JIA 112 (1985) 229-277

PORTFOLIO SELECTION IN THE PRESENCE OF FIXED LIABILITIES: A COMMENT ON 'THE MATCHING OF ASSETS TO LIABILITIES'

By A. D. Wilkie, M.A., F.F.A., F.I.A.

1. This note was inspired by the paper 'The Matching of Assets to Liabilities' presented by A. J. Wise to the Institute in March 1984 (Wise, 1984b). In it he presented a method of looking at the problem of matching which I claimed in the discussion was essentially a portfolio selection approach. However, his approach had a number of novel features. I wish to discuss one of these, approaching it from the conventional portfolio selection viewpoint. I am not aware that this problem has been considered elsewhere in the substantial literature that exists on portfolio selection. Full discussion of the mathematics of the conventional portfolio selection problem is contained in Sharpe (1970) and Szegö (1980), and a general explanation is available in many modern financial text books, and in the Institute paper by Moore (1972).

2. I shall use one of Wise's examples as an illustration, and I shall also use certain aspects of his notation. However, it is convenient to introduce also some notation more usual in discussing portfolio selection.

3. The problem I discuss is that of selecting assets to match fixed liabilities. I say 'fixed' in the sense that the liabilities are not marketable and cannot be disposed of. Their monetary value is, however, a random variable. The usual way in which liabilities are dealt with in the portfolio selection problem is to treat them as marketable securities, which have a known price, and include them simply as negative assets. Wise's problem, and mine, assumes instead that the liabilities are like those of a pension fund or insurance company, which are not marketable and do not have a readily determined market price.

4. The usual portfolio selection problem is that of investing a fixed amount of money now, in order to achieve a desired out-turn at the end of a fixed time horizon. Since the solution does not depend on the size of the initial amount available, the problem reduces to one of choosing the proportions to be invested in each of a variety of available assets. Wise agrees with the approach of looking at a fixed time horizon, in his case the final date of all the liabilities due, and I follow him in this. He is concerned with finding the quantities of available assets that best match, in some sense, the liability. I wish to look instead at the most desirable set of assets having regard also to their present prices. This is the essence of where I differ from Wise, and we both differ from the usual portfolio selection problem.

Wise's elementary example

5. It is worth repeating Wise's elementary example. He assumes that all cash flows occur at the ends of years 1, 2 and 3, and that there are no demographic

factors and no inflation. There are two securities available as assets, and there are known liabilities. Cash flows can be represented by a row vector, whose three elements represent flows at the ends of years 1, 2 and 3. For numerical convenience I shall multiply all Wise's numbers by 100, so the fixed liabilities are represented by:

$$(100, 100, 100),$$

i.e., an amount of 100 is due at the end of each of the three years. The available securities are represented by:

$$\text{security 1: } (10, 100, 0)$$
$$\text{security 2: } (10, 10, 100).$$

We can think of these if we like as fixed interest stocks providing interest of 10 at the end of each year until redemption, an amount of 90 on redemption, and redeemable at the end of years 2 and 3 respectively.

6. So far everything looks fixed with certainty. Wise's seminal contribution is to say: "Let us assume that fixed money amounts can be reinvested in some other security, whose rates of return are random variables." If the proceeds of a fixed money stock are reinvested and rolled up in this way, the amount of the final proceeds from such an investment at the end of the fixed time horizon is a random variable. Wise assumes in his example that this one stochastic investment is comparable with putting cash on deposit, at a rate of interest which, independently each year, can take on the values 8% and 10% with equal probability. One can treat the liability payments as being met by borrowing cash at this same stochastic rate of interest. The accumulated debt is rolled up to the end of the time horizon, so that the final amount of liability is also a random variable.

7. Wise now approaches the problem by defining his primitive elements as a cash flow of unity in each successive year, viz:

$$(1, 0, 0),$$
$$(0, 1, 0),$$
$$(0, 0, 1).$$

While this has algebraic advantages for his method of presentation, I think it conceals the fact that what you are buying are securities, rather than these primitive units.

8. I prefer to work with the securities and the liabilities directly. I call the securities S_1 and S_2 and the liabilities L. The proceeds at the end of year 3 from security S_i, assuming intermediate cash flows reinvested until the end of year 3, are called R_i, a random variable, with expected value E_i, and variance $V_i = \sigma_i^2$. These amounts are all per unit of security purchased. The amount of the rolled-forward liability is also a random variable, R_L, with expected value E_L and variance $V_L = \sigma_L^2$. These amounts are in pounds and squared pounds, rather than per unit. Note that my Es are different from Wise's. The three final returns may be correlated, and their covariances and correlation coefficients are defined:

$$C_{12} = \rho_{12}\,\sigma_1\,\sigma_2$$
$$C_{1L} = \rho_{1L}\,\sigma_1\,\sigma_L$$
$$C_{2L} = \rho_{2L}\,\sigma_2\,\sigma_L.$$

Note also that my C's are not the same as Wise's.

9. The distribution of R_1, R_2 and R_L can easily be derived, by enumerating all possible cases. For example, for the liabilities we have:

Interest rates			*Proceeds by end of:*		
Year 2	*Year 3*	*Probability*	*Year 1*	*Year 2*	*Year 3*
8	8	·25	100	208	324·64
8	10	·25	100	208	328·80
10	8	·25	100	210	326·80
10	10	·25	100	210	331·00

This gives a mean liability, E_L, of 327·81, and a variance V_L of 5·5563, or a standard deviation, σ_L, of 2·3572.

10. Similar calculations allow us to complete the values:

Means	$E_1 = 120{\cdot}881$	$E_2 = 122{\cdot}781$	$E_L = 327{\cdot}810$
Variances	$V_1 = 1{\cdot}241763$	$V_2 = {\cdot}055563$	$V_L = 5{\cdot}556300$
Standard Deviations	$\sigma_1 = 1{\cdot}1143$	$\sigma_2 = {\cdot}2357$	$\sigma_L = 2{\cdot}3572$
Covariances	$C_{12} = {\cdot}243663$	$C_{1L} = 2{\cdot}436630$	$C_{2L} = {\cdot}555630$
Correlation coefficients	$\rho_{12} = {\cdot}9276$	$\rho_{1L} = {\cdot}9276$	$\rho_{2L} = 1{\cdot}0$

Generalization

11. We can generalize a little here. Assume that there are n securities, subscripted by $i = 1, 2, \ldots, n$. Call the liabilities 'security' $n+1 = L$. Denote the returns per unit of security S_i and for the liability as the random row vector:

$$\mathbf{r}' = (R_1, R_2, \ldots, R_n, R_L).$$

Denote the expected returns by the row vector:

$$\mathbf{e}' = (E_1, E_2, \ldots, E_n, E_L).$$

Denote the variance/covariance matrix by:

$$\mathbf{V} = \begin{pmatrix} V_1 & C_{12} & \ldots & C_{1n} & C_{1L} \\ C_{21} & V_2 & \ldots & V_{2n} & V_{2L} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ C_{n1} & C_{n2} & \ldots & V_n & C_{nL} \\ C_{L1} & C_{L2} & \ldots & C_{Ln} & V_L \end{pmatrix}$$

Note that $C_{ji} = C_{ij}$, so the matrix is symmetric. It is also, as all covariance matrices are, positive semi-definite.

Covariances and correlations

12. In our example the covariance matrix is given by:

$$\mathbf{V} = \begin{pmatrix} 1{\cdot}241763 & {\cdot}243663 & 2{\cdot}436630 \\ {\cdot}243663 & {\cdot}055563 & {\cdot}555630 \\ 2{\cdot}436630 & {\cdot}555630 & 5{\cdot}556300 \end{pmatrix}$$

with corresponding correlation matrix:

$$\begin{pmatrix} 1{\cdot}0 & {\cdot}9276 & {\cdot}9276 \\ {\cdot}9276 & 1{\cdot}0 & 1{\cdot}0 \\ {\cdot}9276 & 1{\cdot}0 & 1{\cdot}0 \end{pmatrix}$$

13. It will be noted that the correlation between security 2 and the liability is unity. Clearly 10 units of security 2 provide cash flows (100, 100, 1000), so the final proceeds, $10R_2$, are bound to be exactly 900 higher than R_L. This feature makes Wise's example a special case. One cannot assume in general that any two securities have perfect correlation, either positive or negative, nor that any is perfectly correlated with the liabilities.

Ultimate surplus

14. If, like Wise, we purchase x_1 units of S_1 and x_2 of security S_2, then the final proceeds from the assets will be $x_1R_1+x_2R_2$. Deducting the final amount of liability gives us the ultimate surplus:

$$S=x_1R_1+x_2R_2-R_L.$$

The expected value of S,E, is given by:

$$\begin{aligned} E&=x_1E_1+x_2E_2-E_L, \\ &=120{\cdot}881x_1+122{\cdot}781x_2-327{\cdot}810 \end{aligned} \tag{14.1}$$

in our particular example.

The variance of S,V, is given by:

$$\begin{aligned} V&=x_1^2V_1+2x_1x_2C_{12}+x_2^2V_2-2x_1C_{1L}-2x_2C_{2L}+V_L, \\ &=1{\cdot}241763x_1^2+{\cdot}487326x_1x_2+{\cdot}055563x_2^2 \\ &\quad -4{\cdot}873260x_1-1{\cdot}111260x_2+5{\cdot}556300 \end{aligned} \tag{14.2}$$

in our particular example.

Generalization again

15. If we generalize again to n securities we can define the investments as x_i of security S_i, and include the liabilities with a factor of -1 in the row vector:

$$\mathbf{x}'=(x_1,x_2,\ldots,x_n,x_{n+1}),$$

where $x_{n+1}=-1$.

The expected ultimate surplus is then given by:

$$E=\mathbf{x}'\mathbf{e}=\mathbf{e}'\mathbf{x},$$

and the variance by:

$$V=\mathbf{x}'\mathbf{V}\mathbf{x}.$$

Wise's solution

16. So far Wise and I have done the same, expressed differently. I have used in effect the method of his 'direct solution', described in § 3.13 of his paper. He now defines various optimum portfolios of assets. One can choose either an unconstrained optimum, where no restraints are put on the values of the x_i's, or a constrained solution, where all the x_i's are required to be non-negative. Wise's (unqualified) match minimizes the mean square ultimate surplus, that is the second moment of the ultimate surplus S, given by:

$$E^2+V.$$

Wise calls this E_2, but to avoid confusion I shall call it G. Wise also defines an unbiased match, which minimizes G, subject to E being zero, which gives the minimum variance portfolio, subject to E being zero.

17. This is where I part company with Wise. As I shall explain, I see no reason to restrict ourselves to these solutions. In the usual portfolio selection sense, neither of these solutions is an 'efficient' solution, except in particular circumstances.

Market prices

18. At this point I have to bring in the market prices of the assets, since I shall argue that the rational investor must take account of the prices of securities in order to choose an optimal portfolio. Let the market price per unit of security S_i be P_i. Define the row vector of prices as:

$$\mathbf{p}'=(P_1,P_2,\ldots,P_n,0),$$

where the final 0 shows that we do not have to buy the liability; we are landed with it already.

19. Wise does not quote prices in his example. I shall choose $P_1=400$, $P_2=100$. These are hardly realistic market prices, but I need to use rather extreme prices in order to show the results visually. You will see what happens with realistic prices later.

20. The total cost of buying x_1 of security S_1 and x_2 of security S_2 is clearly given by:

$$P=x_1P_1+x_2P_2, \tag{20.1}$$
$$=400x_1+100x_2,$$

in our particular example, or in the general case by:

$$P=\mathbf{x}'\mathbf{p}=\mathbf{p}'\mathbf{x}.$$

Feasible portfolios

21. I now want to explore the whole range of feasible portfolios, by allowing x_1 and x_2 to take on any values, positive or negative, with no constraints. Given x_1 and x_2 we can calculate the values of E, V and P, from formulae (14.1), (14.2) and

(20.1) above. Thus we can consider the expected ultimate surplus, the variance of ultimate surplus and the price we pay all as functions of x_1 and x_2.

22. Since there are only two securities in our example, we can 'solve' for x_1 and x_2 in terms of P and E, provided that $E_1P_2 - E_2P_1 \neq 0$ to give:

$$x_1 = \frac{P_2E - E_2P + P_2E_L}{E_1P_2 - E_2P_1}, \tag{22.1}$$

$$x_2 = \frac{-P_1E + E_1P - P_1E_L}{E_1P_2 - E_2P_1}, \tag{22.2}$$

or

$$x_1 = -\cdot01080372\ E + \cdot00331623\ P - \cdot88539149,$$
$$x_2 = \cdot00270093\ E - \cdot00326491\ P + 3\cdot54156595,$$

in our particular example.

23. We can then insert these values of x_1 and x_2 into formula (14.2) to give:

$$\begin{aligned} V = \frac{1}{(E_1P_2 - E_2P_1)^2}\Big\{ & (E_2^2V_1 - 2E_1E_2C_{12} + E_1^2V_2)P^2 - 2(P_2E_2V_1 - P_2E_1C_{12} \\ & - P_1E_2C_{12} - P_1E_1V_2)PE + (P_2^2V_1 - 2P_1P_2C_{12} + P_1^2V_2)E^2 \\ & - 2(P_2E_2V_1 - P_2E_1C_{12} - P_1E_2C_{12} + P_1E_1V_2)E_LP \\ & - 2(E_1C_{2L} - E_2C_{1L})(E_1P_2 - E_2P_1)P \\ & + 2(P_2^2V_1 - 2P_1P_2C_{12} + P_1^2V_2)E_LE \\ & + 2(P_1C_{2L} - P_2C_{1L})(E_1P_2 - E_2P_1)E + (P_2^2V_1 - 2P_1P_2C_{12} \\ & + P_1^2V_2)E_L^2 \\ & + 2(P_1C_{2L} - P_2C_{1L})(E_1P_2 - E_2P_1)E_L + (E_1P_2 - E_2P_1)^2V_L \Big\} \end{aligned} \tag{23.1}$$

or $V = \cdot000008972034\ P^2 - \cdot000004407353\ PE + \cdot000001323805\ E^2$
$\quad - \cdot013977449392\ P + \cdot002024504646\ E + 6\cdot077697605108$

in our particular example.

Thus V, which was given by a quadratic form in x_1 and x_2, is also given by a quadratic form in E and P.

