Testing for mean-variance spanning

Nijman, T.E.; de Roon, F.A.

Published in:
Journal of Empirical Finance

Publication date:
2001

Document Version
Peer reviewed version

Citation for published version (APA):
Nijman, T. E., & de Roon, F. A. (2001). Testing for mean-variance spanning: A survey. *Journal of Empirical Finance*, *8*(2), 111-156.

Testing for Mean-Variance Spanning: A Survey

Frans A. de Roon[*] Theo E. Nijman[y][z]

January 2001

Abstract

In this paper we present a survey on the various approaches that can be used to test whether the mean-variance frontier of a set of assets spans or intersects the frontier of a larger set of assets. We analyze the restrictions on the return distribution that are needed to have mean-variance spanning or intersection. The paper explores the duality between mean-variance frontiers and volatility bounds, analyzes regression based test procedures for spanning and intersection, and shows how these regression based tests are related to tests for mean-variance efficiency, performance measurement, optimal portfolio choice, and specification error bounds.

JEL: G110, G120

Keywords: mean-variance spanning, portfolio choice, volatility bounds, performance measurement

[*]CentER for Economic Research and Department of Finance, Tilburg University, P.O.Box 90153, 5000 LE Tilburg, The Netherlands, and CEPR. E-mail: F.A.deRoon@kub.nl

[y]CentER for Economic Research and Department of Econometrics, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: Nyman@kub.nl

[z]Geert Bekaert, Ton Vorst, Bas Werker, and two anonymous referees have provided many helpful comments and suggestions.

1 Introduction

In recent years the finance literature has witnessed an increasing use of tests for mean-variance spanning and intersection, as introduced by Huberman and Kandel (1987). In this paper we will provide a survey of the literature on testing for mean-variance spanning and intersection, as well as of its relationships with volatility bounds, tests for mean-variance efficiency, performance evaluation, and the specification error bounds that have recently been proposed by Hansen and Jagannathan (1997). There exists a vast literature on most of these subjects and the intention here is not to give a complete overview, but merely to illustrate that the concept of mean-variance spanning and intersection provides a framework in which many other results can be understood.

The literature on mean-variance spanning and intersection analyzes the effect that the introduction of additional assets has on the mean-variance frontier. If the mean-variance frontier of the benchmark assets and the frontier of the benchmark plus the new assets have exactly one point in common, this is known as intersection. This means that there is one mean-variance utility function for which there is no benefit from adding the new assets. If the mean-variance frontier of the benchmark assets plus the new assets coincides with the frontier of the benchmark assets only, there is spanning. In this case no mean-variance investor can benefit from adding the new assets to his (optimal) portfolio of the benchmark assets only. For instance, DeSantis (1995) and Cumby and Glen (1990) consider the question whether US-investors can benefit from international diversification. Taking the viewpoint of a US-investor who initially only invests in the US, these authors study the question whether they can enhance the mean-variance characteristics of their portfolio by also investing in other (developed) markets. Similarly, taking the perspective of a US-investor who invests in the US and (possibly) in other developed markets such as Japan and Europe, DeSantis (1994), Bekaert and Urias (1996), Errunza, Hogan and Hung (1998), and DeRoon, Nijman and Werker (2001) e.g., investigate whether the investors can improve upon their mean-variance portfolio by investing in emerging markets. As a final example, Glen and Jorion (1993) investigate whether mean-variance investors with a well-diversified international portfolio of stocks and bonds should add currency futures to their portfolio, i.e., whether or not they should hedge the currency risk that arises from their positions in stocks and bonds.

As shown by DeSantis (1994), Ferson, Foerster, and Keim (1993), Ferson

(1995) and Bekaert and Urias (1996), the hypothesis of mean-variance spanning and intersection can be reformulated in terms of the volatility bounds introduced by Hansen and Jagannathan (1991). In that case, the interest is in the question whether a set of additional assets contains information about the volatility of the pricing kernel or the stochastic discount factor that is not already present in the initial set of assets considered by the econometrician. For instance, in the case of emerging markets, the question is whether considering returns from the US-market together with returns from emerging markets produces tighter volatility bounds on the stochastic discount factor than returns from the US-market only.

The duality between mean-variance frontiers and volatility bounds for the stochastic discount factors will be the subject of Section 2. The analysis provided in that section will allow us to study mean-variance spanning and intersection, both in terms of mean-variance frontiers and in terms of volatility bounds. The concept of mean-variance spanning and intersection will formally be introduced in Section 3. In that section it will be also be shown how simple regression techniques can be used to test for mean-variance spanning and intersection. In Section 4 we will consider how conditioning information can be incorporated in the test procedures. In Section 5 we will show how deviations from mean-variance intersection and spanning can be interpreted in terms of performance measures like Jensen's alpha and the Sharpe ratio, and how the regression tests for intersection can be used to derive the new optimal portfolio weights. In Section 6 we provide a brief discussion of the speci…cation error bound introduced by Hansen and Jagannathan (1997) and how this is related to mean-variance intersection. As with the performance measures in Section 5, speci…cation error bounds are especially of interest when there is no intersection. This paper will end with a summary.

2 Volatility bounds and the duality with mean-variance frontiers

The purpose of this section is to introduce volatility bounds and mean-variance frontiers and to show the duality between these two frontiers. Because mean-variance spanning and intersection can be de…ned from volatility bounds as well as from mean-variance frontiers, this section provides a basis for the analysis of mean-variance spanning and intersection in the remainder of the paper.

