensitivity to an index [e.g., duration in eq. (3)] and that the return on assets may be [eq. (4a)] or may not be [eq. (4b)] in equilibrium as implied by arbitrage pricing theory. For example, temporary mispricing of a particular issue or issues in the bond market could cause a bond to have a positive or negative $\alpha_i$ in eq. (4b). Thus, the assumption underlying eqs. (3) and (4) is a one factor linear return-generating process where the investor is interested in immunizing against this source of risk and where a subset of bonds may be believed to offer a return above or below the equilibrium return.

\footnote{Once again we emphasize that the problem could be solved as a quadratic programming problem with the liabilities included as negative assets. However, this solution is sterile insofar as yielding economic insights for the conditions under which various assets are included.}
We will assume that any mispricing occurs in assets and that liabilities will be fairly priced. We will continue to use the subscript A to stand for all assets and L for the market value (present value) of liabilities. We will also assume that at time zero \((A - L) > 0\) or that the decision maker is solvent.

Before developing the efficient frontier we need to discuss what we mean by duration matching and immunized portfolios. The return on net worth of a combination of risky assets and liabilities is:

\[
R_N = \frac{AR_A - LR_L}{(A - L)}.
\]

Substituting eq. (3) for \(R_A\) and \(R_L\) and rearranging we have:

\[
R_N = \frac{(AR_1 - LR_1) + (AD_A - LD_L)I + Ae_A - Le_L}{(A - L)}.
\]

A duration matched portfolio is a portfolio of assets and liabilities which has zero sensitivity to the factor \(I\). This condition requires that

\[
AD_A - LD_L = 0,
\]

or

\[
D_A = \frac{L}{A} D_L.
\]

Portfolios of assets that meet this condition will be designated as duration matched or immunized portfolios. Note that immunization does not involve holding assets with an average duration equal to that of liabilities, rather the duration of liabilities must be adjusted by the ratio of liabilities to assets.

Note that combinations of duration matched portfolios and the riskless portfolio are still duration matched since each portfolio by itself is immunized. In order to clarify the subsequent analysis we shall use the term duration matched portfolio to stand for all risky duration matched portfolios.

We now examine the combination of the riskless portfolio, the duration matched portfolio of risky assets, and the one period pure discount instrument. The following propositions hold:

**Proposition 1.** If \(x_A = 0\) then the efficient frontier is a single point which represents the riskless portfolio. While investors will place some investment in the cash flow matched portfolio they will not place any funds in any other duration matched portfolio.

We will assume that a portfolio of risky assets can be constructed with any
duration and that eqs. (3) and (4a) hold for each such portfolio. Employing eq. (3) along with eq. (4a) we find the return of an investment \( X \) in a portfolio of risky assets and \( (1-X) \) in the pure discount bond such that the duration on the combination equals the duration of the liabilities is:

\[
R_N = \frac{A(1-X)R_1 + XA(R_1 + e_A) - L(R_1 + e_L)}{(A-L)}
\]

Expected return becomes

\[
R_N = \frac{A(1-X)R_1 + XAR_1 - LR_1}{(A-L)} = R_1 = R_F,
\]

but this is identical to the expected return on the riskless portfolio. However, the risk on this portfolio is not zero. In fact, the variance of return is:

\[
\text{Var}(R_N) = \frac{X^2 A^2 \text{Var}(e_A) + L^2 \text{Var}(e_L)}{(A-L)^2},
\]

where \( \text{Var}(e_A) \) and \( \text{Var}(e_L) \) are the unsystematic variance of the assets and liabilities, respectively.

Thus the investment opportunity set is a horizontal line emanating from \( R_F \). From the efficient set theorem the only efficient portfolio is the riskless portfolio. Thus the optimum investment policy is to cash flow match the liabilities with the remainder invested in the one period government bill.

**Proposition 2.** If \( \alpha_A > 0 \) the efficient frontier consists of combinations of the riskless portfolio and a portfolio of risky assets. All portfolios that include the pure discount instrument in constructing the efficient frontier of risky assets are inefficient.