24. It is convenient to denote this quadratic form in the standard notation for conic sections:

$$V = aP^2 + 2hPE + bE^2 + 2gP + 2fE + c,$$

where

$$\begin{aligned} a &= (E_2^2V_1 - 2E_1E_2C_{12} + E_1^2V_2)/(E_1P_2 - E_2P_1)^2 \\ h &= -(P_2E_2V_1 - P_2E_1C_{12} - P_1E_2C_{12} + P_1E_1V_2)/(E_1P_2 - E_2P_1)^2 \\ b &= (P_2^2V_1 - 2P_1P_2C_{12} + P_1^2V_2)/(E_1P_2 - E_2P_1)^2 \\ g &= -(P_2E_2V_1 - P_2E_1C_{12} - P_1E_2C_{12} + P_1E_1V_2)E_L/(E_1P_2 - E_2P_1)^2 \\ & \quad -(E_1C_{2L} - E_2C_{1L})/(E_1P_2 - E_2P_1) \end{aligned}$$

$$f = (P_2^2V_1 - 2P_1P_2C_{12} + P_1^2V_2)E_L/(E_1P_2 - E_2P_1)^2 + (P_1C_{2L} - P_2C_{1L})/(E_1P_2 - E_2P_1)$$
$$c = (P_2^2V_1 - 2P_1P_2C_{12} + P_1^2V_2)E_L^2/(E_1P_2 - E_2P_1)^2 + 2(P_1C_{2L} - P_2C_{1L})E_L/(E_1P_2 - E_2P_1) + V_L.$$

P–E–V *space*

25. We can now consider feasible portfolios in the P–E–V space. For each point in the P–E plane there is only one possible combination of x_1 and x_2, given by formulae (22.1) and (22.2), which gives us only one value of V, given by formula (23.1). The locus of feasible portfolios is therefore a surface in the P–E–V space, in fact a quadric surface, known as an elliptic paraboloid, which is a sort of oval bowl. If we fix the value of P, we see that the cross-section of the surface is a parabola in the E–V plane. Similarly, if we fix the value of E the cross-section of the surface is a parabola in the P–V plane. If we fix the value of V, the cross-section in the P–E plane is either an ellipse, or has no real points at all if we have chosen a value of V smaller than the minimum possible value. Since V is a variance it is clearly impossible for it to be negative. But in certain special cases, of which our example is one, it is possible for V to be zero. A picture of this surface for our particular example is shown in Figure 1, and cross-sections for planes chosen arbitrarily of $E = 50$, $P = 600$ and $V = 1{\cdot}0$ are given in Figures 2, 3 and 4. The bowl is a very elongated one, more suitable for cooking fish than plum pudding.

26. It is sometimes convenient to use the standard deviation of ultimate surplus, σ, instead of the variance, V. The value of $\sigma^2 = V$ is given by formula (23.1). If we plot the locus of feasible portfolios in the P–E–σ space, we get a different quadric surface, an elliptic hyperboloid, which actually consists of two parts, mirror images of each other in the plane $\sigma = 0$, though we are only interested in the one with positive values of σ. For such a surface, cross-sections in the P or E planes are hyperbolas, and cross sections in the σ plane are again ellipses. If, as in our example, the minimum value of V is zero, then the surface becomes a double elliptic cone, with the points of the two halves of the cone touching at the unique point where $V = \sigma^2 = 0$. There are also other special cases.

Efficient portfolios

27. We have now found what the range of all possible portfolios looks like. Which of them are sensible ones to consider? I shall now define, similarly to the usual portfolio selection definition, an 'efficient' portfolio as a feasible portfolio which is not 'dominated' by any other. A portfolio is dominated by another if:

for the same P and E, the other has a lower V;
for the same P and V, the other has a higher E;
for the same E and V, the other has a lower P;
for the same P, the other has a higher E and lower V;
for the same E, the other has a lower P and lower V;
for the same V, the other has a lower P and a higher E;
the other has a lower P and higher E and lower V.

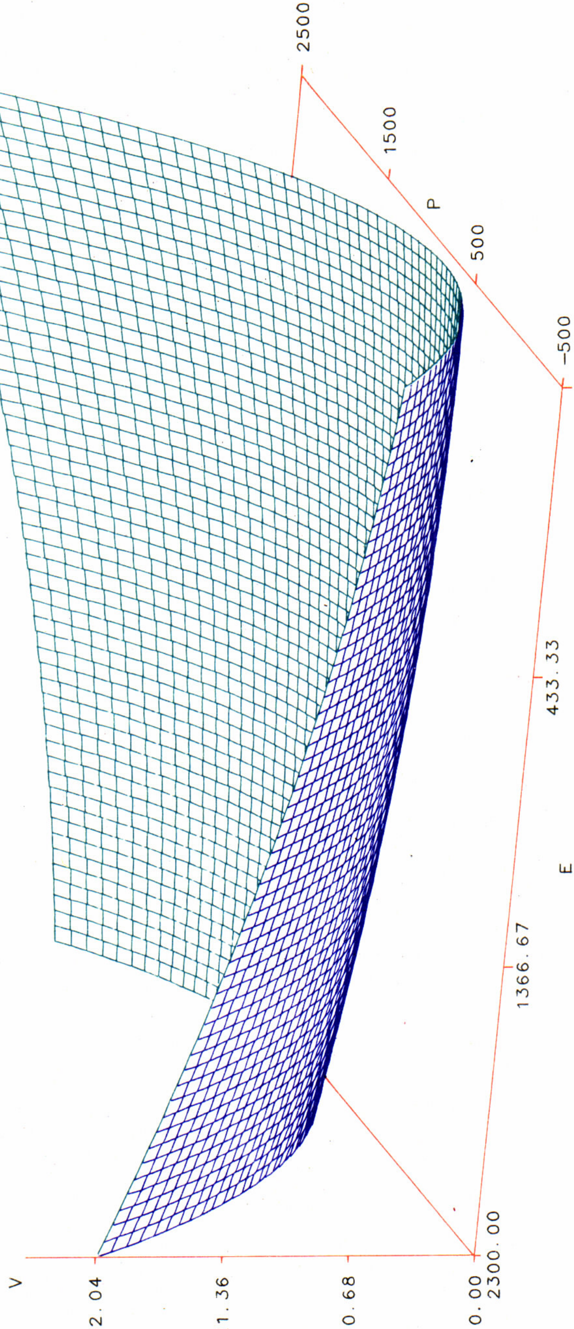

Figure 1
Variance (V) as a function of Price (P) and Expected (E)
Data as in Wise with P1=400 P2=100

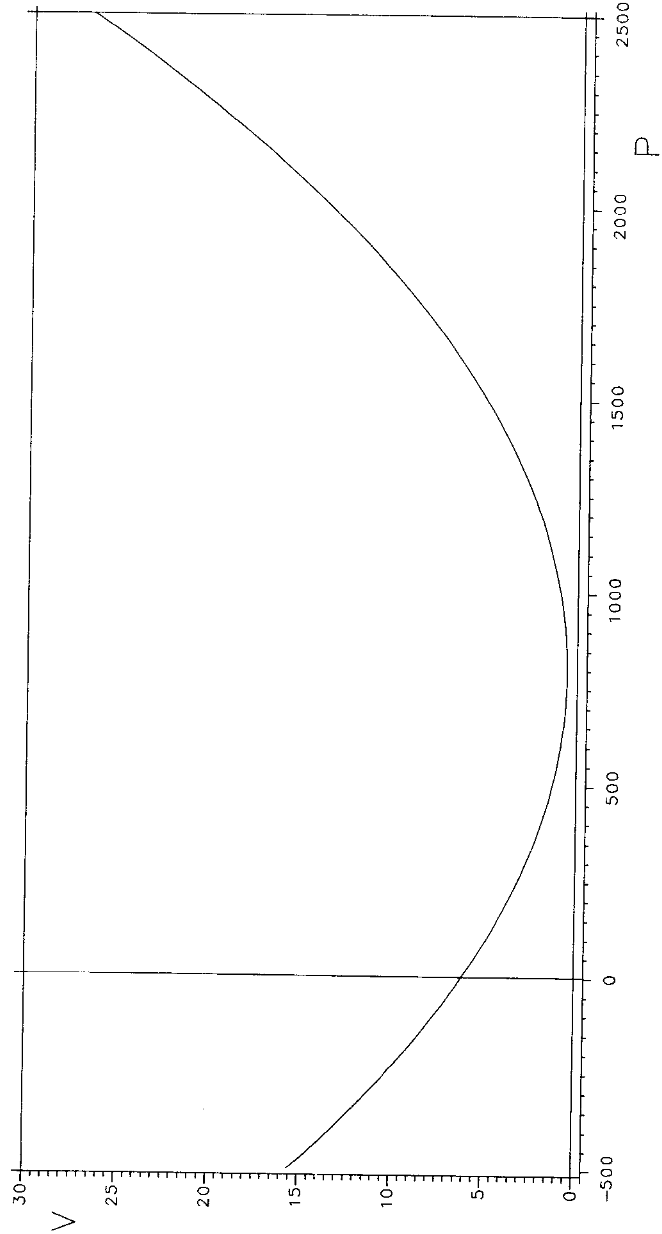

Figure 2
Variance (V) as a function of Price (P) for fixed Expected E=50
Data as in Wise with P1=400 P2=100

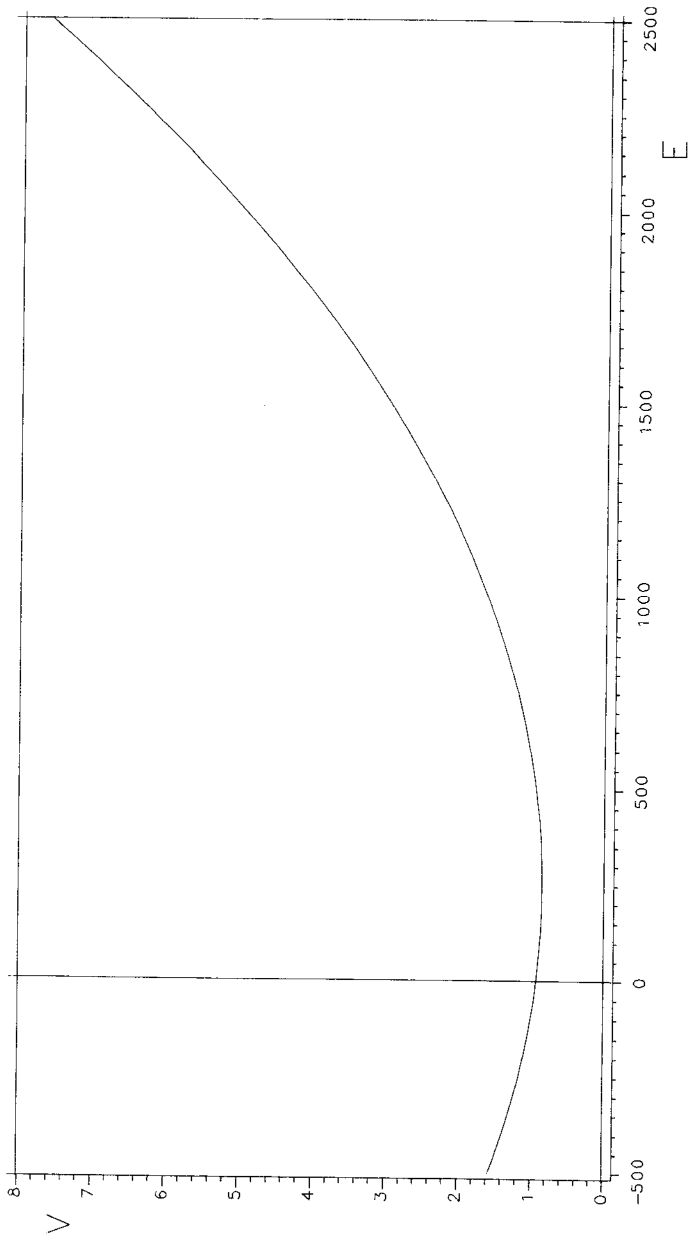

Figure 3
Variance (V) as a function of Expected (E) for fixed Price P=600
Data as in Wise with P1=400 P2=100

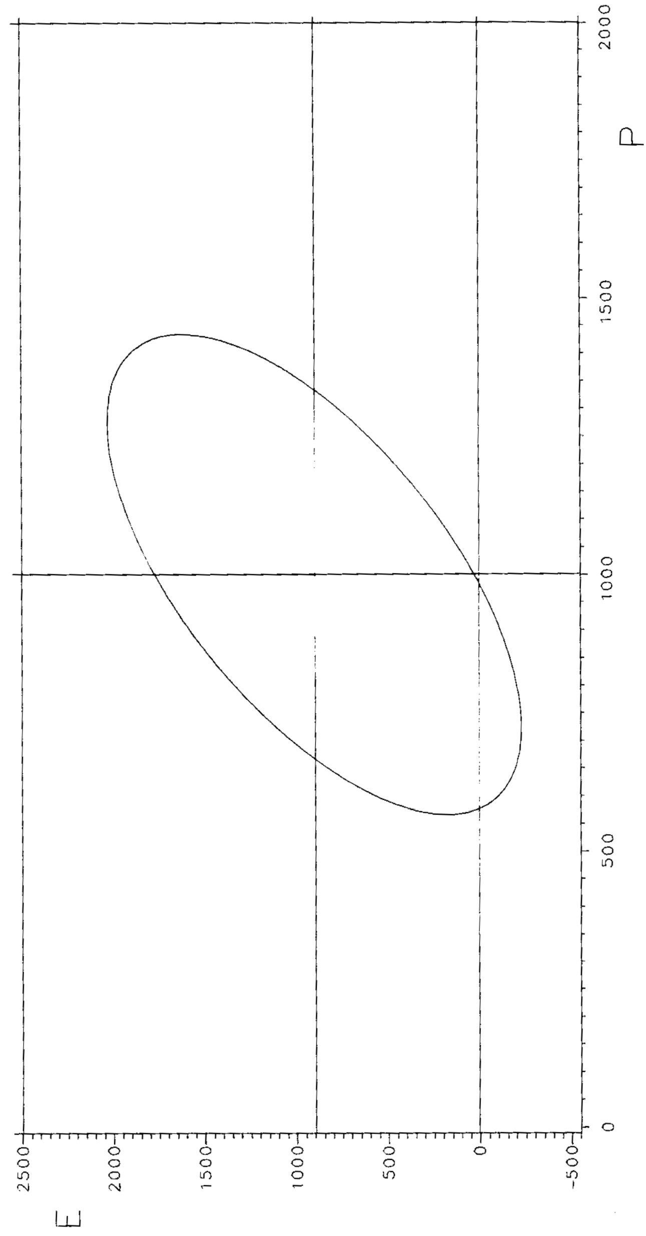

Figure 4

Expected (E) as a function of Price (P) for fixed Variance V=1.0

Data as in Wise with P1=400 P2=100

Thus I assume that investors are in favour of a high expected surplus, E, a low variance of surplus, V, and a low immediate price, P. Investors whose objectives do not coincide with these are not considered here. Consideration of Figure 2, in which expected surplus E is held constant, shows that any portfolio on the right hand arm of the parabola is dominated in this sense by a portfolio at the same level as it on the left hand arm; the expected surplus is the same, the variance is the same and the price is lower for the portfolio on the left hand arm. Similarly in Figure 3, portfolios on the left hand arm are dominated by portfolios on the right hand arm, which have a higher E for the same P and V. In Figure 4, all portfolios have the same V, but portfolios in the right hand part of the ellipse are dominated by those in the left hand part, which give the same E with a lower P; and all portfolios in the lower part of the ellipse are dominated by those in the upper part, which have a higher E for the same P. Thus the only portfolios not dominated by another are those in the 'north-western' sector, which because the ellipse is obliquely placed is a large 'quarter' of it.