2.1 Volatility bounds

Suppose an investor chooses his portfolio from a set of K assets, with current prices given by the K-dimensional vector P_t and with payoffs in the next period given by the vector P_{t+1} (including dividends and the like). Returns $R_{i,t+1}$ are payoffs with prices equal to one. Assuming there are no market frictions such as short sales constraints and transaction costs and assuming that the law of one price holds, there exists a stochastic discount factor or pricing kernel, M_{t+1}, such that[1]

$$E[M_{t+1}R_{t+1} \mid I_t] = \iota_K, \tag{1}$$

where ι_K is a K-dimensional vector containing ones, and I_t is the information set that is known to the investor at time t. In the sequel we will use $E_t[.]$ as shorthand notation for $E[. \mid I_t]$.

Apart from the law of one price, an alternative way to motivate (1) is to look at the discrete time consumption and portfolio problem that an investor solves:

$$\begin{aligned} & \max_{\{w_t, C_t\}} E_t\Big[\sum_{j=0}^{\infty} \rho^j U(C_{t+j})\Big], \qquad (2) \\ \text{s.t. } W_{t+1} &= w_t' R_{t+1}(W_t - C_t), \\ w_t'\iota_K &= 1, \ \forall t \end{aligned}$$

where C_t is consumption at time t, W_t is the wealth owned by the investor at time t, ρ is the subjective discount factor of the investor, and w_t is the K-dimensional vector of portfolio weights that the investor chooses. The function $U(C_t, C_{t+1}, \ldots) = \sum_{j=0}^{\infty} \rho^j U(C_{t+j})$ is a strictly increasing and concave time-separable utility function. The first order conditions of problem (2) imply that

$$M_{t+1} = \rho \frac{U'(C_{t+1})}{U'(C_t)} \Big|_{C_t^{opt}, w_t^{opt}},$$

is a valid stochastic discount factor with $U'(.)$ being the first derivative of U. Thus, one way to think about the stochastic discount factor or pricing

[1] Replacing the law of one price with the stronger condition that there are no arbitrage opportunities, we would also have that $M_{t+1} > 0$.

kernel is as the intertemporal marginal rate of substitution (IMRS). This interpretation of the pricing kernel is more restrictive than the law of one price though, since it also implies that $M_{t+1} > 0$.

In many of the problems we consider in this paper, it is convenient to look at a more simple portfolio problem. Usually we will restrict ourselves to one-period portfolio problems, where the agent maximizes his indirect utility of wealth function (see, e.g., Ingersoll (1987), p.66):

$$
\begin{aligned}
& \max_{\{w\}} E_t[u(W_{t+1})]; \\
\text{s.t. } W_{t+1} &= W_t\, w' R_{t+1}; \\
w'\iota_K &= 1.
\end{aligned}
$$

In this case a valid stochastic discount factor is $W_t \cdot u'(W_{t+1})/\lambda$, with $u'(.)$ being the first derivative of the indirect utility function evaluated at the optimal portfolio choice, and λ the Lagrange multiplier for the restriction that $w'\iota_K = 1$.

The expectation of the stochastic discount factor will be denoted by v_t, i.e., $v_t \equiv E_t[M_{t+1}]$. The name stochastic discount factor refers to the fact that M_{t+1} discounts payoffs differently in different states of the world. To illustrate this, using the definition of covariance, (1) can be rewritten as

$$
\iota_K = E_t[M_{t+1}R_{t+1}] = v_t E_t[R_{t+1}] + Cov_t[R_{t+1}, M_{t+1}]. \tag{3}
$$

The first term in (3) uses v_t to discount the expected future payoffs, while the second term is a risk adjustment (recall that ι_K is the price-vector of the returns R_{t+1}). Accordingly, risk premia are determined by the covariance of asset payoffs with M_{t+1}. If one of the assets is a risk free asset with return R_t^f, then it follows from the conditional expectation in (1) that $R_t^f = 1/v_t$. In the sequel we will usually not impose the presence of such a risk free asset. If a risk free asset is available however, then we can always substitute $1/R_t^f$ for v_t.

Equation (1) is the starting point for most asset pricing models. In fact, differences in asset pricing models can be interpreted as differences in the function that each model assigns to M_{t+1} (see, e.g., Cochrane (1997)). Since each valid stochastic discount factor has to satisfy (1), observed asset returns can be used to derive information about these discount factors. For instance, following Hansen and Jagannathan (1991) it is possible to derive a lower bound on the variance of M_{t+1}, that each valid stochastic discount factor has to satisfy, which is known as the volatility bound. To see this, we start from the unconditional version of (1), and leave out the time subscripts for the expectations and (co)variance operators, as well as for v. In this paper, the expectation of the stochastic discount factor will usually be a free parameter. We will denote all discount factors that satisfy (1) and that have unconditional expectation v with $M(v)_{t+1}$, and derive a lower bound for the variance of each $M(v)_{t+1}$.

Let the unconditional expectation and covariance matrix of the returns R_{t+1} be given by μ_R and Σ_{RR} respectively, and assume that all returns are independently and identically distributed (i.i.d.), so that the expectations and covariances do not vary over time. This assumption will be relaxed in Section 4 of this paper. Given the set of asset returns R_{t+1}, let $m_R(v)_{t+1}$ be a candidate stochastic discount factor that has expectation v and that is linear in the asset returns:

$$m_R(v)_{t+1} = v + \varphi(v)'(R_{t+1} - \mu_R); \tag{4}$$

where we write $\varphi(v)$ to indicate that these coefficients are a function of the expectation of $M(v)_{t+1}$. Substituting (4) into (1), we obtain:

$$\varphi(v) = \Sigma_{RR}^{-1}(\iota_K - v\mu_R): \tag{5}$$

Since both $M(v)_{t+1}$ and $m_R(v)_{t+1}$ satisfy (1) we have that $E[(M(v)_{t+1} - m_R(v)_{t+1})R_{t+1}] = 0$, so the difference between any $M(v)_{t+1}$ that satisfies (1) and $m_R(v)_{t+1}$ is orthogonal to R_{t+1} and therefore to $m_R(v)_{t+1}$ itself. This implies for the variance of $M(v)_{t+1}$ that:

$$\begin{aligned} Var[M(v)_{t+1}] &= Var[m_R(v)_{t+1}] + Var[(M(v)_{t+1} - m_R(v)_{t+1})] \qquad (6) \\ &\geq Var[m_R(v)_{t+1}]; \end{aligned}$$

which shows that $m_R(v)_{t+1}$ has the lowest variance of all valid stochastic discount factors $M(v)_{t+1}$. This minimum variance can be obtained by combining (4) and (5):

$$Var[m_R(v)_{t+1}] = (\iota_K - v\mu_R)'\Sigma_{RR}^{-1}(\iota_K - v\mu_R): \tag{7}$$

Thus, any pricing model that aims to price the assets R_{t+1} correctly, has to yield a pricing kernel that, for a given v, has a variance at least as large as (7). Equivalently, if we know that agents choose their optimal portfolio from the assets that are in R_{t+1}, then (7) gives the minimum amount of variation of their IMRS that is needed to be consistent with the distribution of asset returns. Luttmer (1996) extends this kind of analysis taking into account market frictions such as short sales constraints and transaction costs. For the frictionless markets setting, Snow (1991) provides a similar analysis to derive bounds on other moments of the discount factor as well, and Bansal and Lehmann (1997) provide a bound on the mean of the logarithm of the pricing kernel, using growth optimal portfolios. Balduzzi and Kallal (1997) show how additional knowledge about risk premia may lead to sharper bounds on the volatility of the discount factor and Balduzzi and Robotti (2000) use the minimum variance discount factor to estimate risk premia associated with economic risk variables. Finally, Bekaert and Liu (1999) and Ferson and Siegel (1997) study the use of conditioning information to derive optimally scaled volatility bounds.

2.2 Duality between volatility bounds and mean-variance frontiers

In the previous section we derived the minimum amount of variation in stochastic discount factors that is needed to be consistent with the distribution of asset returns. In this section we will show that there is a close correspondence between these volatility bounds and mean-variance frontiers and that stochastic discount factors that correspond to mean-variance optimizing behavior are the stochastic discount factors with the lowest volatility. Mean-variance optimizing behavior is a special case of the portfolio problem considered before, where the problem the agent faces is $\max_{\{w\}} E[u(W_{t+1})]$, and where $E[u(.)]$ is of the form $f(w'\mu_R, w'\Sigma_{RR}w)$, with f increasing in its first argument and decreasing in its second argument.

For further reference it is useful to define the efficient set variables (see, e.g., Ingersoll (1987)):

$$A \equiv \mu_K' \Sigma_{RR}^{-1} \mu_K, \quad B \equiv \iota_R' \Sigma_{RR}^{-1} \mu_K, \text{ and } C \equiv \iota_R' \Sigma_{RR}^{-1} \iota_R.$$

A mean-variance efficient portfolio w^* is the solution to the problem

$$\max_{\{w\}} L = w'\mu_R - \gamma w'\Sigma_{RR} w - \eta(w'\mu_K - 1),$$

where γ is the coefficient of risk aversion. From the first order conditions of this problem it follows that a portfolio w^* is mean-variance efficient if there exist scalars γ and η such that[2]

$$w^* = \gamma^{-1}\Sigma_{RR}^{-1}(\mu_R - \eta\iota_K). \tag{8}$$

Because of the restriction $w'\iota_K = 1$, it also follows that $\gamma = B - A\eta$, implying that each mean-variance efficient portfolio is uniquely determined when either γ or η is known, unless $\eta = B/A$. It is straightforward to show that for a given mean-variance efficient portfolio w^*, the Lagrange multiplier η equals the expected return on the zero-beta portfolio of w^*, i.e., the intercept of the line tangent to the mean-variance frontier at w^* (in mean-standard deviation space). Since B/A, is the expected return on the global minimum variance (GMV) portfolio, this is the intercept of the asymptotes of the mean-variance frontier, but there are no lines tangent to the frontier originating at this point (see, e.g., Ingersoll (1987, p.86)).

To show the duality between mean-variance frontiers and volatility bounds, take $\varphi(v)$ for a given v, and choose a mean-variance efficient portfolio such that $\eta = 1/v$. It follows from (8) and (5) that

$$w^*(v) = \frac{\Sigma_{RR}^{-1}(\mu_R - \frac{1}{v}\iota_K)}{B - \frac{1}{v}A} = \frac{\Sigma_{RR}^{-1}(\iota_K - v\mu_R)}{A - vB} = \frac{\varphi(v)}{\iota_K'\varphi(v)}, \tag{9}$$

which shows that the vector $\varphi(v)$ is proportional to a mean-variance efficient portfolio with zero-beta return equal to $1/v$. Thus, each point on the volatility bound of stochastic discount factors, i.e., $(v, \mathrm{Var}[m(v)_{t+1}])$ corresponds to a unique point on the mean-variance frontier, (μ_p^*, σ_p^*), and each coefficient vector $\varphi(v)$ corresponds to a unique $w^*(v)$. The only exception to this result is the case where $\iota_K'\varphi(v) = 0$, which is the case if $v = A/B$, or equivalently, $\eta = B/A$. As already noted, this is the case where the zero-beta return equals the expected return on the global minimum variance portfolio (see also Hansen and Jagannathan (1991)). The duality between the mean-variance frontier of R_{t+1} and the volatility bound derived from R_{t+1} can also be seen directly from (5) and (8). Comparing the coefficients $\varphi(v)$ for the minimum variance stochastic discount factor in (5) and the portfolio weights w^* in (8) for $\eta = 1/v$, it can be seen that the coefficients $\varphi(v)$ are proportional to the portfolio weights w^*, where the coefficient of proportionality is equal

[2] More precisely, these are the minimum variance portolios, i.e., the portfolios that have minimum variance for a given expected return. The mean-variance efficient portfolios, i.e., the portfolios that also have maximum expected return for a given variance, require in addition that $\gamma \geq 0$.

to ¡´=°, i.e., w¤ = (¡´=°)'(v). In Appendix A we show graphically which points on the volatility bound correspond to points on the mean-variance frontier.