An investment of \( X \) in the risky assets and \( (1-X) \) in the pure discount bond such that the combination is duration matched to the liabilities has a return of

\[
R_N = \frac{A(1-X)R_1 + XA(R_1 + \alpha_A + e_A) - L(R_1 + e_L)}{(A-L)}.
\]

Note that an extra return is being earned on the assets and not on the liabilities. Thus, we are assuming that eq. (4b) holds for at least some assets.
and that portfolios of all durations offer the same above equilibrium returns. The expected return on the combination is:

$$R_N = R_F + \frac{X \alpha_A}{(A - L)}$$

while the variance of return is:

$$\sigma_N^2 = \frac{X^2 A^2 \text{Var}(e_A) + L^2 \text{Var}(e_L)}{(A - L)^2}.$$  

We can combine this portfolio with the riskless portfolio (as shown in the prior section). These combinations lie along a straight line in expected return standard deviation space with a slope given by

$$\theta = \frac{R_N - R_F}{\sigma_N} = \frac{\alpha_A X A}{\left[X^2 A^2 \text{Var}(e_A) + L^2 \text{Var}(e_L)\right]^{1/2}}.$$  

The preferred portfolio is the one that maximizes this slope. Taking $d\theta/dX$ to determine the proportions to put in the portfolio offering above equilibrium returns and recalling $R_F = R_1$ we get:

$$\frac{d\theta}{dX} = \frac{A \alpha_A L^2 \text{Var}(e_L)}{\sigma_N^3} > 0.$$  

Thus, the slope is maximized when $X$ is made as large as possible. This implies that the investor should form the risky portfolio by short selling the one period spot rate (borrow at this rate) and use the proceeds to buy the duration matched portfolio. If short selling of the one period bond is not

---

16 If we assumed there existed only one portfolio of assets with above equilibrium returns then the solution to this problem would be trivial. Enough of the one period bond would be held so that the duration of the portfolio plus the riskless bond equaled the duration of the liabilities.

17 Note that since the $A - L$ cancel in the objective function arrived at before rearranging, the problem would be unchanged if we used as an objective, dollar return on net worth rather than percent return.
feasible (the more realistic situation) the investor should form the risky portfolio by placing 100% of the money in the risky duration matched portfolio. Thus, the efficient frontier consists of a straight line which represents combinations of the riskless portfolio and the duration matched portfolio. The proportion of bonds that the investor places in each depends on his or her risk return tradeoff. It is perfectly rational for an investor to choose to place a fraction of the funds in a cash flow matched portfolio (as part of the riskless portfolio) and a fraction in the more risky duration matched portfolio.

Many institutions offer both cash flow matching and duration matching portfolios to their customers. In this section we have seen that if all assets are priced in equilibrium, no investor should hold any immunized (duration matched portfolios), except for the cash flow matched portfolio. On the other hand, if some assets are priced out of equilibrium it will almost always be advantageous to cash flow match some portion of liabilities while at the same time investing part of the funds in a portfolio which is duration matched but not cash flow matched.

We will now examine the efficient frontier when we consider non-immunized strategies.

4. Optimal bond portfolio strategies in general

In this section we will expand the analysis of the efficient frontier to include additional assets and non-duration matched strategies. We will continue to assume the investor can place funds in the riskless portfolio.

4.a. The portfolios considered

We assume an investor can select an optimum portfolio by combining the riskless portfolio with three other portfolios:
1. A particular portfolio of efficiently priced assets.
2. A unique portfolio of assets with excess returns.
3. The cash flow matched portfolio.

In a separate paper [Elton and Gruber (1989)] we present a set of mutual fund theorems for alternative multi index models which can be used to show that restricting the choice to these portfolios is optimal. The economic intuition behind this results is easy to see. Our assumption concerning the return generating model means that risk is divided into diversifiable and nondiversifiable components. Since diversifiable risk is not priced, the choice among portfolios of efficiently priced assets should involve minimizing residual risk. Thus the minimum risk portfolio is in the decision set. In Appendix A we prove that the portfolio that minimizes residual risk is unique. With perfect factor replication it is the replicating portfolio. The
riskless portfolio is included because it is the portfolio which allows us to eliminate the diversifiable risk associated with the investors' liabilities. The uniqueness of a portfolio of positive \( \alpha \) assets (designated as the special portfolio) is rigorously shown in a latter section when we derive its composition. The inclusion of this portfolio in the decision set is motivated by the fact that it provides excess returns, returns which may more than compensate for the fact that it has residual risks.

The final portfolio the cash flow matched portfolio allows different amounts of the one period pure discount bond to be held since the only difference between it and the riskless portfolio is that the riskless portfolio contains some investment in the one period pure discount bond.