The minimum variance portfolio

28. Figure 5 shows the projection of the elliptic paraboloid onto the P–E plane, with lines of constant V, the 'contour lines', shown as concentric ellipses. Each ellipse represents a lower variance than those outside it, and the lowest possible variance is given by the centre of the ellipses, which represents the point where the 'nose' of the elliptic paraboloid approaches most nearly (in this case touches) the plane $V=0$. This point is given by (P_C, E_C, V_C) where

$$P_C = \frac{hf-bg}{ab-h^2} = \frac{P_1(V_2C_{1L}-C_{12}C_{2L}) + P_2(V_1C_{2L}-C_{12}C_{1L})}{(V_1V_2-C_{12}^2)}$$

$$E_C = \frac{hg-af}{ab-h^2} = \frac{E_1(V_2C_{1L}-C_{12}C_{2L}) + E_2(V_1C_{2L}-C_{12}C_{1L})}{(V_1V_2-C_{12}^2)} - E_L$$

$$V_C = \frac{abc+2fgh-af^2-bg^2-ch^2}{ab-h^2} = V_L - \frac{(V_1C_{2L}^2-2C_{12}C_{1L}C_{2L}+V_2C_{1L}^2)}{(V_1V_2-C_{12}^2)}$$

$$= V_L\left(\frac{1-(\rho_{12}^2+\rho_{1L}^2+\rho_{2L}^2)+2\,\rho_{12}\rho_{1L}\rho_{2L}}{1-\rho_{12}^2}\right).$$

At this point x_1, x_2 are equal to x_1^c, x_2^c, where

$$x_1^c = \frac{V_2C_{1L}-C_{12}C_{2L}}{V_1V_2-C_{12}^2} = \frac{\sigma_L(\rho_{1L}-\rho_{12}\rho_{2L})}{\sigma_1(1-\rho_{12}^2)},$$

$$x_2^c = \frac{V_1C_{2L}-C_{12}C_{1L}}{V_1V_2-C_{12}^2} = \frac{\sigma_L(\rho_{2L}-\rho_{12}\rho_{1L})}{\sigma_2(1-\rho_{12}^2)},$$

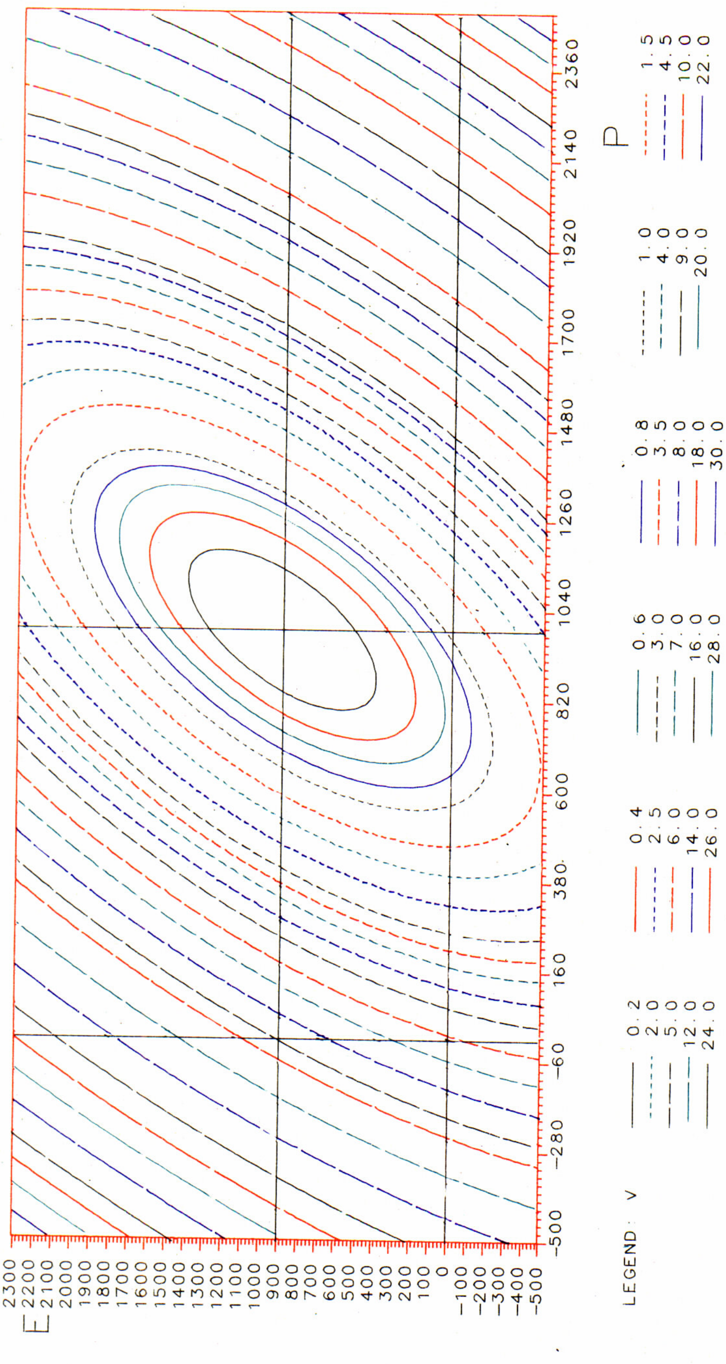

Figure 5
Contour lines of Variance as a function of Price (P) and Expected (E)
Data as in Wise with P1=400 P2=100

In our example, the values are:

$$\begin{aligned} P_C &= 1000 \\ E_C &= 900 \\ V_C &= 0 \\ x_1^c &= 0 \\ x_2^c &= 10 \end{aligned}$$

It is worth noting that the minimum possible variance depends only on V_L and on the various correlation coefficients; that the values of x_1 and x_2 at this point introduce further only σ_1 and σ_2 respectively; that P_C depends additionally only on the prices and E_C only on the expected values.

Minimum variance as a function of price

29. We can also consider the portfolios that give the minimum variance for a given price, regardless of expected surplus. One way of deriving these is to consider the projection of the elliptic paraboloid onto the plane $E=0$. The outline or profile of the projection of the surface is itself a parabola, giving V as a quadratic in P. It can be calculated as:

$$V = \frac{(ab-h^2)P^2 + 2(bg-hf)P + (bc-f^2)}{b}$$

$$= \frac{\left\{ \begin{aligned} &(V_1V_2 - C_{12})P^2 - 2[P_1(V_2C_{1L} - C_{12}C_{2L}) + P_2(V_1C_{2L} - C_{12}C_{1L})]P \\ &\quad + V_L(P_2^2V_1 - 2P_1P_2C_{12} + P_1^2V_2) - (P_1C_{1L} - P_2C_{1L})^2 \end{aligned} \right\}}{(P_2^2V_1 - 2P_1P_2C_{12} + P_1^2V_2)} \qquad (29.1)$$

or $V = \cdot0000053037\ P^2 - \cdot0106073501\ P + 5\cdot3036750484$, in our particular example. This profile is shown in Figure 6, in which the minimum variance parabola is seen to be the outer envelope of all possible cross-sections of the surface for various values of E, a few of which are also shown.

30. The minimum variance portfolio for any value of P can readily be calculated from this expression, and is just what we have found above as (P_C, E_C, V_C). Each of these portfolios represented by the above parabola lie along the line joining the westernmost and easternmost points of the ellipses, viz:

$$E = -\frac{hP + f}{b},$$

and the quantities of each security in these portfolios can be derived as

$$x_1 = \frac{(P_1V_2 - P_2C_{12})P + (P_2C_{1L} - P_1C_{2L})P_2}{(P_2^2V_1 - 2P_1P_2C_{12} + P_1^2V_2)}$$

$$x_2 = \frac{(P_2V_1 - P_1C_{12})P + (P_1C_{2L} - P_2C_{1L})P_1}{(P_2^2V_1 - 2P_1P_2C_{12} + P_1^2V_2)}$$

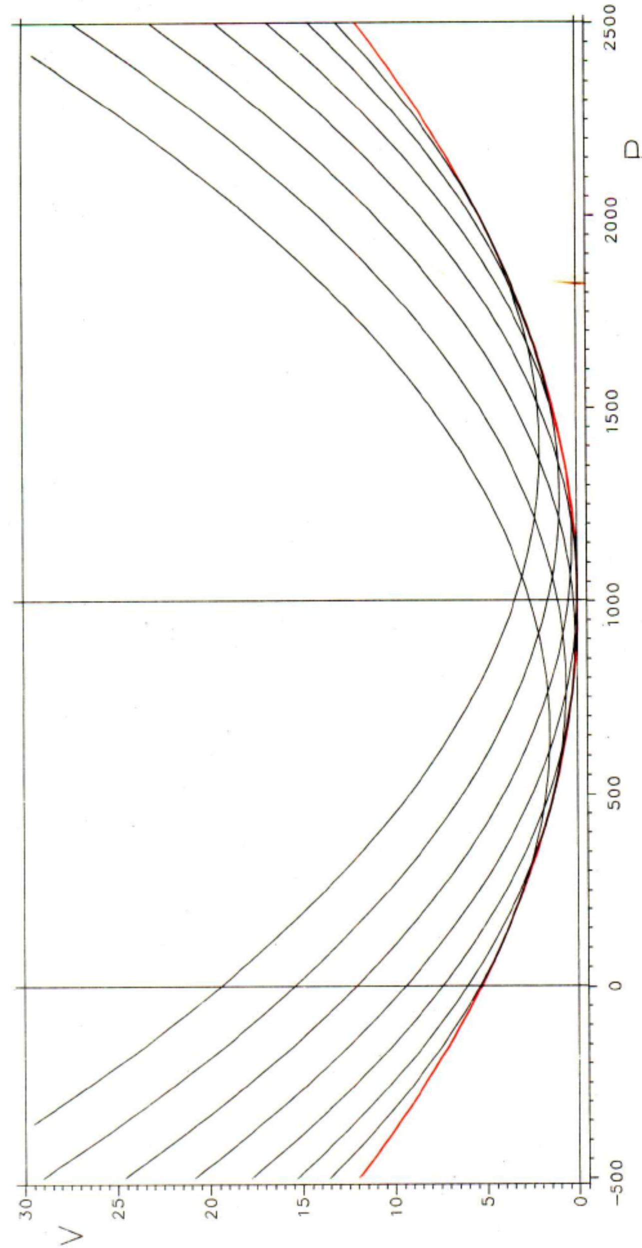

Figure 6
Profile of Feasible Region projected onto plane E=0
With cross-sections of surface for fixed E=−500 (500) 2500

Minimum variance as a function of expected surplus

31. Similarly the portfolios that give the minimum variance for a given expected surplus, regardless of price, are given by the projection of the profile of the surface onto the plane $P=0$, or:

$$V=\frac{(ab-h^2)E^2+2(af-hg)E+(ac-g^2)}{a}$$

$$=\left\{V_1V_2-C_{12})E^2-2[E_1(V_2C_{1L}-C_{12}C_{2L})+E_2(V_1C_{2L}-C_{12}C_{1L})]E+V_L(E_2^2V_1-2E_1E_2C_{12}+E_1^2V_2)-(E_1C_{2L}-E_2C_{1L})^2\right\}\Big/(E_2^2V_1-2E_1E_2C_{12}+E_1^2V_2)$$

or $V=\cdot0000007825\ E^2-\cdot0014085831\ E+\cdot6338623827$, in our particular example. This profile is shown in Figure 7. It may be noted that this profile does not depend at all on the values of P_1 and P_2, just as the profile in paragraph 29 above does not depend on the values of E_1, E_2 and E_L. The minimum values of V for the two profiles are, however, equal, and as noted above do not depend on any of the prices or the expected values.

32. The portfolios represented by this profile lie along the line joining the southernmost and northernmost points of the ellipses, viz:

$$E=-\frac{aP+g}{h}.$$

Region of efficient portfolios

33. The elliptic paraboloid represents all feasible portfolios. The region of efficient portfolios is restricted to the 'north-western' quadrant of each ellipse, or the region to the north-west of the straight lines connecting the centre of the ellipses with the westernmost and the northernmost points of each ellipse, viz:

$$E=-\frac{hP+f}{b}$$

and

$$E=-\frac{aP+g}{h}$$

or

$$E=1\cdot664654P-764\cdot6538$$

and

$$E=4\cdot071393P-3171\cdot3930,$$

in our example. These are the lines of minimum variance regardless of E and minimum variance regardless of P respectively, just discussed above. They meet at the centre of the ellipses, and are shown in Figure 8, along with a few of the ellipses.

34. The region of the efficient portfolios in the P–E–V space is therefore that quadrant of the elliptic paraboloid sliced off by the planes

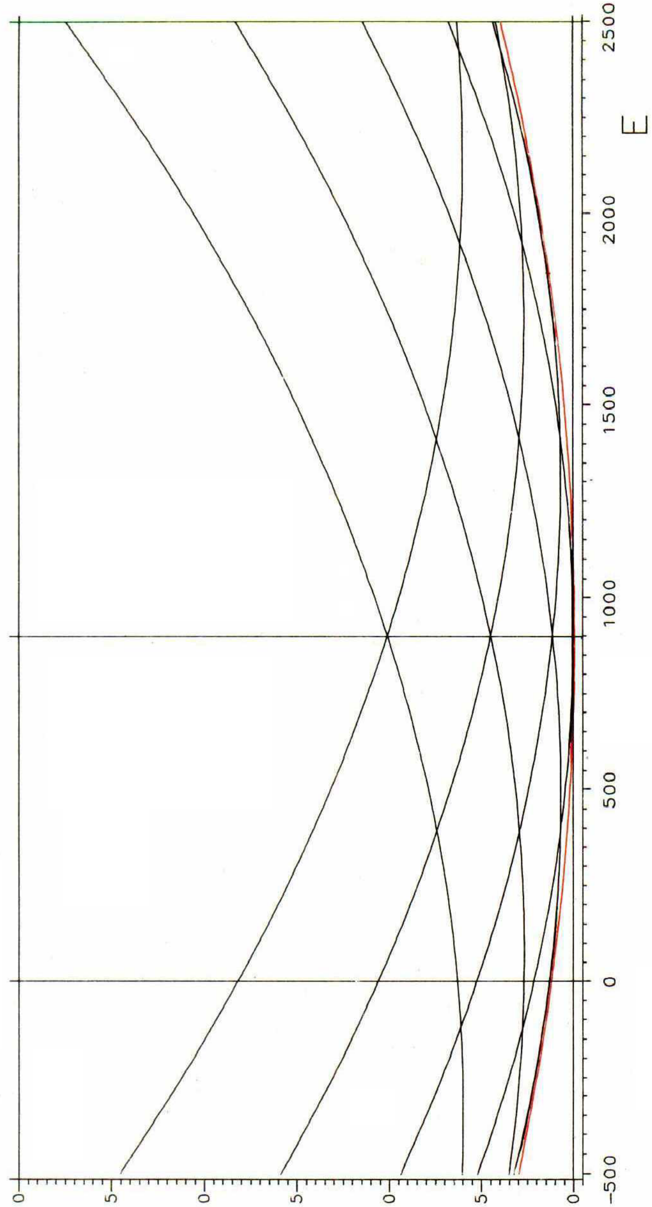

Figure 7
Profile of Feasible Region projected onto plane P=0
With cross-sections of surface for fixed P= 250 (250) 1750

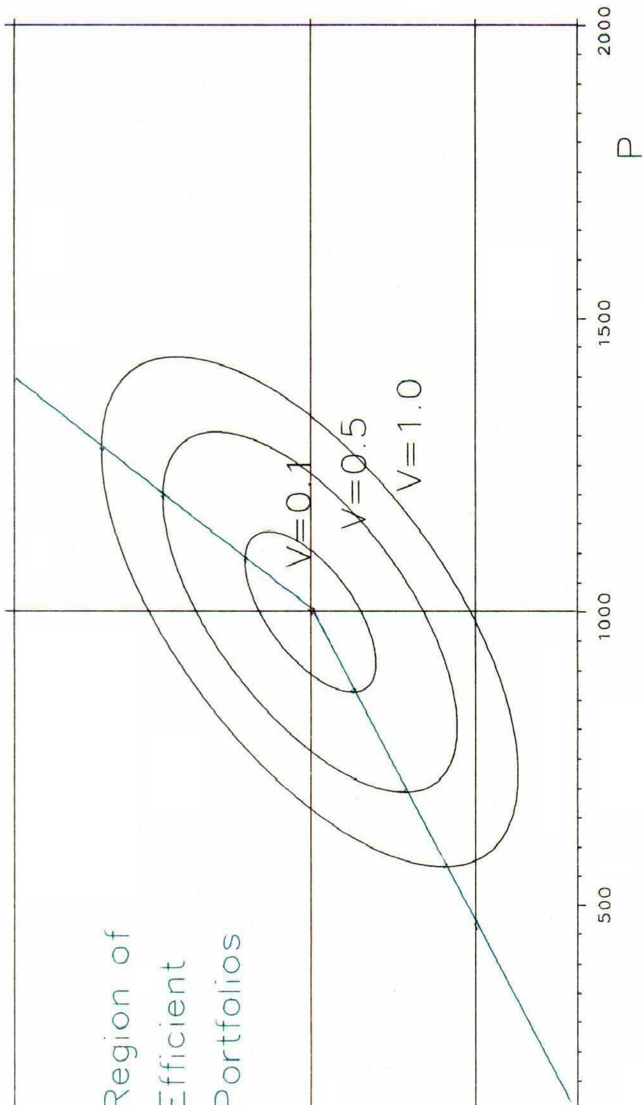

Figure 8
Expected (E) as a function of Price (P)
for fixed Variances V=0.1, 0.5, 1.0

$$bE+hP+f=0$$

and

$$hE+aP+g=0,$$

and lying to the 'north-west' of these planes, i.e. in the direction of higher E, lower P.