Summarizing, finding stochastic discount factors that have the lowest variance of all stochastic discount factors that price a set of asset returns R_{t+1} correctly is tantamount to finding mean-variance efficient portfolios for these same assets R_{t+1}. In the remainder of this paper we will study the effects of adding new assets to the set of assets available to investors. Although most of the results will be stated in terms of mean-variance frontiers and mean-variance efficient portfolios, it should be kept in mind that there is always a dual interpretation in terms of volatility bounds.

3 Mean-variance spanning and intersection

In the previous section we considered the volatility bounds and mean-variance frontiers that can be derived from a given set of K assets with return vector R_{t+1}. Suppose now that an investor takes an additional set of N assets with return vector r_{t+1} into account in his portfolio problem. The question we are interested in is under what conditions mean-variance efficient portfolios derived from the set of returns R_{t+1} are also mean-variance efficient for the larger set of $K + N$ assets $(R_{t+1}; r_{t+1})$. This problem was addressed in the seminal paper of Huberman and Kandel (1987). If there is only one value of ° or ´ for which mean-variance investors can not improve their mean-variance efficient portfolio by including r_{t+1} in their investment set, the mean-variance frontiers of R_{t+1} and $(R_{t+1}; r_{t+1})$ have exactly one point in common, which is referred to as intersection. In this case we will say that the mean-variance frontier of R_{t+1} intersects the mean-variance frontier of $(R_{t+1}; r_{t+1})$, or simply that R_{t+1} intersects $(R_{t+1}; r_{t+1})$. If there is no mean-variance investor that can improve his mean-variance efficient portfolio by including r_{t+1} in his investment set, the mean-variance frontiers of R_{t+1} and $(R_{t+1}; r_{t+1})$ coincide, which is referred to as spanning. In this case we will say that (the mean-variance frontier of) R_{t+1} spans (the mean-variance frontier of) $(R_{t+1}; r_{t+1})$.

As suggested by the previous section, and as shown by Ferson, Foerster, and Keim (1993), DeSantis (1994), Ferson (1995) and Bekaert and Urias (1996), the concept of mean-variance spanning and intersection has a dual interpretation in terms of volatility bounds. In terms of volatility bounds mean-variance spanning means that the volatility bound derived from the returns R_{t+1} is the same as the bound derived from $(R_{t+1}; r_{t+1})$. Therefore, the minimum variance stochastic discount factors for R_{t+1}, $m_R(v)_{t+1}$, are also the minimum variance stochastic discount factors for $(R_{t+1}; r_{t+1})$, and the asset returns r_{t+1} do not provide information about the necessary volatility of stochastic discount factors that is not already present in R_{t+1}. As will be shown formally below, mean-variance intersection is equivalent to saying that the volatility bounds derived from R_{t+1} and $(R_{t+1}; r_{t+1})$ have exactly one point in common. Thus, in case of intersection there is exactly one value of v for which the minimum variance stochastic discount factor does not change, whereas for all other values of v it does.

In finite samples it will in general be the case that adding assets causes a shift in the estimated mean-variance frontier and the estimated volatility bound. This shift may very well be the result of estimation error however, and the main question is whether the observed shift is too large to be attributed to chance. Therefore, to answer the question whether or not the observed shift in the mean-variance frontier is significant in statistical terms, in this section we will also show how regression analysis can be used to test for spanning and intersection.

3.1 Spanning and intersection in terms of mean-variance frontiers

To state the problem formally, the hypothesis of mean-variance intersection means that there is a portfolio w^* which is mean-variance efficient for the smaller set R_{t+1} and which is also mean-variance efficient for the larger set $(R_{t+1}; r_{t+1})$. In the sequel, variables that refer to the smaller set R_{t+1} (r_{t+1}) will be referred to with a subscript R (r), or with their dimension K (N), whereas variables that refer to the larger set $(R_{t+1}; r_{t+1})$, will not have any subscript or will have their dimension as subscript, $K + N$. Thus, w_R is a K-dimensional vector with portfolio weights for the assets in R_{t+1}, and w is a $(K + N)$-dimensional vector with portfolio weights for all the available assets $(R_{t+1}; r_{t+1})$. The hypothesis of mean-variance intersection comes down to the statement that there exists a mean-variance efficient portfolio w^* of the form

$$w^* = \begin{pmatrix} w_R^* \\ 0_N \end{pmatrix}; \tag{10}$$

i.e., there exist scalars γ and η, such that

$$\mu - \eta\iota_{K+N} = \gamma\Sigma \begin{pmatrix} w_R^* \\ 0_N \end{pmatrix}. \tag{11}$$

If such a portfolio w^* exists, there is one point on the mean-variance frontier of R_{t+1} that also lies on the mean-variance frontier of (R_{t+1}, r_{t+1}). Using obvious notation, μ consists of two subvectors μ_R and μ_r, and Σ consists of submatrices Σ_{RR}, Σ_{Rr}, Σ_{rR}, and Σ_{rr}. The first K rows of (11) imply that

$$\mu_R - \eta\iota_K = \gamma\Sigma_{RR}w_R^* \iff w_R^* = \gamma^{-1}\Sigma_{RR}^{-1}(\mu_R - \eta\iota_K). \tag{12}$$

Equation (12) simply says that w_R^* is indeed mean-variance efficient for the smaller set R_{t+1}.