In the next section of the paper we will solve for the optimum composition of the risky portfolio. We shall study two cases. In the first we shall assume that the factor portfolio is perfectly correlated with the factor driving bond returns. In the second case shall assume that correlation is less than perfect and therefore the factor portfolio has residual risk.

4.6. The ease of a perfectly correlated factor portfolio

We will continue to assume that the investor makes an allocation decision between a risky and a riskless portfolio. To determine the composition of the risky portfolio, the investor tries to maximize the slope of the line passing through the riskless portfolio and a risky portfolio in expected return standard deviation space. Using the subscript \( N \) to denote the risky portfolio:

\[
\text{Max } \theta = \frac{\bar{R}_N - R_F}{\sigma_N}.
\]

Define

- \( X_c \) = fraction of assets invested in the cash flow matching portfolio,
- \( X_p \) = fraction of assets invested in the factor portfolio,
- \( X_s = 1 - X_c - X_p \) = fraction of assets in the special portfolio,
- \( D_c, D_p, D_s \) = the duration of the cash flow, factor, and special portfolio, respectively,
- \( L \) = market value of liabilities,
- \( A \) = market value of assets (\( A \) is assumed greater than \( L \) at time zero).

Utilizing eqs. (4a) and (4b) and the definitions above, the expected return on a risky portfolio composed of assets invested in portfolio C, P, and S and containing liabilities \( L \) is
\[ \bar{R}_N = \frac{AX_* \sigma_* + D_N \bar{l}}{A - L} + R_F, \]

where \( D_N \) is the duration of net worth and is given by

\[ D_N = [A \{X_cD_L + X_pD_p + (1 - X_c - X_p)D_s\} - D_L L]. \]

The variance of the risky portfolio is:

\[ \sigma^2_N = \left[ \frac{D^2_N \sigma^2_I + (X_cA - L)^2 \sigma^2_L + X^2_c A^2 \sigma^2_p}{(A - L)^2} \right], \]

where

\( \sigma^2_I \) = variance of the return on the index \( I \),
\( \sigma^2_L \) = variance of the residuals (after the effect of the factor has been removed) of the cash match portfolio and the liabilities,
\( \sigma^2_S \) = residual variance of the special portfolio,
\( \sigma^2_p \) = residual variance of the factor portfolio.

We make the assumption in this section that the factor portfolio is perfectly correlated with the factor \( I \), hence \( \sigma^2_p = 0 \) and \( \sigma^2_s \) does not appear in the variance equation. Furthermore, the residual risks of all portfolios, from the one factor model, are assumed to be uncorrelated with one another. Finally, note that the cash flow matched portfolio and the liabilities should have the same duration.

Substituting the expression for expected return and variance into \( \theta \), noting \( R_F = R_I \), and substituting \( 1 - X_c - X_p \) for \( X_* \) yields

\[ \theta = \frac{X_c(1 - X_c - X_p) + D_N \bar{l}}{[D^2_N \sigma^2_I + (X_cA - L)^2 \sigma^2_L + (1 - X_c - X_p)^2 A^2 \sigma^2_p]^{1/2}}. \]

To find optimum proportions take the derivative of \( \theta \) with respect to \( X_c \) and then with respect to \( X_p \), setting each equal to zero. The derivative with respect to \( X_c \) is:

\[ \frac{d\theta}{dX_c} = \frac{I_A(D_L - D_s) - \sigma_* A}{(\sigma^2_N)^{1/2} (A - L)} \left( \frac{\bar{R}_N - R_F}{\sigma^2_N} \right) \left( \frac{D_N A(D_L - D_s) \sigma^2_I}{(A - L)^3} \right) + A(X_cA - L) \sigma^2_L - A^2 (1 - X_c - X_p) \sigma^2_p = 0. \]

Substituting for \( R_N \) and \( \sigma^2_N \) and simplifying yields

\[ \bar{I}(X_cA - L) \sigma^2_L [X_p(D_s - D_p)A - D_s(A - L)] \]
\[ + \bar{I}(1 - X_c - X_p)A \sigma_x^2 [D_L(A - L) \]
\[- X_p(D_L - D_p)A] - \alpha_\ell \sigma_N^2 [D_L(A - L) - X_p A(D_L - D_p)] \]
\[- \alpha_x (X_c A - L) \sigma_x^2 (A - L + X_p A) = 0. \]  