Wise's unbiased match

35. Now consider Wise's unbiased match, that is, the portfolio that minimizes $G=E^2+V$, subject to $E=0$, i.e. minimizes V subject to $E=0$. Consider the plane $E=0$. This cuts the quadric surface in a parabola. The minimum value of V in this parabola is found at the point where the line $E=0$ touches the contour line with lowest V (in the projection onto the P–E plane), among all those contour lines that it cuts. This must be at the 'southernmost' or 'northernmost' tip of the corresponding ellipse. In this example it is at the 'southernmost' tip, since the centre of the concentric ellipses has positive E ($E_C=900$) and so is north of the line $E=0$. The portfolio represented by this point cannot be an efficient one, since it is dominated by all the portfolios that, for the same V, have both higher E and lower P, and by the one portfolio that, for the same V, has higher E and the same P. This is shown in Figure 9.

36. Note that the 'inefficiency' of Wise's unbiased match does not depend (in general) on the prices of the securities, since the expected return for the minimum variance portfolio (E_C) does not depend on the values of P_1 and P_2. Only in a further special case discussed in § 67 below is Wise's unbiased match an efficient portfolio.

37. A similar argument shows that Wise's (unqualified) match is also not an efficient portfolio. This is the portfolio with minimum $G=E^2+V$. Suppose that it is found at some value of $E=E_G$. Then the argument of § 35 can be repeated, replacing the line $E=0$ by $E=E_G$. In general this line touches a contour ellipse at its 'southernmost' point (in Wise's example), which again is not an efficient portfolio.

38. In a different example the unbiased match and the unqualified match may touch the appropriate contour ellipses at their 'northernmost' points. This will occur if the expected surplus of the minimum variance portfolio, E_C, is negative. In this case Wise's matches are efficient portfolios, but not the only ones, and not necessarily the ones that the investor may prefer.

Investors' preferences

39. How does the investor choose among all the possible efficient portfolios? Consider the series of parallel planes $E-\lambda P-\mu V=k$, with λ, $\mu \geqslant 0$, for various values of k. The perpendicular to these planes runs in the desirable direction, viz higher E, lower P, lower V. If we choose that plane that gives us the highest k, but still contains a feasible portfolio, we shall find that that plane contains just one feasible (and efficient) portfolio, and that it just touches the quadric surface. One

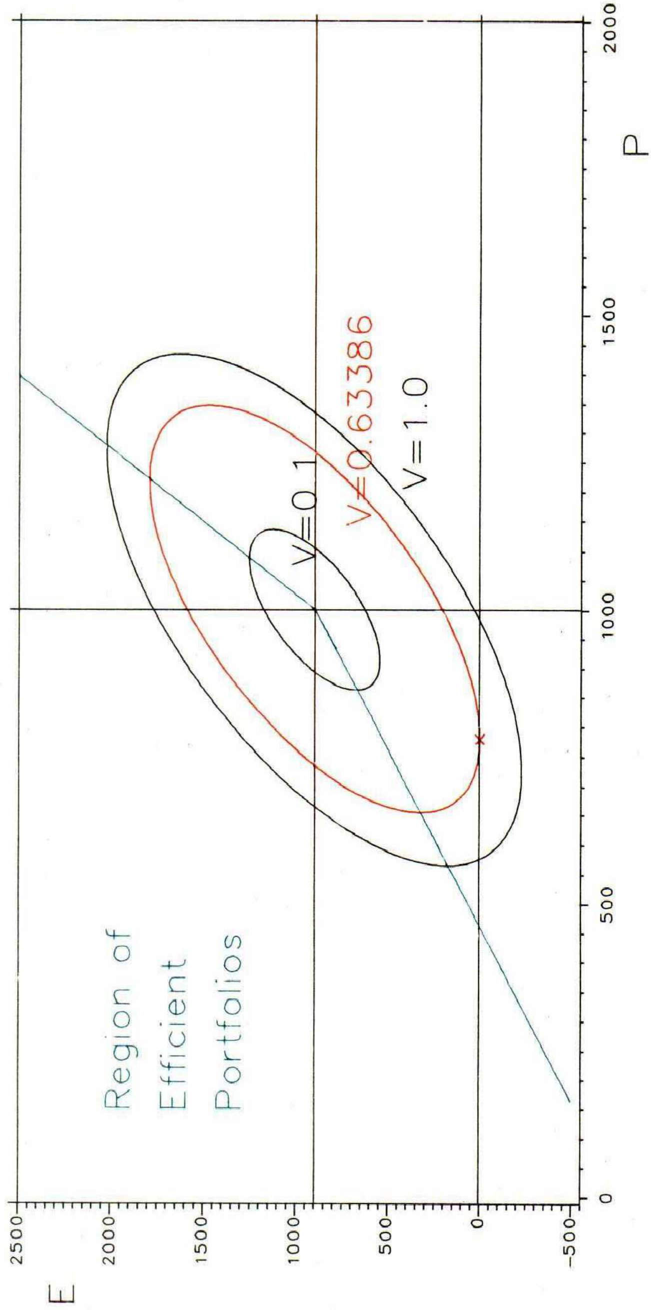

Figure 9
Expected (E) as a function of Price (P)
for fixed Variances V=0.1, 0.63386, 1.0
X=Wise's Unbiased Match

can vizualize turning the bowl upside down and approaching it with a plane surface held at a fixed angle. If it touches smoothly, the slopes of the surface at this point are the same as the slopes of the plane, and the various partial derivatives represent the marginal trade-offs between the various desirable or undesirable features.

40. For example, keeping V constant, we have

$$\frac{\partial E}{\partial P} - \lambda = 0, \text{ or } \frac{\partial E}{\partial P} = \lambda.$$

Thus, at the tangential point, the marginal extra expected surplus per unit of price is equal to λ, or the marginal rate of return is $\lambda - 1$. The investor could thus choose a fixed V and travel round the appropriate ellipse from its 'westernmost' point at which $\lambda = \infty$, so that the marginal expected return per unit price is infinite, moving towards its 'northernmost' point, at which $\lambda = 0$, and stopping when the slope $\partial E/\partial P$ is such as to suit his particular preference for additional expected return, which he may base on the expected return obtainable elsewhere in the market for investments with a similar variance.

41. Whichever fixed value of V he has chosen, he will find that $\partial E/\partial P = \lambda$ at a point which lies on a straight line from the centre of the ellipses and has slope

$$-\frac{a + h\lambda}{h + b\lambda}.$$

At his (temporarily) chosen point, he will find that $\partial E/\partial V$ has some particular value. He can then travel outwards or inwards along the straight line (thus keeping $\partial E/\partial P$ constant), until $\partial E/\partial V$ is equal to his chosen μ. It may be better to express this as $\partial V/\partial E$, and equate it to $1/\mu$. This gives us the marginal trade-off between variance and expected surplus. Alternatively we can consider

$$\frac{\partial V}{\partial P} = -\frac{\lambda}{\mu}$$

and carry out the same exercise. This gives the marginal trade-off between variance and price, i.e. the amount the investor is willing to marginally increase his price by in order to reduce the variance of his ultimate surplus.

42. Lines of constant $\partial E/\partial P = \lambda$, as noted above, radiate out from the centre of the ellipses and have equation:

$$P(a + h\lambda) + E(h + b\lambda) + (g + f\lambda) = 0.$$

Lines of constant

$$\frac{\partial V}{\partial E} = \frac{1}{\mu}$$

lie parallel to the line joining the 'western' and 'eastern' points of the ellipse (along which $\partial V/\partial E = 0$), and are given by

$$hP + bE + f = 1/2\mu.$$

If the values of λ and μ are given these can readily be solved, giving

$$P = P_C + \frac{h + b\lambda}{2\mu\,(ab - h^2)},$$

and

$$E = E_C + \frac{a + h\lambda}{2\mu\,(ab - h^2)}.$$

43. For example, in our particular case, if the investor were to choose $\lambda = 1{\cdot}2$, $\mu = 4{,}800$, so that $-\lambda/\mu = -{\cdot}00025$, indicating that he was prepared to pay, at the margin, an extra £1 of price for an extra £1·2 of expected surplus, and an extra £1 of price for a reduction of £^{2}·00025 of variance, his chosen investment position would be such that

$$\begin{aligned} P &= 1{,}009{\cdot}1 \\ E &= 993{\cdot}9 \\ V &= {\cdot}008638 \\ x_1 &= -{\cdot}2233 \\ x_2 &= 10{\cdot}9844. \end{aligned}$$

Alternatively, if he chose $\lambda = 2{\cdot}0$, $\mu = 2{,}000$, so that $-\lambda/\mu = -{\cdot}001$, his chosen investment position would be such that

$$\begin{aligned} P &= 984{\cdot}2 \\ E &= 1{,}062{\cdot}5 \\ V &= {\cdot}048538 \\ x_1 &= -{\cdot}4914 \\ x_2 &= 11{\cdot}8076. \end{aligned}$$

These points are marked as points L and M respectively in Figure 10.

Investors' preferences in P–E–σ *space*

44. An alternative is to consider the surface in the P–E–σ space, and the parallel planes $E - \lambda P - v\sigma = k$. We can then consider the marginal trade-off between price, expected return and standard deviation (rather than variance), and choose a portfolio to maximize k above. In general we may be able to choose a portfolio to give us values so that

$$\frac{\partial E}{\partial P} = \lambda,\ \frac{\partial \sigma}{\partial E} = \frac{1}{v} \text{ or } \frac{\partial \sigma}{\partial P} = \frac{-\lambda}{v}.$$

However, in this particular case the surface in the P–E–σ plane is an elliptical

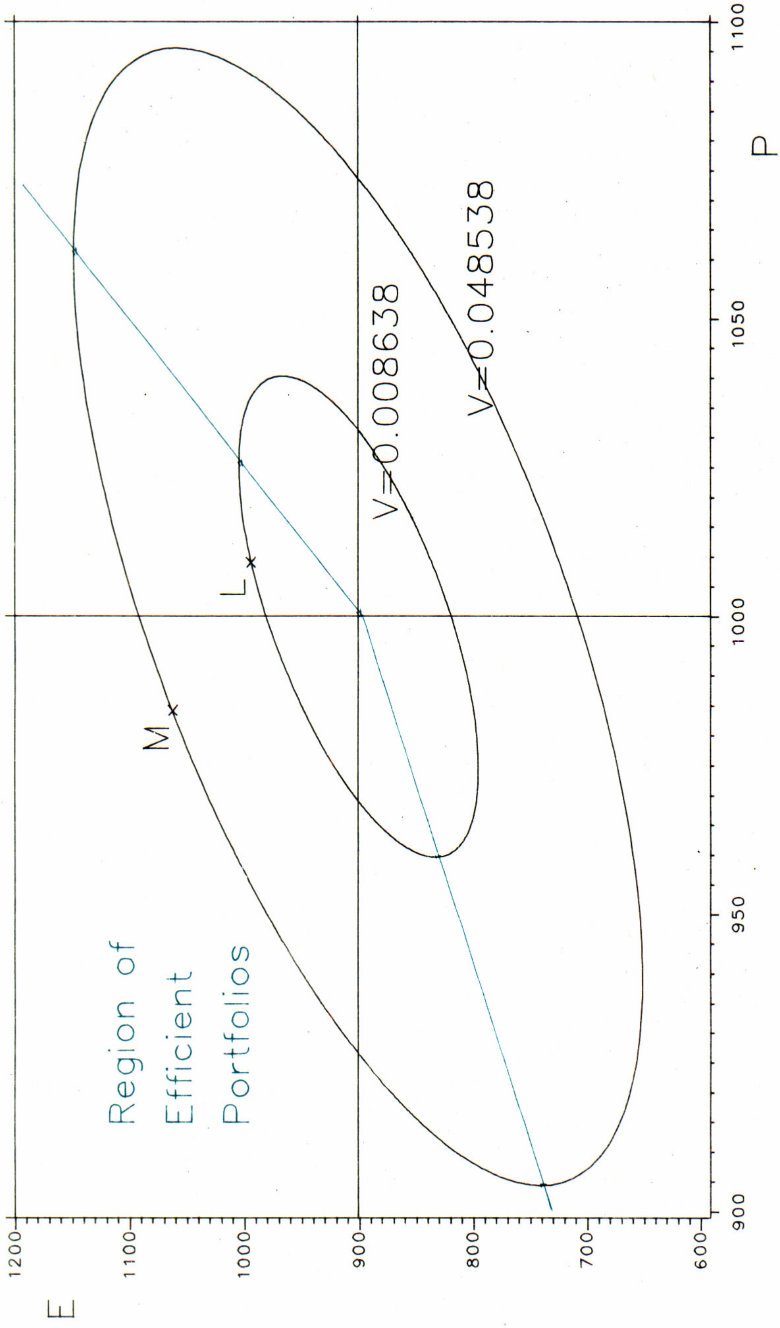

Figure 10
Optimum unconstrained Portfolios L and M for two investors

cone, which makes it rather a special case. We can still choose a radial line so that $\partial E/\partial P=\lambda$, just as in the previous case. But if we imagine a vertical cone, and try to approach it with a moveable plane held at a fixed angle, we can see that, if the angle is shallow, we first touch the cone at the vertex; if the angle is increased to exactly a certain value, we shall touch all the way along one side of the cone; and above that value, we shall not touch it at all except at infinity.

45. Thus, choice of λ determines the radial line from the centre of the ellipses along which we seek an optimum. If v is above a certain critical value (which depends on λ), so that the investor places a high value on avoiding risk, his optimum will be at the vertex of the cone, i.e. with $\sigma=0$. If v equals this critical value, he will be indifferent between all points on the radial line determined by λ. If v is lower than this critical value, so that he places a low value on avoiding risk, he will wish to go infinitely far along the radial line, thus accepting an infinite standard deviation, and also getting an infinite expected surplus. Whether he gets this at a positively or a negatively infinite price depends on the direction of the radial line.