The next step is to derive the restrictions on the distribution of R_{t+1} and r_{t+1} that are equivalent to mean-variance intersection. In order to do so, substitute (12) in the last N rows of (11) to obtain:

$$\begin{aligned} \mu_r - \eta\iota_N &= \Sigma_{rR}\Sigma_{RR}^{-1}(\mu_R - \eta\iota_K) \iff \\ & (\mu_r - \beta\mu_R) + (\beta\iota_K - \iota_N)\eta = 0, \end{aligned} \tag{13}$$

with $\beta \equiv \Sigma_{rR}\Sigma_{RR}^{-1}$. Thus, if there is a portfolio that is mean-variance efficient for the smaller set R_{t+1} that is also mean-variance efficient for the larger set (R_{t+1}, r_{t+1}), there must exist an η such that the restriction in (13) holds. It follows immediately from the derivation above that this η is the zero-beta return that corresponds to the portfolio w_R^* (and w^*).

If there is mean-variance spanning then all mean-variance efficient portfolios w^* must be of the form (10), i.e., (11) must be true for all values of η and the corresponding γ's. Going through the same steps, if (11) must hold for any η, (13) must hold for any η, and this can only be the case if

$$\mu_r - \beta\mu_R = 0 \text{ and } \beta\iota_K - \iota_N = 0, \tag{14}$$

which are the restrictions imposed by the hypothesis of spanning. If these restrictions on the distribution of R_{t+1} and r_{t+1} hold, every point on the mean-variance frontier of R_{t+1} is also on the mean-variance frontier of (R_{t+1}, r_{t+1}) and the two frontiers coincide.

3.2 Spanning and intersection in terms of volatility bounds

In the previous section we defined mean-variance spanning and intersection from the properties of mean-variance efficient portfolios and we derived the equivalent restrictions on the distribution of asset returns, which have previously been derived by Huberman and Kandel (1987). In this section we analyze mean-variance intersection and spanning from the properties of minimum variance stochastic discount factors that price the assets in R_{t+1} and in $(R_{t+1}; r_{t+1})$ correctly and we show that this imposes the same restrictions on the distribution of the asset returns. In terms of volatility bounds, the hypothesis of intersection is that there is a value of v such that the minimum variance stochastic discount factor for R_{t+1}, i.e., $m_R(v)_{t+1}$, is also the minimum variance stochastic discount factor for the larger set $(R_{t+1}; r_{t+1})$. The discount factor $m_R(v)_{t+1}$ as defined by (4) and (5) is the minimum variance stochastic discount factor for this larger set if it also prices r_{t+1} correctly. If $m_R(v)_{t+1}$ prices both R_{t+1} and r_{t+1} correctly, the difference between $m_R(v)_{t+1}$ and any other $M(v)_{t+1}$ that prices R_{t+1} and r_{t+1} correctly is orthogonal to R_{t+1} and r_{t+1}, implying that $m_R(v)_{t+1}$ must have the lowest variance among all stochastic discount factors $M(v)_{t+1}$, by the same reasoning that leads to (6).

Thus, the hypothesis of intersection for volatility bounds can be stated as:

$$\exists v \text{ s.t. } E[r_{t+1} m_R(v)_{t+1}] = \iota_N. \tag{15}$$

To show that this hypothesis imposes the same restrictions on the distribution of R_{t+1} and r_{t+1} as in (13), substitute (4) and (5) into (15):

$$\begin{aligned} E[r_{t+1}(v + (R_{t+1} - \mu_R)'\Sigma_{RR}^{-1}(\iota_K - v\mu_R))] &= \iota_N; \quad \Leftrightarrow \\ (\mu_r - \Sigma_{rR}\Sigma_{RR}^{-1}\mu_R)v + (\Sigma_{rR}\Sigma_{RR}^{-1}\iota_K - \iota_N) &= 0; \quad \Leftrightarrow \\ (\mu_r - \beta\mu_R)v + (\beta\iota_K - \iota_N) &= 0. \end{aligned} \tag{16}$$

Dividing both sides of (16) by v shows that the hypothesis of intersection in terms of volatility bounds indeed implies the same restrictions as the hypothesis of intersection in terms of mean-variance frontiers, if we choose $\eta = 1/v$. This could be expected beforehand, since from the duality between mean-variance frontiers and volatility bounds in (9) we already knew that the vector $\varphi_R(v)$ that defines $m_R(v)_{t+1}$, is proportional to a mean-variance efficient portfolio with zero-beta return $\eta = 1/v$. The hypothesis that w^* is of the form $(w_R^{p\prime}\ 0_N')'$ is therefore equivalent the hypothesis that $\varphi(v)$ is of the form $(\varphi_R(v)'\ 0_N')'$.

By the same logic, the hypothesis of spanning in terms of volatility bounds, requires that $m_R(v)_{t+1}$ prices the returns r_{t+1} for all values of v:

$$E[r_{t+1} m_R(v)_{t+1}] = \iota_N, \quad \forall v, \tag{17}$$

since in that case the entire volatility bound derived from (R_{t+1}, r_{t+1}) coincides with the volatility bound derived from (R_{t+1}) only. This requirement implies that (16) holds for all values of v, and this can only be the case if the restrictions in (14) hold.