(5)

Similarly taking the derivative of \( \theta \) with respect to \( X_p \) and simplifying yields
\[ \bar{I}(D_p - D_c)(X_c A - L)^2 \sigma_x^2 + \bar{I} A(1 - X_c - X_p) \sigma_x^2 [4D_p(1 - X_c) \]
\[ + (X_c A - L)D_L] - \alpha_\ell \sigma_N^2 [4D_p(1 - X_c) + (X_c A - L)D_L] \]
\[- \alpha_x (X_c A - L)^2 \sigma_x^2 = 0. \]  

(6)

There are two solutions to these two questions. One involves imaginary values for the investment in assets and hence can be dismissed. The other solution is

\[ X_c = \frac{L}{A}. \]
\[ X_p = \frac{\bar{I} \sigma_x^2 - \alpha_\ell \sigma_N^2}{\alpha_x \sigma_N^2 (D_p - D_c)} \frac{A - L}{A}, \]
\[ X_s = (1 - X_c - X_p) = \frac{\alpha_\ell D_p \sigma_N^2}{\alpha_x \sigma_N^2 (D_p - D_c) + \bar{I} \sigma_x^2} \frac{A - L}{A}. \]

Several properties of the optimal solution are interesting. The most intriguing is that given assets larger in size than liabilities, the value of \( X_c \) shows that it always pays to hold a cash flow matched portfolio exactly equal in size to the liabilities. It is worth noting that since the riskless portfolio also contains a cash flow matched portfolio equal to liabilities, no matter what mixture of the risky portfolio and the riskless portfolio the investor selects, he or she will always exactly cash flow match liabilities. The intuition behind this proof is that since cash flow matching is the only way to get rid of all of the residual risk on liabilities, it pays to do so.

The remainder of the assets are invested in part in the special portfolio

\[ 18 \text{The reader can validate this as a solution by noting that for } X_c = L/A \text{ the first and fourth terms of eqs. (5) and (6) are equal to zero and for either equation the other two terms reduce to the equation which is shown in the text. Examining the second derivative shows that the denominator in the expression for } X_s \text{ and } X_p \text{ must be positive for a maximum. If the denominator is negative the second derivative is negative, the first positive and the solution involves placing the maximum amount in the special asset.} \]
and in part in the factor portfolio. Let us examine the allocation of the extra assets \( A - L \) in greater detail. The fraction of these extra assets put in the special portfolio is:

\[
\frac{X_s}{X_s + X_p} = \frac{\alpha_s D_p \sigma^2_s}{\alpha_s \sigma^2_s (D_p - D_s) + T\sigma^2_s}.
\]  

(7)

A more intuitively appealing form can be derived by expressing this relationship in terms of the average duration \( (D_N) \) of the portfolio bought with surplus funds \( (A - L) \). In this case \( D_N \) becomes,

\[
D_N = \frac{X_s}{X_s + X_p} - D_s + \frac{X_p}{X_s + X_p} - D_p.
\]

Substituting in the expression for \( X_s \) and \( X_p \) and simplifying and rearranging:

\[
\frac{D_N}{\alpha_s \sigma^2_s} = \frac{1}{\alpha_s \sigma^2_s (D_p - D_s) + T\sigma^2_s}.
\]

Substituting this into eq. (7) yields

\[
\frac{X_s}{X_s + X_p} = \frac{\alpha_s D_N \sigma^2_s}{\alpha_s \sigma^2_s (D_p - D_s) + T\sigma^2_s}.
\]

(8)

The numerator is a reward risk ratio for the special portfolio while the denominator is a reward risk ratio for an efficiently priced portfolio of duration \( D_N \). Thus the greater the reward risk ratio for the special portfolio relative to the ratio for an efficiently priced portfolio the more is invested in the special portfolio. Up to now we have assumed that the special portfolio could be constructed without consideration of the allocation between other portfolios or the characteristics of the factor portfolio. We will now prove this.