46. In the general case, the minimum standard deviation is not zero, and the surface in the P–E–σ space is a proper elliptic hyperboloid. In this case, the 'nose' of the surface is rounded, not pointed, and provided v is above a certain value (which still depends on λ) a 'smooth' contact will be made. If the value of v equals the critical value, the planes $E-\lambda P-v\,\sigma=k$ are parallel to one of the asymptotes of the hyperboloid, and the optimum portfolio is again infinitely far in that direction, as is the case when v is less than this critical value.

Region of positive portfolios

47. It will be noted that the solutions for the specimen values of λ and μ in the numerical example above produced negative values of x_1, implying that the investor had to purchase a negative quantity of security 1, or 'sell it short'. For some investment intermediaries this may be a practicable possibility. An insurance company, for example, may issue fixed interest stock and could presumably obtain about as good a market price for such an issue as it would have to pay to obtain similar fixed interest stock of a similar security. However, insurance companies seldom in fact do this, and pension funds may not have powers to do it (though they could insure the whole scheme with an insurance company that did have such powers). Any investment intermediary is therefore concerned also with considering the portfolios that are feasible with non-negative holdings of the available securities, one of which Wise calls the 'positive match'. Generalizing, I shall call all feasible portfolios with non-negative x_i's positive feasible portfolios, and the efficient portfolios among these the set of positive efficient portfolios.

48. The region of positive feasible portfolios is, in our present case, subject to the constraints

$$x_1 \geqslant 0,$$

and

$$x_2 \geqslant 0,$$

and is bounded by the portfolios in which

$$x_1 = 0,$$

and

$$x_2 = 0, \text{ respectively.}$$

If $x_1 = 0$, then

$$P = x_2 P_2,$$
$$E = x_2 E_2 - E_L,$$

and the line of such portfolios in the P–E plane is given by

$$E = PE_2/P_2 - E_L.$$

Similarly, if $x_2 = 0$, the line of such portfolios is given by

$$E = PE_1/P_1 - E_L.$$

These two lines meet at $P = 0$, $E = -E_L$ (where $x_1 = 0$, $x_2 = 0$) and the positive feasible region is the 'wedge' lying to the 'northeast' of this point, between these lines.

49. This region is marked on Figure 11. It will be seen that in this case the minimum variance portfolio is a positive feasible portfolio (at which $x_1 = 0$, $x_2 = 10$); that south-west of this point the positive feasible region includes part of the original (unconstrained) efficient region; and that north-east of the minimum variance point the positive feasible region is wholly contained within the original region of inefficient portfolios. Remembering the definition of an efficient portfolio, we see that the region of positive efficient portfolios contains that part of the original efficient region just mentioned, bounded by the planes along which $x_1 = 0$ (the upper line in Figure 11) and $x_2 = 0$ (the lower line in Figure 11), as far as the minimum variance portfolio, together with all portfolios lying along that plane (line) to the north-east of the centre. The investor still has a considerable choice of portfolios.

50. In this particular case, the constraint lines included the minimum variance portfolio (just) and included a part of the original efficient region. In other examples, these lines might lie wholly inside or wholly outside the original efficient region.

Optimization in the positive region

51. An investor who wishes to maximize $E - \lambda P - \mu V$, subject to the constraints we have introduced, will need to discover where the parallel planes $E - \lambda P - \mu V = k$ just touch the positive efficient region. For certain values of λ and μ, the planes will make a smooth contact in the region that is part of the unconstrained efficient region; for this investor, the constraints will not be binding. Otherwise the planes will make a 'sharp' contact somewhere along the boundary of the

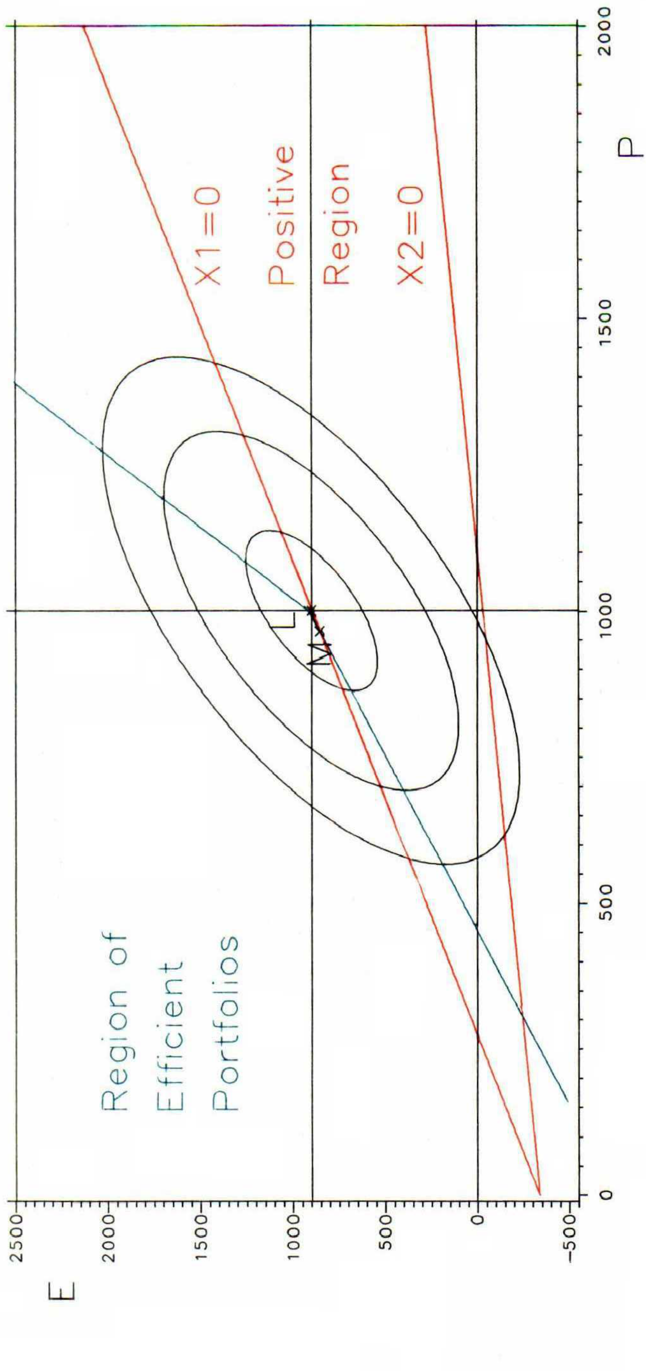

Figure 11
Boundaries of Positive Region are lines where X1=0 and X2=0
Also optimum Positive Portfolios L and M for two investors

positive efficient region, either along the line where $x_1=0$ or where $x_2=0$. We consider the former case from first principles.

52. If $x_1=0$, we have

$$P=x_2P_2$$
$$E=x_2E_2-E_L$$
$$V=x_2^2\mathrm{L}V_2-2x_2C_{2L}+V_L$$

Let $F=E-\lambda P-\mu V=(x_2E_2-E_L)-\lambda(x_2P_2)-\mu(x_2^2V_2-2x_2C_{2L}+V_L)$
F is maximized if:

$$\frac{\partial F}{\partial x_2}=E_2-\lambda P_2-2\mu x_2V_2+2\mu C_{2L}=0$$

$$x_2=\frac{E_2-\lambda P_2+2\mu C_{2L}}{2\mu V_2}$$

At this point

$$P=P_2\left(\frac{E_2-\lambda P_2+2\mu C_{2L}}{2\mu V_2}\right),$$

$$E=E_2\left(\frac{E_2-\lambda P_2+2\mu C_{2L}}{2\mu V_2}\right)-E_L,$$

and

$$V=\frac{(E_2-\lambda P_2)^2}{4\mu^2V_2}-\frac{C_{2L}^2}{V_2}+V_L.$$

53. The investors in the numerical example in paragraph 43 find themselves constrained by the non-negativity constraints. Their optimum positive portfolios are found, for $\lambda=1{\cdot}2$, $\mu=4{,}800$, at:

$$P=1{,}000{\cdot}5$$
$$E=900{\cdot}6$$
$$V=\cdot000002$$
$$x_1=0$$
$$x_2=10{\cdot}0052$$

and, for $\lambda=2{\cdot}0$, $\mu=2{,}000$, at:

$$P=965{\cdot}3$$
$$E=857{\cdot}3$$
$$V=\cdot006707$$
$$x_1=0$$
$$x_2=9{\cdot}6526.$$

These points are marked as L and M in Figure 11. It may be noted that the first investor chooses a point a little to the north-east of the minimum variance portfolio but very close to it, the second a little further away to the south-west.

The $k\sigma$-solvency region

54. A further approach to choosing an optimal portfolio for a particular investor is the use of what I shall call the '$k\sigma$-solvency region'. Assume that the asset and liability returns and hence the ultimate surplus are all normally distributed, with the mean and variance that we have already assumed. This is not true in the particular example I have been using, which is discussed below, but it is not difficult to see how we could generalize the example so that, instead of two rates of interest each year with equal probabilities, we had an appropriate normal distribution of rates of interest. Now assume that the investor wishes to have a particular (large) probability that his ultimate surplus is positive, so that he is able to meet his ultimate liability, or is 'solvent'. Correspondingly he wishes a particular (small) probability, the complement of the first, that his ultimate surplus is negative or he is 'insolvent'.

55. He can ensure this by choosing assets so that the probability that the ultimate surplus is negative is less than or equal to some small probability α, or that $P(S<0)\leqslant\alpha$, or $E-k\sigma\geqslant 0$, where $P(z<-k)=\alpha$, where Z is a unit normal variate. For example, if he wishes a 1/100 probability of insolvency, or a 99/100 probability of solvency, he will wish to ensure that $E-2{\cdot}326\sigma\geqslant 0$, because $P(z<-2{\cdot}326)=0{\cdot}01$ for a unit normal variate. The region (if any) within which $E-k\sigma\geqslant 0$ I shall call the $k\sigma$-solvency region. Its shape and size (and even existence) depend on the value of k.

56. Consider the plane $E-k\sigma=0$ in the P–E–σ space. If $k=0$, this is the plane $E=0$. As $k\to\infty$, it becomes the plane $\sigma=0$. The plane cuts the elliptic hyperboloid, the surface giving the region of feasible portfolios, in a conic section. The projection of this conic section onto the plane $\sigma=0$ is also a conic section, in general of the same class, which outlines the $k\sigma$-solvency region on the P–E plane. Consider also the projection of the elliptic hyperboloid onto the plane $P=0$. This is bounded by a hyperbola (or by two straight lines). The projection of the plane $E-k\sigma=0$ is the line $E-k\sigma=0$. We can start by considering these projections in the E–σ plane.

57. If $k=0$, the condition becomes the plane $E=0$, which cuts the surface in a hyperbola; but the projection of this surface onto the P–E plane is just the line $E=0$. The $k\sigma$-solvency region with $k=0$ is just the region 'north' of the line $E=0$. As k increases, the line $E-k\sigma$ tilts towards the E-axis, continuing to cut the surface in a hyperbola, whose projection onto the P–E plane is also a hyperbola. Consider the case where the vertex of the hyperboloid is 'north' of the line $E=0$, i.e. $E_C>0$. At some value of k, the line $E=k\sigma$ is parallel to the asymptote of the hyperbola, and the plane intersects the hyperboloid in a parabola. For higher values of k, the intersection is an ellipse, whose projection is also an ellipse. As k increases the plane $E-k\sigma=0$ will first touch the surface (if it is a proper hyperboloid) and then miss it altogether. In our particular example, where the

hyperboloid is a cone with vertex on the plane $\sigma=0$, as $k\to\infty$, the ellipse tends towards a single point.

58. A general example is shown in Figure 12, and the specific example in Figures 13 and 14, which show the lines $E-k\sigma=0$ for several values of k and the projections of the $k\sigma$-solvency region onto the P–E plane for $k=10$ (a hyperbola) and $k=2000$ (an ellipse). In general, this projection is given by substituting $\sigma=E/k$ in the general equation of the ellipse, giving:

$$\left(\frac{E}{k}\right)^2 = aP^2 + 2hPE + bE^2 + 2gP + 2fE + c$$

or

$$aP^2 + 2hPE + \left(b-\frac{1}{k^2}\right)E^2 + 2gP + 2fE + c = 0.$$

It can be seen that the new discriminant of this conic section, its '$ab-h^2$', is given by

$$a\left(b-\frac{1}{k^2}\right)-h^2.$$

If k is infinite, this is unchanged from the original '$ab-h^2$', but as k reduces, the value of the discriminant also reduces. It equals zero, making the boundary of the $k\sigma$ – solvency region a parabola, when

$$k^2=\frac{a}{ab-h^2}.$$

Below this value of k the discriminant is negative, making the boundary of the $k\sigma$-solvency region a hyperbola. As already noted, when $k=0$, the boundary of the $k\sigma$-solvency region is just the line $E=0$. In our particular example, this critical value of k is 1,130·4, an extraordinarily high value for any investor to wish to choose. Any reasonable $k\sigma$-solvency region in this case would therefore be bounded by a hyperbola.

59. Consideration of the $k\sigma$-solvency region does not provide us with one optimum portfolio, but it does put a constraint on the portfolios acceptable to the investor. He may then choose the portfolio which maximizes, say, $E-\lambda P$, subject to the constraint $E-k\sigma\geqslant 0$. In the general case, of course, he may not be able to satisfy some particular $k\sigma$ constraint at all, if k is too big, and he may find that any $k\sigma$-solvency region does not intersect with the region of positive portfolios. If he cannot meet his desired solvency requirement he may then choose, for example, the portfolio which maximises $E-k\sigma$, regardless of price.