3.3 Intersection and mean-variance efficiency of a given portfolio

A question that is of obvious interest both from a portfolio choice perspective and from an asset pricing perspective, is the question whether or not a given portfolio w^p is mean-variance efficient or not. From a portfolio choice perspective, an investor will be interested in whether or not his portfolio has the desired properties of a mean-variance efficient portfolio. From an asset pricing perspective, the frequently analyzed question is, e.g., whether or not the market portfolio is mean-variance efficient as the CAPM predicts. Alternative asset pricing models may identify other portfolios as being mean-variance efficient. For instance, in the Consumption-CAPM the portfolio that mimics aggregate per-capita consumption is mean-variance efficient and the Intertemporal-CAPM implies that a combination of the market portfolio and the portfolios hedging changes in the investment-opportunity set is mean-variance efficient.

Denote the return on some portfolio w^p by R^p_{t+1} and its expectation by μ^p. The question whether or not w^p is mean-variance efficient with respect to the $N+1$ assets (R^p_{t+1}, r_{t+1}), is obviously a special case of the question whether or not there is mean-variance intersection with $K = 1$ and $R_{t+1} = R^p_{t+1}$, since intersection in this case simply means that the portfolio w^p is on the mean-variance frontier of (R^p_{t+1}, r_{t+1}). Therefore, if w^p is mean-variance efficient for the set (R^p_{t+1}, r_{t+1}), the following restrictions on the distribution of R^p_{t+1} and r_{t+1} should hold:

$$\mu_r = \eta\iota_N + \beta^p(\mu^p - \eta), \tag{18}$$

where β^p is the N-dimensional vector $Cov[r_{t+1}; R^p_{t+1}]/Var[R^p_{t+1}]$, and $\mu^p = E[R^p_{t+1}]$. When testing for mean-variance efficiency, R^p_{t+1} is usually the return on a portfolio of r_{t+1}.

What we want to establish in this section however, is that the hypothesis that the mean-variance frontier of R_{t+1} ($K \geq 1$) intersects the frontier of $(R_{t+1}; r_{t+1})$ at a given value of $\eta = 1/v$, is tantamount to the hypothesis that the portfolio w^*_R that is mean-variance efficient for R_{t+1} and that has η as its zero-beta rate is also mean-variance efficient with respect to $(R_{t+1}; r_{t+1})$. Denote the return on w^*_R as R^*_{t+1} and its expectation as μ^*. Recall that the portfolio w^*_R is given by the first K rows of (11)

$$w^*_R = \gamma^{-1}\Sigma^{-1}_{RR}(\mu_R - \eta\iota_K);$$

from which

$$w^{*\prime}_R(\mu_R - \eta\iota_K) = \gamma w^{*\prime}_R\Sigma_{RR}w^*_R \;, \quad \gamma = \frac{\mu^* - \eta}{Var[R^*_{t+1}]}.$$

Substituting these relations into (11) and defining $\beta^* \equiv Cov[r_{t+1}; R^*_{t+1}]/Var[R^*_{t+1}]$, results in

$$0 = (\mu_r - \beta^*\mu^*) + (\beta^* - \iota_N)\eta. \qquad (19)$$

These are the same restrictions as (18) for $w^p = w^*$. Thus, the hypothesis of intersection indeed implies the same restrictions on the distribution of R_{t+1} and r_{t+1} as the hypothesis that w^*_R is mean-variance efficient with respect to r_{t+1}.

3.4 Testing for spanning and intersection

So far we derived the restrictions implied by the hypotheses of mean-variance intersection and spanning for the distribution of R_{t+1} and r_{t+1}. Huberman and Kandel (1987) showed how regression analysis can be used to test these hypotheses. To see how regression analysis can be used to test for intersection, start from (13):

$$\mu_r - \eta\iota_N = \beta(\mu_R - \eta\iota_K).$$

Replacing the expected returns μ_r and μ_R with realized returns r_{t+1} and R_{t+1}, gives the regression

$$r_{t+1} = \alpha + \beta R_{t+1} + \varepsilon_{t+1}, \qquad (20)$$

with $\alpha = \mu_r - \beta\mu_R$, $\varepsilon_{t+1} = u_{r,t+1} - \beta u_{R,t+1}$, $u_{r,t+1} \equiv r_{t+1} - \mu_r$ and $u_{R,t+1} \equiv R_{t+1} - \mu_R$. It can readily be checked that under the null hypotheses of spanning and intersection $\mathrm{Cov}[\varepsilon_{t+1}; R_{t+1}] = 0$. Notice that α is an N-dimensional vector of intercepts, β is an $N \times K$-dimensional matrix of slope coefficients, and ε_{t+1} is an N-dimensional vector of error terms. The restrictions imposed by the hypothesis of intersection in (13) can now be stated as

$$\alpha - \eta(\iota_N - \beta\iota_K) = 0. \tag{21}$$

With intersection there are two cases of interest. First, we may be interested in testing for intersection for a given value of the zero-beta rate η. In that case the restrictions in (21) should hold for this specific value of η, which is a set of linear restrictions. In the sequel we will mainly be interested in this case. Second, the interest may be in the question whether there is intersection at some unknown point of the frontier, i.e., for some unknown value of η. In that case the hypothesis is that there exists some η such that the restrictions in (21) hold. This hypothesis can be stated as

$$\alpha_i/(1 - \beta_i\iota_K) = \alpha_j/(1 - \beta_j\iota_K); \quad i, j = 1, \ldots, N;$$

where β_i is the ith row of β. Thus, the hypothesis that there is intersection at some point of the frontier imposes a set of nonlinear restrictions on the regression parameters in (20). Notice that given estimates of α_i and β_i an estimate of the zero-beta rate for which there is intersection can be obtained from $\alpha_i/(1 - \beta_i\iota_K)$. Also note, that testing whether there is intersection at some unknown point of the frontier only makes sense if $N \geq 2$, since there is always intersection if $N = 1$. (Because there is always one efficient portfolio for which the weight in the new asset is zero.)