4.6.1. Composition of special portfolio

In the proof we assume the previous result that an amount of assets equal to liabilities is invested in a cash flow matching portfolio. Therefore the allocation of the remaining assets can be treated as a separate portfolio problem. The slope of the tangency line connecting the riskless portfolio and the efficient frontier is
E.J. Elton and M.J. Gruber, Optimal investment strategies

\[ \theta = \frac{\sum x_i X_i + D_N \bar{I}}{\left[D_N^2 \sigma_I^2 + \sum X_i^2 \sigma_i^2 \right]^{1/2}}, \]

where

- \( X_k \) = the proportion invested in special asset \( k \),
- \( D_k \) = the duration of asset \( k \),
- \( D_N = [D_p(1 - \sum X_i) + \sum X_i D_i] \),
- and other terms are as before.

Taking the derivative with respect to \( X_k \) and setting it equal to zero results in

\[ \alpha_k \left[D_N^2 \sigma_I^2 + \sum X_i^2 \sigma_i^2 \right] - D_N \sigma_I^2 (D_k - D_p) \sum x_i X_i \]

\[ - \sum x_i X_i (X_k \sigma_k^2) + \bar{I} (D_k - D_p) \sum X_i^2 \sigma_i^2 - D_N \bar{I} X_k \sigma_k^2 = 0. \]

Multiplying through by \( X_k \) and adding across all \( k \) results in

\[ [D_N \sigma_I^2 \sum \alpha_i X_i - \bar{I} \sum X_i^2 \sigma_i^2 ] = 0. \]

Substituting this into the first order condition and simplifying yields

\[ X_k = \frac{\alpha_k}{\sigma_k^2} \left[ \frac{D_N^2 \sigma_I^2 + \sum X_i^2 \sigma_i^2}{\sum \alpha_i X_i + D_N \bar{I}} \right] = \frac{\alpha_k}{\sigma_k^2} \lambda. \]

where \( \lambda = \text{total risk/total excess return} \).

Note that this is identical to the Treynor and Black (1973) result derived in another context. Since the total amount invested in the special asset is the sum of \( X_i \), the fraction \( X_k \) represents of the alpha portfolio is given by

\[ \frac{X_k}{\sum X_k} = \frac{(\alpha_k/\sigma_k^2) \lambda}{\sum (\alpha_i/\sigma_i^2) \lambda} = \frac{\alpha_k/\sigma_k^2}{\sum \alpha_i/\sigma_i^2}. \]

Thus, the composition of the special portfolio can be determined independently from the composition of the factor portfolio and depends only on characteristics of the special assets. In particular, note that characteristics of the liability structure such as its duration do not affect the composition of the special portfolio. Throughout this section, we have assumed the existence of such a portfolio. We have now shown it can be calculated and that it is unique, it is independent of the factor portfolio, and it is independent of the liability structure.

Before leaving this section we should examine one more solution. We will first delineate those conditions that cause us to short sell the factor portfolio.
After defining these conditions we will derive the appropriate solution to the portfolio problem when these short sales are not allowed.

4.b.2. The optimal solution when the factor portfolio cannot be sold short

Up to now we have assumed that the factor portfolio could be sold short. The factor portfolio is sold short if \( X_s/(X_s + X_p) \) is greater than one. From eq. (8) this condition is

\[
\frac{x_s}{\sigma_s^2} > \frac{\bar{t}D_N}{D_N^2 \sigma_f^2}.
\]

The left-hand side is the reward to risk ratio of the special portfolio. The right hand side of the inequality is the reward to risk ratio of an efficiently priced portfolio of duration \( D_N \). If this inequality holds, 100% of the free assets would be placed in the special portfolio or greater than 100% if short sales are allowed.

If the previous analysis calls for a positive amount of funds to be placed in the factor portfolio the solution is unchanged by a short sales constraint. If the solution calls for the factor portfolio to be sold short and short sales are not allowed, the solution can be obtained by setting the fraction of assets invested in the factor portfolio equal to zero and solving the restructured problem:

\[
\text{Max } \theta = \frac{x_s A(1 - X_c) + D_N \bar{t}}{D_N^2 \sigma_f^2 + (X_c A - L)^2 \sigma_f^2 + (1 - X_c)^2 A^2 \sigma_t^2}^{1/2},
\]

where

\[
D_N = A[X_c D_c + (1 - X_c) D_{\bar{t}}].
\]

The first order condition is:

\[
X_c = \frac{L}{A} + \frac{A - L}{A} \frac{D_c(\bar{t} \sigma_f^2 - x_s D_c \sigma_f^2)}{D_c(\bar{t} \sigma_f^2 - x_s D_c \sigma_f^2) + \sigma_t^2(\bar{t} D_c + x_s + x_s D_{\bar{t}})}.
\]  

When the factor portfolio could be sold short, the optimum amount to invest in cash flow matching was \( L/A \) or place a sum of money exactly equal to the liabilities in the exact match portfolio. The sign of the second term determines whether more or less is invested in the cash flow match portfolio when short sales are not allowed.