The absolute solvency region

60. In our particular example, there are in fact only four possible outcomes,

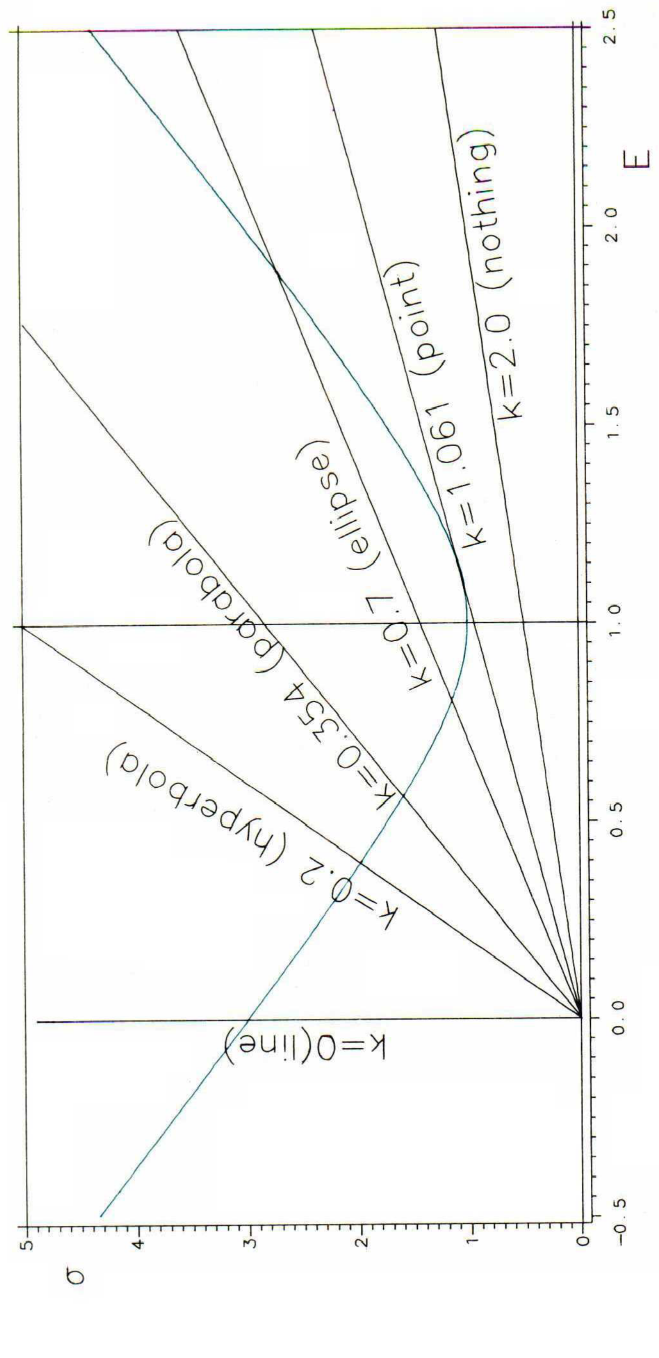

Figure 12. Profile of Feasible Region projected onto plane P=0 Together with k−σ lines for k such that the projections onto the P−E plane are line, hyperbola, parabola, ellipse, point and nothing Artificial data

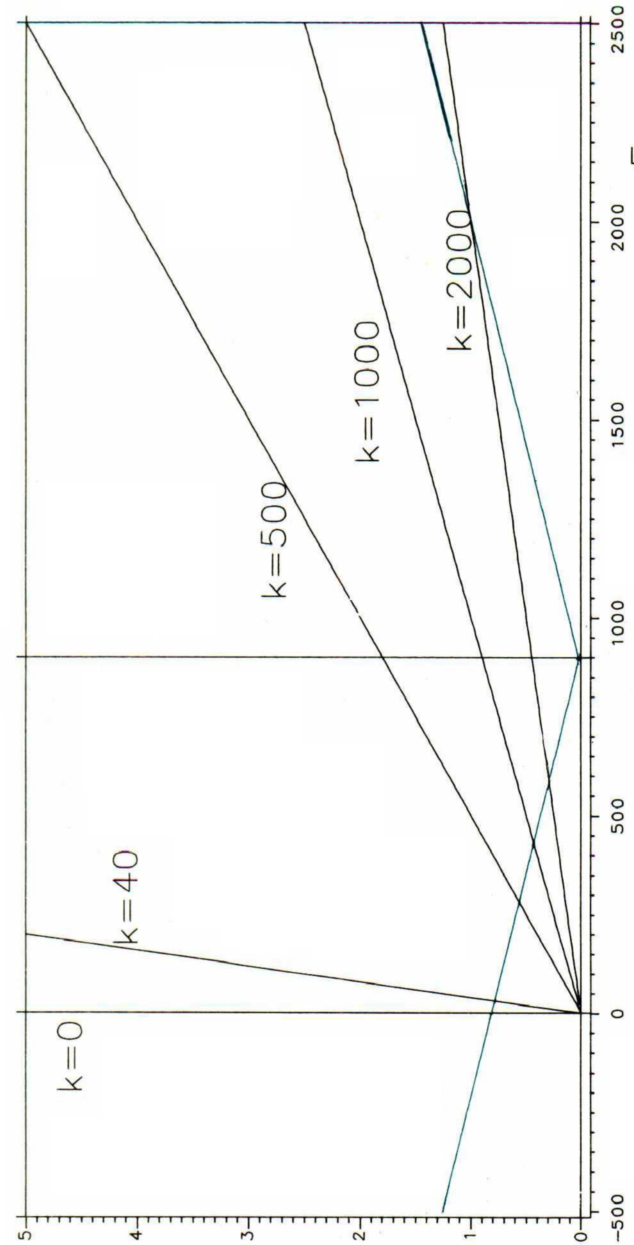

Figure 13
Profile of Feasible Region projected onto plane P=0
Together with k−σ lines for k=0, 40, 500, 1000 & 2000
Data as in Wise with P1=400 P2=100

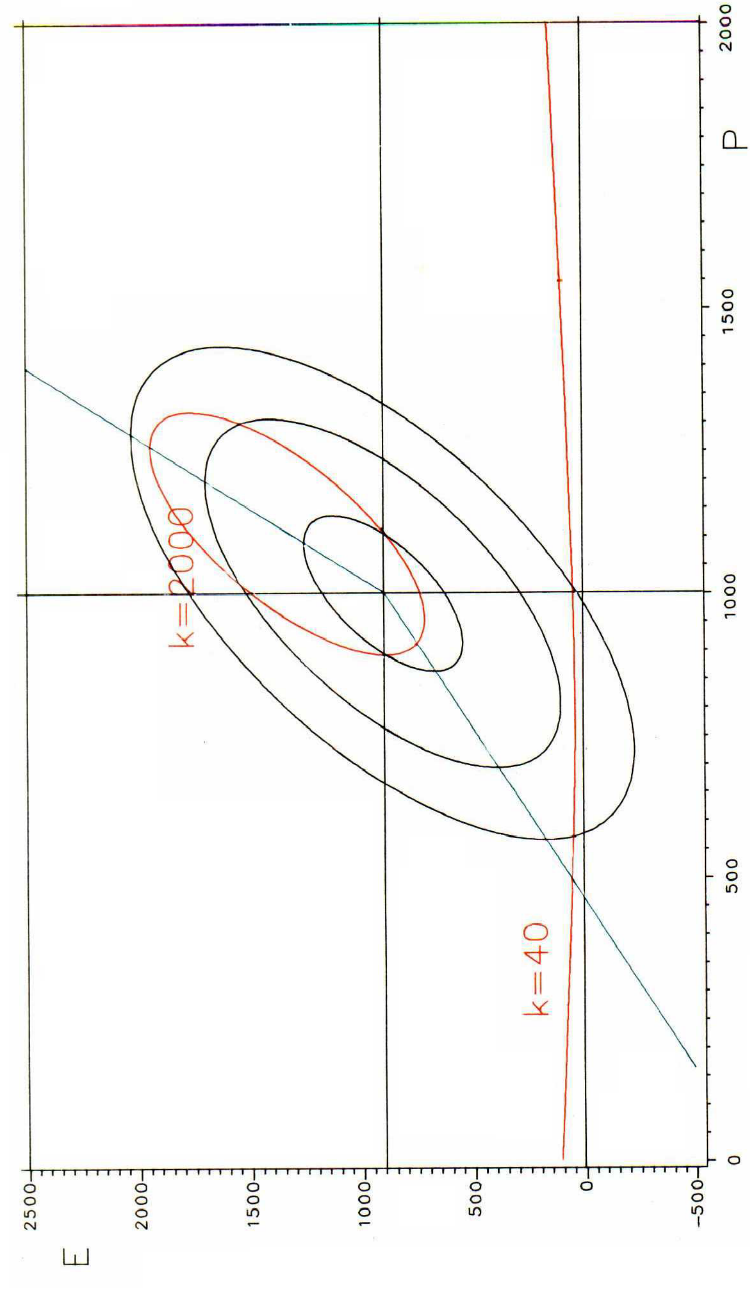

Figure14
Boundaries of kσ–Solvency Regions for k=40 and k=2000

each one of which has an equal probability of occurring. It is easy to see that the ultimate surplus is certainly positive, if it is positive in each of the possible outcomes. This will be the case if all four of the following conditions are met (see Wise, 3.14):

$$119{\cdot}66x_1+122{\cdot}46x_2 \geqslant 324{\cdot}64$$
$$121{\cdot}88x_1+122{\cdot}88x_2 \geqslant 328{\cdot}80$$
$$119{\cdot}88x_1+122{\cdot}68x_2 \geqslant 326{\cdot}80$$
$$122{\cdot}10x_1+123{\cdot}10x_2 \geqslant 331{\cdot}00.$$

Substituting for x_1 and x_2 according to formulae (22.1) and (22.2) and substituting numerical values we get:

$$E \geqslant {\cdot}003001\,P-3{\cdot}17$$
$$E \geqslant -{\cdot}002995\,P+0{\cdot}99$$
$$E \geqslant {\cdot}002985\,P-1{\cdot}01$$
$$E \geqslant -{\cdot}003001\,P+3{\cdot}19.$$

61. The portfolios which satisfy all of these inequalities can be represented either in the x_1–x_2 plane or in the P–E plane as lying in the region formed by the intersection of the half-planes lying to the appropriate side of the corresponding equalities. In this case, the portfolios lie to the north of the boundary shown in Figure 15. We can call such a region the region of 'absolute solvency', and if it exists it is of significance to the investor. He may, for example, choose the positive portfolio that provides him with absolute solvency for the minimum price, which in this case is found at the intersection of the line $x_1=0$ with the boundary of the region of absolute solvency. This portfolio gives:

$$\begin{aligned} P &= 268{\cdot}87 \\ E &= 2{\cdot}33 \\ V &= 2{\cdot}9700 \\ x_1 &= 0 \\ x_2 &= 2{\cdot}68887 \end{aligned}$$

However, neither Wise nor I wish to limit ourselves to such cases, and have chosen in general to work with expected values and variances, without considering the distributions in further detail.

Price as a constraint

62. Yet a further constraint that may affect any particular investor is the current amount of funds available, or the price, P, he can pay for securities to meet the liability. His feasible region is then bounded by the plane

$$P \leqslant P_{MAX}$$

where P_{MAX} is the maximum price he is able to pay. He may further be constrained by a minimum price, if, for example, there is already a pension fund

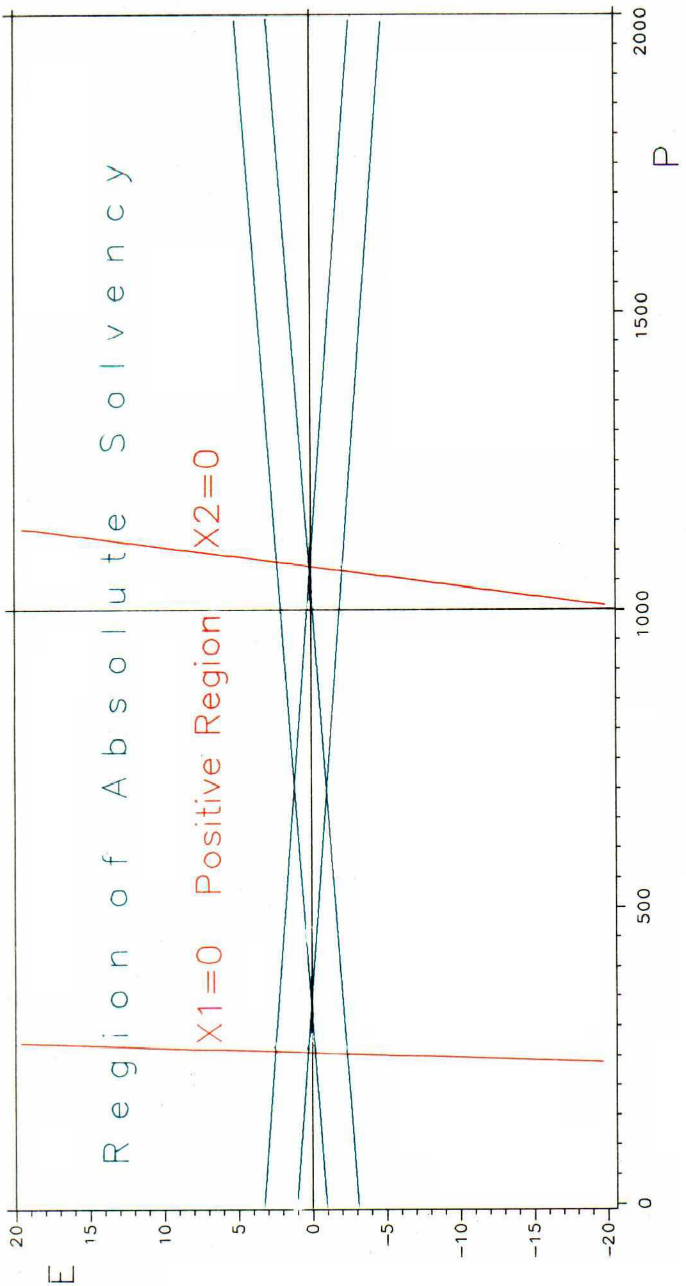

Figure 15
Boundaries of Positive Region and of Region of Absolute Solvency

from which he may not withdraw funds, though he may be able to add a contribution, so that

$$P_{MIN} \leqslant P \leqslant P_{MAX}.$$

63. Such constraints bound the region which the investor may consider by the two planes $P = P_{MIN}$ and $P = P_{MAX}$. If the investor, following one or other of the criteria discussed above, finds that his optimum portfolio satisfies these bounds, then the price is not, in fact, constraining him. If, however, he finds that his optimum portfolio (so far) breaks these bounds, then he must search for the optimum along one or other boundary. This reduces almost to the conventional portfolio selection problem, where there is simply a trade-off between E and V or E and σ, i.e. the investors maximize $E - \lambda V$ or $E - \nu\sigma$, for the fixed price $P = P_{MAX}$ or $P = P_{MIN}$ as appropriate.

The investors' optimum portfolio

64. The various considerations that have been put forward above give ways in which a rational investor might find his way to an optimum portfolio that suits him. They do not prescribe any particular method, or any particular portfolio. The particular portfolio that is optimum for the investor depends on his choice of criterion, and on the values of λ,μ or λ,ν or k, that he considers appropriate. The optimum portfolio also depends on whether or not he is bound by the positive constraints $x_i \geqslant 0$, and on whether or not he is bound by any price constraints. However, the rationale discussed above should eliminate any portfolio that is not efficient (within the relevant constraints) in favour of efficient portfolios, for all investors who are willing to measure the distribution of ultimate surplus by the mean and variance alone, and to ignore higher moments of the distribution (in so far as these are not determined by the mean and variance anyway).

A special case

65. In § 22 above we saw that we could solve the equations

$$x_1 P_1 + x_2 P_2 = P,$$
$$x_1 E_1 + x_2 E_2 = E + E_L,$$

for x_1 and x_2 in terms of E and P, provided that $E_1 P_2 - E_2 P_1 \neq 0$. I now wish to consider this special case.

$$E_1 P_2 - E_2 P_1 = 0 \text{ gives}$$

$$\frac{E_1}{P_1} = \frac{E_2}{P_2} = r \text{ say,}$$

that is, the returns per unit price on the two assets are equal. We find that, whatever the values of x_1 and x_2,

$$E = rP - E_L,$$

so that the expected surplus is determined only by the total price paid, and not by the proportions of the two securities that make up this price. The variance, however, does depend on the mixture of the two assets within the price.