Recall that the hypothesis of spanning implies that (21) holds for all values of η. Therefore, going through the same steps, the restrictions imposed by the hypothesis of spanning can be stated as

$$\alpha = 0 \text{ and } \beta\iota_K - \iota_N = 0. \tag{22}$$

The restrictions in terms of the regression model in (20) are intuitively very clear. For instance, the spanning restrictions in (22) state that if there is spanning, then each return of the additional assets, $r_{i,t+1}$, $i = 1, 2, \ldots, N$, can be written as the return of a portfolio of the benchmark assets $\beta_i R_{t+1}$, $\beta_i\iota_K = 1$, plus an error term $\varepsilon_{i,t+1}$ which has expectation zero and which is orthogonal to the returns R_{t+1}. Since such an asset can only add to the variance of portfolios of R_{t+1}, and not to the expected return, mean-variance optimizing agents will not include such an asset in their portfolio. A similar interpretation holds for the intersection restrictions.

If the returns series R_{t+1} and r_{t+1} are stationary and ergodic, consistent estimates of the parameters α and β in (20) are easily obtained using OLS. In writing down the test statistics for (21) and (22), it is convenient to use a different specification of (20), in which all the coefficients α and β are stacked into one big vector:

$$r_{t+1} = \left(I_N \otimes \left(1 \quad R'_{t+1}\right)\right) b + \varepsilon_{t+1}; \tag{23}$$

where $b = \mathrm{vec}\left(\left(\alpha \quad \beta\right)'\right)$, a $(K+1)N$-dimensional vector. If $\hat{b}$ is the OLS estimate of b and $\hat{\Omega}$ is a consistent estimate of the asymptotic covariance matrix of $\hat{b}$, the hypotheses of intersection and spanning can be tested using a standard Wald test. Defining

$$H(\eta)_{int} \equiv I_N \otimes \left(1 \quad \eta\iota'_K\right) \quad \text{and} \tag{24a}$$

$$h(\eta)_{int} \equiv H(\eta)_{int}\hat{b} - \eta\iota_N, \tag{24b}$$

the Wald test-statistic for intersection can be written as

$$\xi_W^{int} = h(\eta)'_{int}\left(H(\eta)_{int}\hat{\Omega}H(\eta)'_{int}\right)^{-1} h(\eta)_{int}. \tag{25}$$

Similarly, defining

$$H_{span} \equiv I_N \otimes \begin{pmatrix} 1 & 0'_K \\ 0 & \iota'_K \end{pmatrix} \quad \text{and} \tag{26a}$$

$$h_{span} \equiv H_{span}\hat{b} - \iota_N \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}; \tag{26b}$$

the Wald test-statistic for spanning can be written as

$$\xi_W^{span} = h'_{span}\left(H_{span}\hat{\Omega}H'_{span}\right)^{-1} h_{span}. \tag{27}$$

Under the null hypotheses and standard regularity conditions, the limit distribution of ξ_W^{int} will be χ^2_N and the limit distribution of ξ_W^{span} will be χ^2_{2N}. The test statistics in (25) and (27) have interesting economic interpretations in terms of performance measures. The relationship between tests for intersection and spanning and performance evaluation will be discussed in detail in Section 5.3.

Chen and Knez (1996) and Hall and Knez (1995) propose a test for intersection that is based on (15). Define the deviation from the equality in (15) to be $\lambda(v)$:

$$\lambda(v) \equiv E[m_R(v)_t r_t] - \iota_N. \tag{28}$$

In Section 5.1 we will interpret $\lambda(v)$ scaled by v as a generalization of the well-known Jensen measure. Given an estimate of the parameters $\varphi_R(v)$ using the sample equivalent of (5):

$$\hat{\varphi}_R(v) = \left(\frac{1}{T}\sum_{t=1}(R_t - \overline{R})(R_t - \overline{R})'\right)^{-1}\left(\iota_K - v\overline{R}\right);$$

with $\overline{R}$ the sample mean of R_t, define $\hat{\lambda}(v)_t$ as

$$\hat{\lambda}(v)_t \equiv r_t(v + \hat{\varphi}_R(v)'(R_t - \overline{R})) - \iota_N.$$

A test for the hypothesis of intersection, $\lambda(v) = 0$, can now be based on

$$\xi^{int}_{CK} = \left(\frac{1}{T}\sum_{t=1}\hat{\lambda}(v)_t\right)'\left(\widehat{Var}[\hat{\lambda}(v)_t]\right)^{-1}\left(\frac{1}{T}\sum_{t=1}\hat{\lambda}(v)_t\right); \tag{29}$$

where the estimate $\widehat{Var}[\hat{\lambda}(v)_t]$ can for instance be obtained using the method suggested by Newey and West (1987). The limit distribution of the test-statistic ξ^{int}_{CK} is also χ^2_N. Since for $\eta = 1/v$, we have

$$\begin{aligned}\left(\frac{1}{T}\sum_{t=1}\hat{\lambda}(v)_t\right)/v &= \frac{1}{T}\sum_{t=1}\left[r_t + r_t\left(R_t - \overline{R}\right)'\hat{\varphi}_R(v)\eta - \eta\iota_N\right]\\ &= \hat{\alpha} + \left(\hat{B}\iota_K - \iota_N\right)\eta;\end{aligned}$$

it follows that

$$\left(\frac{1}{T}\sum_{t=1}\hat{\lambda}(v)_t\right)/v = H(\eta)_{int}\hat{\theta} - h(\eta)_{int};$$

and that the only difference in the Wald test-statistic in (25) and the statistic proposed in (29) is the way in which the covariance matrix is estimated.