Taking the second derivative of \( \theta \) with respect to \( X_c \) gives
\[
\frac{d^2\theta}{dX_e^2} = -D_e(\sigma_s^2 - \alpha_s D_s \sigma_s^2) - \sigma^2 (\sigma^2 + \alpha_s D^2).
\]

For the solution for \(X_e\) to be a maximum this must be less than zero. The second derivative has the same functional form as the denominator of the fraction in eq. (9). For the second derivative to be negative, the denominator of the coefficient of \(A - L\) in eq. (9) must be positive. From eq. (7) we know that one case where we desire to sell the factor portfolio short is where \(\sigma_s^2 - \alpha_s D_s \sigma_s^2\) is negative. This is the numerator of the coefficient of \(A - L\) in eq. (9). Thus the second term of eq. (9) is negative and we cannot get a solution to exactly cash flow match a set of funds equal to liabilities. Rather we invest more than \(A - L\) in the special portfolio and invest less than \(L\) in the cash flow matched portfolio. This is an extremely important result. Most institutions cannot or will not short sell bonds as part of their optimal portfolio. Given this constraint, it is generally optimum to still engage in some cash flow matching but to only invest a fraction of the liabilities in this manner.

4.c. Optimal investment policy when a factor portfolio has residual risk

In this section we will assume that there does not exist a portfolio perfectly correlated with the factor. The assets we will consider are a cash flow match portfolio, a portfolio of efficiently priced assets with minimum residual risk at any duration level and a portfolio of inefficiently priced assets. In the last section we showed that when short sales were allowed, the cash flow matched portfolio should be held in an amount equal to the liabilities. The same conclusion holds when a zero residual risk factor portfolio does not exist. While we can formally show this, the conclusion is easy to see given the analysis in the last section. When the minimum risk portfolio has positive residual risk rather than zero risk, this portfolio is less attractive. Since the characteristics of other portfolios are held constant, making the minimum risk portfolio less attractive should not reduce the amount in the exact match portfolio. Thus, the proportion in the exact match portfolio should be at least \(L\). For investment beyond \(L\), the minimum risk factor portfolio dominates the exact match portfolio. Any asset in the exact match portfolio can be included in the minimum residual risk at any duration level. At any particular duration the return and systematic risk for the minimum risk portfolio and the exact match portfolio are the same. Thus, the minimum risk portfolio dominates the cash flow match portfolio for proportions in the cash flow match portfolio above \(L\).

In the remainder of this section we will examine how to split an amount \(A - L\) between the special portfolio and minimum risk factor portfolio. Once
again we will maximize the slope of a line connecting the riskless portfolio (exact match plus pure discount) and the efficient frontier. We will retain the symbols used in the last section with the following exception. We will deal with the portfolio problem from the viewpoint of allocating the funds $A - L$. We will express all $X_i$ as a fraction of this amount of funds. Therefore, $X_s + X_p = 1$. The expected return is given by

$$\bar{R}_N = x_s X_s + D_N \bar{I},$$

where

$$D_N = [D_s X_s + (1 - X_s) D_p].$$

The variance requires more explanation. The systematic risk is $D_s^2 \sigma_s^2$. We believe the following is a reasonable representation of risk:19

$$\sigma_p^2 + X_s^2 (\sigma_s^2 - \sigma_p^2), \quad \sigma_s^2 > \sigma_p^2.$$ 

This representation states that the minimum residual risk is obtained when 100% is invested in the minimum risk factor portfolio but its risk is $\sigma_p^2$ rather than zero as assumed in the previous analysis. The $(\sigma_s^2 - \sigma_p^2)$ is the incremental residual risk when some of the money is placed in the special portfolio. Given this representation, the following problem should be solved:

$$\text{Max } \theta' = \frac{x_s X_s + D_N \bar{I}}{[D_s^2 \sigma_s^2 + \sigma_p^2 + X_s^2 (\sigma_s^2 - \sigma_p^2)]^{1/2}}.$$ 