66. If we put $x_2=(P-x_1P_1)/P_2$, and substitute in equation (14.2) for V, we get V in terms of a quadratic in x for fixed P:

$$\begin{aligned} V = \frac{1}{P_2^2}\{ & x_1^2(P_2^2V_1-2P_1P_2C_{12}+P_1^2V_2) \\ & + 2x_1(PP_2C_{12}-PP_1V_2-P_2^2C_{1L}+P_1P_2C_{2L}) \\ & + (P^2V_2-2PP_2C_{2L}+P_2^2V_L)\} \end{aligned}$$

By differentiating with respect to x_1, setting $\partial V/\partial x_1=0$, and solving for x_1 we get the value of x_1 that gives us the minimum variance portfolio for that fixed P:

$$x_1 = \frac{P(P_1V_2-P_2C_{12})-P_2(P_1C_{2L}-P_2C_{1L})}{(P_2^2V_1-2P_1P_2C_{12}+P_1^2V_2)}$$

From this we get the corresponding value of x_2:

$$x_2 = \frac{P(P_2V_1-P_1C_{12})-P_1(P_2C_{1L}-P_1C_{2L})}{(P_2^2V_1-2P_1P_2C_{12}+P_1^2V_2)}$$

and the minimum possible value of V for fixed P as

$$V = \frac{\left\{\begin{aligned}&(V_1V_2-C_{12}^2)P^2-2\,[P_1(V_2C_{1L}-C_{12}C_{2L})+P_2(V_1C_{2L}-C_{12}C_{1L})]P \\ &+V_L(P_2^2V_1-2P_1P_2C_{12}+P_1^2V_2)-(P_1C_{2L}-P_2C_{1L})^2\end{aligned}\right\}}{(P_2^2V_1-2P_1P_2C_{12}+P_1^2V_2)} \qquad (66.1)$$

This gives the minimum possible values of V as a quadratic in P, a parabola.

67. As $E_2P_1-E_1P_2$ approaches 0, the elliptic paraboloid giving the surface of feasible portfolios in the P–E–V space, which is a sort of oval bowl shape, becomes thinner and narrower, eventually folding up flat. When $E_2P_1-E_1P_2=0$ the region of feasible portfolios is restricted to the plane $E=rP-E_L$, and to an area of that plane above the parabola in V and P defined by equation (66.1) above. This parabola gives the minimum value of V for any P (and E) and hence is the locus of efficient portfolios. If we compare any two points on this parabola, we cannot say that either dominates the other, so all are efficient portfolios.

68. Strictly the parabola defined by equation (66.1) lies in the P–V plane, and is the projection onto the plane $E=0$ of the locus of efficient portfolios described above. It is the same equation as that of the projection, in the general case, of the profile of the elliptic paraboloid discussed in § 29, and defined by equation (29.1) above. If we substitute for P, P_1 and P_2 in terms of E, E_1 and E_2 we can obtain an expression for V in terms of E, yet another parabola, which gives the projection

onto the plane $P=0$, both of the profile of the elliptic hyperboloid in the general case, discussed in § 31 above, and of the efficient parabola in this special case.

69. In this case, the ellipses in the P–E plane collapse into line segments, of which only the end points are efficient portfolios. In this special case, Wise's (unconstrained) match and his unbiased match are efficient portfolios, though neither is necessarily optimum for any particular investor. The investor can still consider a $k\sigma$-solvency region, which may be a section of the set of efficient portfolios, limited necessarily to E above some value, and either limited to E below some other value, or unlimited upwards, depending on the value of k. We can also consider maximizing the functions $E-\lambda P-\mu V$ or $E-\lambda P-\nu\sigma$, but the contact of the planes with the locus of efficient portfolios will fix μ (or ν) in terms of λ, so the investor has only one 'degree of freedom' in this case. If he is constrained only to one fixed price, then he has only one efficient portfolio, that given in § 66 above.

Risk-free securities and portfolios

70. In our example, the portfolio consisting of 10 units of security 2 as an asset, after deducting the fixed liability, gives an ultimate surplus, at our time horizon of three years, of exactly 900, with total certainty: $V=0$. Wise, in his written reply to the discussion, correctly points out that it would be ridiculous to cover a liability that, by normal actuarial discounting methods of valuation, cannot be more than 300 in present value, by such 'overkill'. But, just as many investment intermediaries may hold negative quantities of assets by creating corresponding liabilities, so such an intermediary may also create new securities and sell them on the market.

71. Thus, if the investor creates a third security, represented by:

$$\text{security 3: } (0, 0, 100),$$

then the portfolio that consists of 10 units of security 2 minus 9 units of security 3 will have the representation:

$$(100, 100, 100),$$

and will exactly match the specified liability, both in the sense that the asset-income in each year will exactly meet the liability-outgo, and in the sense that the ultimate surplus will have zero mean and zero variance. If the market price of security 3 is P_3, then the price of this exactly matching portfolio will be:

$$10P_2-9P_3,$$

and this gives an exact present market value for the liability.

72. At the time horizon, one unit of security 3 provides £100 with certainty: $E_3=100$ and $V_3=0$. We can describe security 3 as a risk-free security in respect of this time-horizon. We could also describe it as a fixed interest zero-coupon bond, or a 'bullet bond' in U.S. terminology, but the concept of a risk-free security extends beyond fixed interest bonds to any security that has zero variance,

measured either in fixed money or in index-linked real terms or in any other terms. Just as the new security 3 is a risk-free security, so the portfolio consisting of 10 units of security 2 minus the liability is a risk-free portfolio, and so would any portfolio that produced zero variance at the specified time-horizon. If the corresponding risk-free security has a known market value, so likewise does the risk-free portfolio, and this may determine the market price of the liabilities even though they are unmarketable by themselves.

Absolute and risk-free matching

73. There is no disagreement between Wise and me about the concept of absolute matching. However, I think it may be useful to extend the term 'absolute match' for a given liability to cover any feasible portfolio that gives zero expected surplus and zero variance of surplus, whatever happens on the way there. Similarly, the 'risk-free match' can be defined as any feasible portfolio that gives zero variance of surplus, even if the expected surplus is non-zero. In both cases, the set of feasible portfolios may be subject to whatever constraints are imposed by the circumstances.

Realistic prices

74. In § 19 above I chose market prices of $P_1 = 400$, $P_2 = 100$, and observed that these were hardly realistic. More reasonable prices would be, for example, $P_1 = 93$, $P_2 = 95$. At these prices the redemption yields are 9·21% and 8·92%, respectively, so that security 1 is now a little 'cheap' and security 2 a little 'dear'. However, the expected returns per unit of price are now almost equal, respectively $E_1/P_1 = 1{\cdot}2998$ and $E_2/P_2 = 1{\cdot}2924$, so the surface of feasible portfolios has nearly 'folded up flat' and the ellipses of constant V in the P–E plane are extremely thin. The region of positive portfolios in Figure 16 is almost indistinguishable from a straight line, and each ellipse has a major axis which is 3,083 times the length of the minor axis, so in the Figure they are almost indistinguishable from line segments. Wise's unbiased match is at the 'southern' extremity of such an ellipse, but that is practically indistinguishable from the 'western' extremity, which is truly an efficient portfolio. Similarly, the region of positive portfolios is an extremely thin wedge, indistinguishable from a straight line in the Figure. It was for these reasons that I started with unrealistic prices, to show the principles more distinctly.

Generalization to n *securities*

75. I shall now turn to the general case in which there may be many securities. I shall first consider the unconstrained case, with n securities and one liability (subscripted $n+1$). From §§ 11, 15, 19 and 20 above we see that, to find the optimum portfolio for a given $E = E^*$ and $P = P^*$, we need to minimize:

$$V = \mathbf{x}'\mathbf{V}\mathbf{x}$$

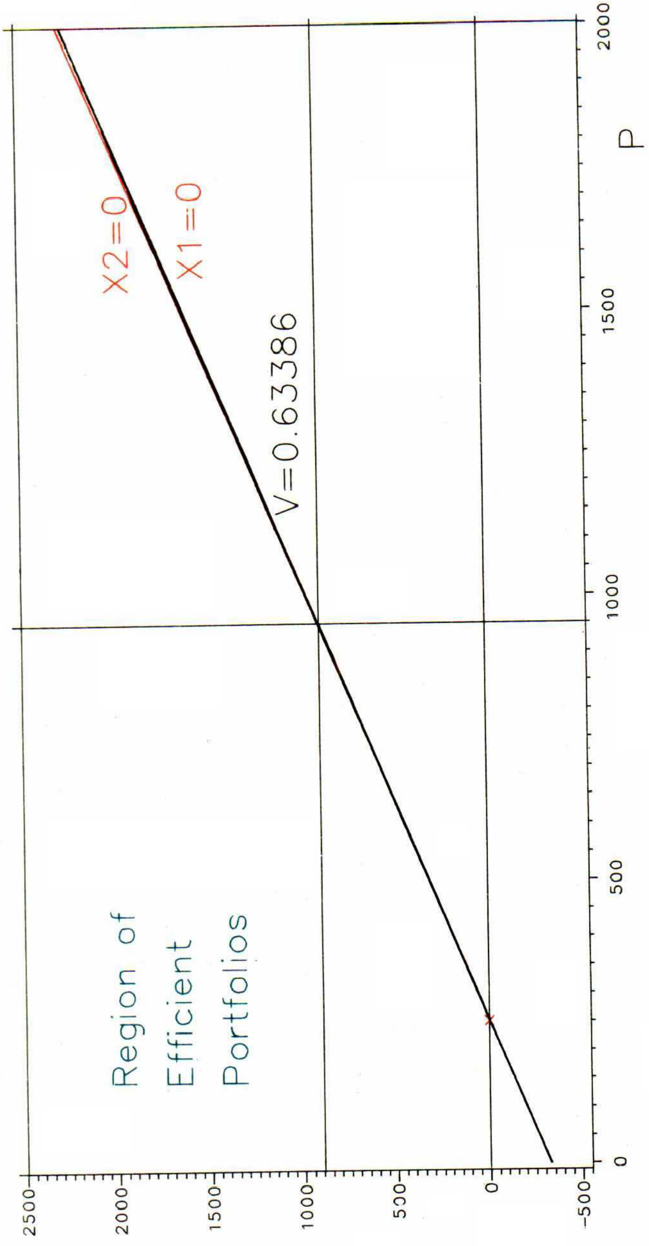

Figure 16
Expected (E) as a function of Price (P)
for fixed Variance 0.63386
X=Wise's Unbiased Match
Data as Wise with P1=93 P2=95

subject to

$$E = \mathbf{x}'\mathbf{e} = \mathbf{e}'\mathbf{x} = E^*$$
$$P = \mathbf{x}'\mathbf{p} = \mathbf{p}'\mathbf{x} = P^*$$

and $x_{n+1} = x_L = -1$, which I shall represent by

$$Q = \mathbf{x}'\mathbf{q} = \mathbf{q}'\mathbf{x} = 1$$

where $\mathbf{q}' = (0, 0, \ldots, 0, -1)$.

76. We can find the value of $\mathbf{x}$ that minimises V by setting up the Lagrangian

$$\begin{aligned} L &= V - 2g_1(E - E^*) - 2g_2(P - P^*) - 2g_3(Q - 1), \\ &= \mathbf{x}'\mathbf{V}\mathbf{x} - 2g_1(\mathbf{e}'\mathbf{x} - E^*) - 2g_2(\mathbf{p}'\mathbf{x} - P^*) - 2g_3(\mathbf{q}'\mathbf{x} - 1), \end{aligned} \qquad (76.1)$$

differentiating L with respect to each of the x_i ($i = 1, \ldots, n+1$) and also g_1, g_2 and g_3, setting each of the partial derivatives equal to zero, and solving for $\mathbf{x}$, g_1, g_2 and g_3, viz:

$$\frac{\partial L}{\partial \mathbf{x}} = 2\mathbf{x}'\mathbf{V} - 2g_1\mathbf{e}' - 2g_2\mathbf{p}' - 2g_3\mathbf{q}' = 0 \qquad (76.2)$$

$$\frac{\partial L}{\partial g_1} = -\mathbf{e}'\mathbf{x} + E^* = 0 \qquad (76.3)$$

$$\frac{\partial L}{\partial g_2} = -\mathbf{p}'\mathbf{x} + P^* = 0 \qquad (76.4)$$

$$\frac{\partial L}{\partial g_3} = -\mathbf{q}'\mathbf{x} + 1 = 0 \qquad (76.5)$$

Provided $\mathbf{V}$ is not singular, so that $\mathbf{V}^{-1}$ exists, we can postmultiply (76.2) by $\mathbf{V}^{-1}$ and arrange to get:

$$\mathbf{x}' = g_1\mathbf{e}'\mathbf{V}^{-1} + g_2\mathbf{p}'\mathbf{V}^{-1} + g_3\mathbf{q}'\mathbf{V}^{-1} \qquad (76.6)$$

Postmultiplying (76.6) by $\mathbf{e}$ and substituting in (76.3) we get:

$$\mathbf{x}'\mathbf{e} = \mathbf{e}'\mathbf{x} = g_1\mathbf{e}'\mathbf{V}^{-1}\mathbf{e} + g_2\mathbf{p}'\mathbf{V}^{-1}\mathbf{e} + g_3\mathbf{q}'\mathbf{V}^{-1}\mathbf{e} = E^* \qquad (76.7)$$

Postmultiplying (76.6) by $\mathbf{p}$ and substituting in (76.4) we get:

$$\mathbf{x}'\mathbf{p} = \mathbf{p}'\mathbf{x} = g_1\mathbf{e}'\mathbf{V}^{-1}\mathbf{p} + g_2\mathbf{p}'\mathbf{V}^{-1}\mathbf{p} + g_3\mathbf{q}'\mathbf{V}^{-1}\mathbf{p} = P^* \qquad (76.8)$$

Postmultiplying (76.6) by $\mathbf{q}$ and substituting in (76.5) we get:

$$\mathbf{x}'\mathbf{q} = \mathbf{q}'\mathbf{x} = g_1\mathbf{e}'\mathbf{V}^{-1}\mathbf{q} + g_2\mathbf{p}'\mathbf{V}^{-1}\mathbf{q} + g_3\mathbf{q}'\mathbf{V}^{-1}\mathbf{q} = 1 \qquad (76.9)$$

77. Equations (76.7), (76.8) and (76.9) are of the form:

$$\begin{aligned} \alpha g_1 + \zeta g_2 + \varepsilon g_3 &= E \\ \zeta g_1 + \beta g_2 + \delta g_3 &= P \\ \varepsilon g_1 + \delta g_2 + \gamma g_3 &= 1, \end{aligned}$$

where we have dropped the * on E and P, to denote general values of E and P, and we have put

$$\begin{aligned} \mathbf{e}'\mathbf{V}^{-1}\mathbf{e} &= \alpha \\ \mathbf{p}'\mathbf{V}^{-1}\mathbf{p} &= \beta \\ \mathbf{q}'\mathbf{V}^{-1}\mathbf{q} &= \gamma \\ \mathbf{p}'\mathbf{V}^{-1}\mathbf{e} = \mathbf{e}'\mathbf{V}^{-1}\mathbf{p} &= \zeta \\ \mathbf{q}'\mathbf{V}^{-1}\mathbf{e} = \mathbf{e}'\mathbf{V}^{-1}\mathbf{q} &= \varepsilon \\ \mathbf{p}'\mathbf{V}^{-1}\mathbf{q} = \mathbf{q}'\mathbf{V}^{-1}\mathbf{p} &= \delta \end{aligned}$$

Provided that the determinant

$$\Delta = \begin{vmatrix} \alpha & \zeta & \varepsilon \\ \zeta & \beta & \delta \\ \varepsilon & \delta & \gamma \end{vmatrix} = \alpha\beta\gamma + 2\delta\varepsilon\zeta - \alpha\delta^2 - \beta\varepsilon^2 - \gamma\zeta^2 \neq 0,$$

we get the solution

$$g_1 = \frac{P(\delta\varepsilon - \gamma\zeta) + E(\beta\gamma - \delta^2) + (\delta\zeta - \beta\varepsilon)}{\Delta}$$

$$g_2 = \frac{P(\alpha\gamma - \varepsilon^2) + E(\delta\varepsilon - \gamma\zeta) + (\varepsilon\zeta - \alpha\delta)}{\Delta}$$

$$g_3 = \frac{P(\varepsilon\zeta - \alpha\delta) + E(\delta\zeta - \beta\varepsilon) + (\alpha\beta - \zeta^2)}{\Delta}$$

whence

$$\begin{aligned} \mathbf{x} = \{&[P(\delta\varepsilon - \gamma\zeta) + E(\beta\gamma - \delta^2) + (\delta\zeta - \beta\varepsilon)]\mathbf{V}^{-1}\mathbf{e} \\ &+ [P(\alpha\gamma - \varepsilon\gamma^2) + E(\delta\varepsilon - \gamma\zeta) + (\varepsilon\zeta - \alpha\delta)]\mathbf{V}^{-1}\mathbf{p} \\ &+ [P(\varepsilon\zeta - \alpha\delta) + E(\delta\zeta - \beta\varepsilon) + (\alpha\beta - \zeta^2)]\mathbf{V}^{-1}\mathbf{q}\}/\Delta, \end{aligned}$$

and $V = \mathbf{x}'\mathbf{V}\mathbf{x}$

$$\begin{aligned} = \{&P^2(\alpha\gamma - \varepsilon^2) + 2PE(\delta\varepsilon - \gamma\zeta) + E^2(\beta\gamma - \delta^2) \\ &+ 2P(\varepsilon\zeta - \alpha\delta) + 2E(\delta\zeta - \beta\varepsilon) + (\alpha\beta - \zeta^2)\}/\Delta, \end{aligned}$$

which gives V as a quadratic function of P and E. (The above development follows that given by Szegö (1980), Chapter 2, for the ordinary portfolio selection case).