The first order condition rearranged is

$$X_s = \frac{x_s (\sigma_p^2 + D_s^2 \sigma_s^2) + \bar{I} \sigma_p^2 (D_s - D_p)}{\bar{I} (\sigma_s^2 - \sigma_p^2) D_p - x_s (D_s - D_p) D_N \sigma_I}.$$ 

Examining the second derivative shows that for this to be a maximum:

$$\frac{x_s}{(\sigma_s^2 - \sigma_p^2)} < \frac{\bar{I} (D_s - D_p)}{(D_s - D_p)^2 \sigma_I^2}.$$ 

The left hand side is the special return per unit of incremental residual risk from holding the special assets. The right hand side is the change in systematic risk. If the above inequality does not hold, the first order condition is a maximum and $X_s$ should be made as large as possible. Thus,

19There are a number of correlation patterns of residuals that would result in this representation of residual risk.
for the minimum risk portfolio to be held the inequality must hold. This requires that $D_r$ be greater than $D_p$. In addition, the special return to residual risk has to be sufficiently small relative to the incremental systematic return to risk. Therefore, the amount invested in the minimum risk portfolio and factor portfolio depends on systematic return and risk as well as the residual risk and extra return.

Changing the objective function to analyze the proportion to place in each special asset involves replacing $X_iD_i$ with $\sum X_iD_i$, $X_i$ with $\sum X_i$, and $X_i(\sigma_i^2 - \sigma_p^2)$ with $\sum X_i(\sigma_i^2 - \sigma_p^2)$. Taking the derivative of this objective function with respect to each $X_k$, aggregating across all $X_k$ to simplify, results in the following expression for $X_k$:

$$X_k = \frac{\alpha_k}{\sigma_k^2} \left( \frac{\text{Total Risk}}{\text{Total Return}} \right) + \left( \frac{D_k - D_p}{\sigma_k^2} \frac{\sigma_p^2}{D_p} \right),$$

where

- Total Risk $= D_k^2 + \sigma_p^2 + \sum X_i(\sigma_i^2 - \sigma_p^2)$,
- Total Return $= D_k^2 + X_i\alpha_i$.

Thus, the optimum proportion to invest in asset $k$ is a function of the characteristics of the minimum risk factor portfolio. The characteristics of this portfolio need to be known in order to determine the composition of the special portfolio. When the factor portfolio had zero residual risk, separation was possible. When the factor portfolio has positive residual risk, separation is impossible. Nevertheless, the special portfolio is unique and can be prespecified as assumed in the earlier analysis.

**Conclusion**

It is very common today to see insurance companies offering their pension fund customers a choice from among at least three sets of portfolios; an actively managed bond portfolio, a portfolio which matches the duration of a customer specified set of cash flows (liabilities), and a passive index fund. The question we have answered in this paper is when should a customer select one of these investments by itself or a portfolio of these investments and what is the design of the optimal portfolio.

We have shown that if a portfolio is to be fully immunized and all assets are priced in equilibrium, it never pays to put funds in an immunized portfolio which is not cash flow matched. If assets exist which offer an above equilibrium return, then in addition to placing a fraction of funds in a cash flow matched portfolio, there is an economic rationale for placing a fraction of funds in a non-cash flow matched, but duration matched portfolio.
In constructing the efficient frontier of risky assets for an investor who did not require full immunization, we considered combinations of a cash flow matched portfolio, a factor portfolio, and a special portfolio of above equilibrium return assets. We showed that there was a single combination of assets yielding equilibrium returns that was preferred and a single preferred combination of assets yielding above equilibrium returns so that our framework was general. We derived the optimal proportions in each portfolio. We showed that, whether short sales are allowed or not, it is always optimal to exact cash flow match the liabilities. Additional funds are allocated between the special assets and factor portfolio in a specified manner. When the solution calls for the factor portfolio to be sold short and short sales are forbidden the optimum investment policy consists of combinations of the exact match portfolio and the special portfolio.

While all of our analysis is general, the reader who believes that all assets are priced in equilibrium needs only set alpha to zero for the special portfolio. He will find that all proofs go through but that the fraction of assets invested in the special portfolio is zero.

References

Elton, E.J. and M.J. Gruber, 1989, Portfolio analysis with a non-normal multi-index return generating prices, New York University's Stern School of Business, working paper.