78. Thus in the general case the boundary of the region of feasible portfolios that gives the lowest V for any chosen P and E is again a quadric surface, an elliptic paraboloid. In the general case it is possible to reach a feasible portfolio with any higher value of V, which was not possible in the earlier example with only two securities, but any such feasible portfolio is dominated by the portfolio with the same P and E on the surface itself. Not all the boundary portfolios are efficient. Just as in the earlier example, only those that lie in the 'northwest quadrant' of this surface are efficient. The general features of the region of efficient portfolios, the choice of optimal portfolio for an investor, the use of the $k\sigma$-solvency region, and the constraints on price follow just as in the earlier discussion.

Constraints

79. Just as before, we may wish to discover the positive region and the region of positive efficient portfolios, subject to the constraints $x_i \geqslant 0$, for $i = 1, \ldots, n$ (necessarily $x_{n+1} = -1$). This takes us into a more complicated problem of quadratic programming which I cannot discuss here. Wise (1984a) gives an algorithm for solving his particular case. I do not know whether it will also solve my more general one. The general outline of the solution is, however, apparent, though this development starts from the other end from Wise's. Consider first the surface of unconstrained minimum variance portfolios. Assume that some part of this surface contains portfolios that satisfy the positive constraints. This part will be bounded along some edge by portfolios in which some x_i, say x_n, $= 0$. We now omit security n, and discover the surface of minimum variance portfolios for all other securities. Inside the region already discovered as positive, this will necessarily have higher variances than those portfolios including security n; along the edge where x_n equalled 0 they will coincide; outside the region already discovered this new surface will contain some portfolios with $x_i \geqslant 0$, $i = 1, \ldots, n-1$. But as we go further away we shall find another edge where another x_i, say x_{n-1}, $= 0$. We omit security $n-1$ and continue. If we are unlucky to start with, so that no part of the region of unconstrained portfolios meets the positive constraints, we need to omit each security in turn to find some region that contains positive portfolios; or omit each pair of securities in turn, etc. Eventually the positive minimum variance portfolio, if any, for every combination of P and E can be discovered. Computer methods for solving such quadratic programming problems are well developed.

Conclusion

80. This note is long enough already without my attempting to tackle any more of Wise's examples. I hope the detailed consideration I have given to his one elementary example will show the line of future development, particularly in the more general cases. As already noted, full discussion of the mathematics of the portfolio selection problem with fixed price and with no liabilities is contained in Sharpe (1970) and in Szegö (1980), and what is needed is more generalization of their basic approaches. Although our elementary example uses fixed money liabilities and fixed interest securities, Wise moves on to examples where liabilities and securities are dependent on other random variables, such as price inflation or earnings increases. My own development is also applicable to these, indeed particularly so. The fixed interest case becomes a special case, with the overall minimum $V = 0$ (and $\mathbf{V}$ singular), whenever the number of available independent securities equals the number of independent years, so that 'overkill' is always possible. A possible objection to my approach is that when the prices of assets change so also does the whole structure; but unless the prices change greatly, the portfolio that was previously efficient is not likely to be far removed from efficiency. This problem requires empirical investigation.

ACKNOWLEDGEMENTS

81. Much of the work for this note was carried out while I was on a visit to the University of Waterloo, Ontario, in July and August 1984, sponsored by the Canadian Life and Health Insurance Association, to whom I am very grateful. I am grateful also for the lively discussions I had then with Phelim Boyle, Harry Panjer, Allan Brender, Keith Sharp and others, and also with Andrew Wise at various times since his own paper was presented. Finally, I should like to thank Mrs Eleanor Runciman for preparing a faultless typescript of a very difficult paper. Readers may also like to note that the Figures in the paper were produced using the graph plotting package of a statistical computer system called SAS.

REFERENCES

MOORE, P. G. (1972) Mathematical models in portfolio selection. *J.I.A.* **98**, 103.

SHARPE, W. F. (1970) *Portfolio Theory and Capital Markets*. McGraw-Hill, New York.

SZEGÖ, G. P. (1980) *Portfolio Theory with Application to Bank Asset Management*. Academic Press, New York.

WISE, A. J. (1984a) A Theoretical Analysis of the Matching of Assets to Liabilities. *J.I.A.* **111**, 375.

WISE, A. J. (1984b) The Matching of Assets to Liabilities. *J.I.A.* **111**, 445.

APPENDIX

A1. Older actuaries may not remember, and younger ones may never have learnt, much about analytical geometry. This appendix therefore states, without proof, a few relevant facts about the conic sections used in the note.

A2. The quadratic equation:

$$F(x,y)=ax^2+2hxy+by^2+2gx+2fy+c=0$$

represents the general conic section in the x–y plane. The three main forms this may take are the ellipse, the hyperbola and the parabola. There are also a number of special cases.

A3. The main special case is when $F(x,y)$ factorizes into two factors linear in x and y. This is so when the determinant

$$D=\begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix}=abc+2fgh-af^2-bg^2-ch^2=0.$$

In this case, the equation represents two straight lines in the x–y plane; these, however, may be not real lines, but imaginary ones, depending on the values of the coefficients, a, b, c, etc.

A4. If the determinant is not zero, the form the conic section takes depends on the sign of the discriminant

$$\begin{vmatrix} a & h \\ h & b \end{vmatrix}=ab-h^2.$$

If the discriminant is positive the equation represents an ellipse, if negative a hyperbola, and if zero is parabola.

A5. The archetypal ellipse is given by

$$\frac{x^2}{p^2}+\frac{y^2}{q^2}=1,$$

which represents an ellipse with its centre at the origin, its axes along the x and y axes, and with half-axes of length p and q. If $p=q$ we get

$$x^2+y^2=p^2,$$

the equation of a circle of radius p.

A6. The archetypal hyperbola is given by

$$\frac{x^2}{p^2}-\frac{y^2}{q^2}=1,$$

which represents a hyperbola with its centre at the origin, asymptotes given by the lines

$$\frac{x}{p} = \pm \frac{y}{q},$$

and with the 'nose' or vertex of each arm at a distance p from the centre.

A7. The archetypal parabola is given by the familiar

$$y = px^2,$$

which has its vertex at the origin, and its axis along the positive y axis. Alternatively:

$$x = qy^2,$$

has its vertex at the origin and its axis along the positive x axis.

A8. The general form of the ellipse can be transformed into the archetypal form first by changing the coordinate origin to the centre of the ellipse, then by rotating the coordinate axes. The centre of the general ellipse is given by:

$$x_C = \frac{hf - bg}{ab - h^2}$$

$$y_C = \frac{hg - af}{ab - h^2}$$

and the equation of the ellipse relative to that centre is given by

$$ax^2 + 2hxy + by^2 + c^* = 0$$

where $c^* = F(x_C, y_C)$.

A9. If the angle that the axis of the ellipse that lies in the first quadrant makes with the x axis is designated θ, we can derive θ from

$$\tan 2\theta = \frac{2h}{a - b}.$$

If $h = 0$, the ellipse is already 'square on' to the axes. If $a = b$, the axes are at 45° to the coordinate axes. If $h = 0$ and $a = b$ we have a circle.

A10. If $\tan 2\theta$ is positive, θ is less than 45°, and the equations of the axes of the ellipse, relative to the centre, are:

$$y = x\sqrt{\frac{S^2 + (a-b)S - 2h^2}{2h^2}},$$

and

$$y = x\Big/\sqrt{\frac{S^2 - (a-b)S - 2h^2}{2h^2}},$$

where

$$S = \sqrt{(a-b)^2 + 4h^2}.$$

A11. If $\tan 2\theta$ is negative, θ is greater than 45°, and the equations of the axes of the ellipse, relative to the centre, are

$$y = x\sqrt{\frac{S^2-(a-b)S-2h^2}{2h^2}},$$

and

$$y = -x/\sqrt{\frac{S^2+(a-b)S-2h^2}{2h^2}},$$

where S is defined as above.

A12. The lengths of the half-axes are:

$$\sqrt{\frac{2c^*}{(a+b)-S}} \quad \text{and} \quad \sqrt{\frac{2c^*}{(a+b)+S}}$$

and their ratio is

$$\sqrt{\frac{(a+b)+S}{(a+b)-S}}.$$

A13. A straight line through the centre of the ellipse, a diameter, divides it symmetrically, meeting it at two opposite points. The tangents to the ellipse at these points are parallel. If these tangents have slope m, the equation of the diameter, relative to the original coordinate axes, is:

$$y(h+bm)+x(a+hm)+(g+fm)=0.$$

and the points on the ellipse at opposite ends of this diameter (x_A,y_A) and (x_B,y_B) are given by:

$$x_A = x_C - \frac{(hm+b)\sqrt{-D/(a^2m^2+2hm+b)}}{(ab-h^2)}$$

$$y_A = y_C + \frac{(am+h)\sqrt{-D/(a^2m^2+2hm+b)}}{(ab-h^2)}$$

and

$$x_B = x_C + \frac{(hm+b)\sqrt{-D/(a^2m^2+2hm+b)}}{(ab-h^2)}$$

$$y_B = y_C - \frac{(am+h)\sqrt{-D/(a^2m^2+2hm+b)}}{(ab-h^2)}$$

where D is the determinant defined in A3 above.

A14. If $m=0$, the line joins the 'northern' and 'southern' extremities of the ellipse, i.e. those points with maximum and minimum values of y, and is:

$$hy+ax+g=0.$$

If m is infinite, the line joins the 'western' and 'eastern' extremities of the ellipse, i.e. those points with minimum and maximum values of x, and is

$$by+hx+f=0.$$

A15. The 'northern' and 'southern' extremities are given by:

$$x_N = x_C - h\sqrt{-aD}\Big/a(ab-h^2) \qquad x_S = x_C + h\sqrt{-aD}\Big/a(ab-h^2)$$

$$y_N = y_C + \sqrt{-aD}\Big/(ab-h^2) \qquad y_S = y_C - \sqrt{-aD}\Big/(ab-h^2)$$

and the 'eastern' and 'western' extremities are given by:

$$x_E = x_C + \sqrt{-bD}\Big/(ab-h^2) \qquad x_W = x_C - \sqrt{-bD}\Big/(ab-h^2)$$

$$y_E = y_C - h\sqrt{-bD}\Big/b(ab-h^2) \qquad y_W = y_C + h\sqrt{-bD}\Big/b(ab-h^2)$$

These expressions assume that a and b, which must be of the same sign if $ab-h^2$ is to be positive, are both positive; if not, the whole equation of the ellipse can be multiplied by -1.

A16. We can thus see that the proportions and the alignment of the ellipse are determined by a, b and h; that given these, the position of the centre is determined by f and g; and that given these, the size of the ellipse is determined by c. Since the whole equation can be multiplied or divided by a constant and still represent the same figure, one of the coefficients can be chosen arbitrarily.

A17. Similar results to the above can be derived for the hyperbola and the parabola, but since they are not needed in the development in the note they have been omitted.

A18. A quadratic equation in x, y and z defines a quadric surface in x–y–z space. The only general quadric surfaces discussed in the note are of the form

$$z = ax^2 + 2hxy + by^2 + 2gx + 2fy + c,$$

an elliptic paraboloid, and

$$z^2 = ax^2 + 2hxy + by^2 + 2gx + 2fy + c,$$

an elliptic hyperboloid of two sheets. In both cases, we assume that $(ab - h^2) > 0$, so that the right hand side represents an ellipse, and that the right hand side is greater than or equal to zero for all values of x and y, so that the surface does not cross the plane $z = 0$.

A19. It is easy to see that the 'contour lines' of these surfaces, projected onto the plane $z = 0$, form a series of concentric ellipses of the same alignment and proportions, with sizes determined by $(c - z)$ and $(c - z^2)$ respectively.

A20. A cross-section of either surface in the plane of a constant y gives a curve in the x–z plane of the forms respectively:

$$z = ax^2 + 2(hy + g)x + (by^2 + 2fy + c),$$

a parabola, or

$$z^2 = ax^2 + 2(hy + g)x + (by^2 + 2fy + c) = 0,$$

a hyperbola. Both curves are 'square on' to the x-axis, with 'noses' pointing to the z-axis. The hyperbola is symmetric about the z-axis, with centre

$$x_C = -(hy + g)/a,$$
$$z_C = 0.$$

The same expressions give the point on the z-axis which the nose of the parabola points to or touches.

A21. If each surface is projected onto the plane $y = 0$, its profile is given by a parabola or hyperbola, respectively:

$$z = \frac{(ab - h^2)x^2 + 2(bg - hf)x + (bc - f^2)}{b}$$

and

$$z^2 = \frac{(ab - h^2)x^2 + 2(bg - hf)x + (bc - f^2)}{b}$$

Projected onto the plane $x = 0$, the profiles are given correspondingly by:

$$z \text{ or } z^2 = \frac{(ab - h^2)y^2 + 2(af - hg)y + (ac - g^2)}{a}.$$

A22. The noses of the quadric surfaces are given by

$$x_C = \frac{hf - bg}{ab - h^2}$$

$$y_C = \frac{hg - af}{ab - h^2}$$

$$z_c \text{ or } z_C^2 = \frac{abc + 2fgh - af^2 - bg^2 - ch^2}{ab - h^2}$$

A23. Certain partial derivatives for the elliptic paraboloid are:

$$\frac{\partial y}{\partial x} = -\frac{ax + hy + g}{hx + by + f}, \text{ i.e. keeping } z \text{ constant,}$$

$$\frac{\partial z}{\partial x} = 2(ax + hy + g), \text{ i.e. keeping } y \text{ constant,}$$

and

$$\frac{\partial z}{\partial y} = 2(hx + by + f), \text{ i.e. keeping } x \text{ constant,}$$

For the elliptic hyperboloid the first of these is the same, and the others are:

$$\frac{\partial z}{\partial x} = (ax + hy + g)/z,$$

and

$$\frac{\partial z}{\partial y} = (hx + by + f)/z.$$