A disadvantage of the test originally proposed by Chen and Knez (1996) is that they test for intersection for a very specific stochastic discount factor, which corresponds to the minimum second moment portfolio. This discount factor can be found by projecting the kernel M_{t+1} on the asset returns only, excluding the constant. The corresponding portfolio on the mean-variance frontier is the one with the minimum second moment among all portfolios on the frontier, and can graphically be found as the tangency point between the mean-variance frontier and a circle with its centre at the origin. The problem with this portfolio is that it is located at the inefficient part of the frontier, implying that the test used by Chen and Knez (1995) is for intersection at an inefficient portfolio. Therefore it is economically not very interesting, unless a risk free asset is included. Since in the test statistic in (29) the discount factor $m_R(v)_{t+1}$ results from a projection of M_{t+1} on R_{t+1} plus a constant, this test allows us to test for intersection at any mean-variance efficient portfolio, so this test does not suffer from the problem of the test originally suggested by Chen and Knez. Dahlquist and Söderlind (1999), who use the test proposed by Chen and Knez to evaluate the performance of Swedish mutual funds, also acknowledge this problem and add a constant to the set R_{t+1} such that the conditional mean of $m_r(v)_{t+1}$, equals one over the risk free rate, i.e., $v_t = 1/R_{f;t}$.

The distinction between the Wald tests in (25) and (27) on the one hand and the tests proposed by Chen and Knez in (29) is similar to the distinction between tests based on the (traditional) regression methodology and on the SDF methodology as discussed in Kan and Zhou (1999). Their simulations suggest that in small samples tests based on the regression methodology have better size and power properties than tests based on the SDF methodology, which indicates that the test in (25) may be preferred to (29).

Alternative tests for the hypotheses of intersection and spanning are suggested, e.g., by Huberman and Kandel (1987), who propose a likelihood ratio test, and by Snow (1991) and DeSantis (1995), who propose a Generalized Method of Moments (GMM) procedure. This latter procedure is also identical to the region subset test suggested by Hansen, Heaton and Luttmer (1995) which is equivalent to a test for intersection. A comparison of the small sample properties of various test-procedures can be found in Bekaert and Urias (1996). Their small sample results suggest that the likelihood test for spanning as proposed by Huberman and Kandel has better power properties than the GMM-based tests, while it also has a size distortion that is in most cases not worse than for the GMM-based tests. The GMM-based test or region subset test is based on the observation that under the null hypotheses of spanning or intersection, the kernel that prices R_{t+1} and r_{t+1}

correctly is of the form

$$
\begin{aligned}
m(v)_{t+1} &= v + \varphi_R(v)'(R_{t+1} - \mu_R) + \varphi_r(v)'(r_{t+1} - \mu_r), \\
\text{with } \varphi_r(v) &= 0.
\end{aligned}
$$

Given that $\varphi_r(v) = 0$, a GMM-estimate of the K parameters in $\varphi_R(v)$ can be obtained by using the $K + N$ sample moments

$$
g_T(\varphi_R(v)) = \frac{1}{T}\sum_{t=1}^{T}\left\{\begin{pmatrix} R_t \\ r_t \end{pmatrix}(v + \varphi_R(v)'(R_t - \overline{R})) - \iota_{K+N}\right\} = \frac{1}{T}\sum_{t=1}^{T} g_t(\varphi_R(v)).
$$

A consistent estimate of $\varphi_R(v)$ can therefore be obtained by solving

$$
\min_{\varphi_R(v)} g_T(\varphi_R(v))' W_T g_T(\varphi_R(v)) = J_T(\varphi_R(v)); \tag{30}
$$

where W_T is a symmetric nonsingular weighting matrix. Notice that the GMM-estimate of the K parameters $\varphi_R(v)$ obtained from (30) is based on $K + N$ moment restrictions. The N overidentifying restrictions are derived from the hypothesis that $m_R(v)_{t+1}$ must also price the N additional assets r_{t+1}. Intersection for a given value of v can now be tested by using the fact that under the null-hypothesis and regularity conditions $T J_T(a_R(v))$ is asymptotically χ^2_N-distributed. Since spanning implies that (15) holds for (at least) two different values of v, the GMM-based test can easily be extended by estimating two vectors $\varphi_R(v_1)$ and $\varphi_R(v_2)$ simultaneously ($v_1 \neq v_2$) using (30). In this case there are $2K$ parameters to be estimated with $2(K + N)$ moment conditions. The test for spanning is therefore a test for the $2N$ overidentifying restrictions and will asymptotically be χ^2_{2N}-distributed under the null-hypothesis of spanning.

4 Testing for spanning and intersection with conditioning information

The purpose of this section is to incorporate conditioning information in tests for intersection and spanning. Until now we assumed that returns are independently and identically distributed (i.i.d.). However, there is ample evidence that asset returns are to some extent predictable. For instance, stock and bond returns can be predicted from variables like lagged returns, dividend yields, short term interest rates, and default premiums (see, e.g., Ferson (1995)) and futures returns can be predicted from hedging pressure variables (see e.g. DeRoon, Nijman and Veld (2000)) as well as from the spread between spot and forward prices (see, e.g., Fama (1984)). Kirby (1998) analyzes whether predictability of security returns is consistent with rational asset pricing. He shows that the covariance between the pricing kernel implied by an asset pricing model and conditioning variables, restricts the slope coefficients in a regression of security returns on those same conditioning variables. In Section 4.1 we will show how conditional information can be used in a straightforward way by using scaled returns (see, e.g., Cochrane (1997) and Bekaert and Urias (1996)). Although this is a fairly general and intuitive way of incorporating conditional information, a disadvantage of this method is that the dimension of the estimation and testing problem increases quickly. In Section 4.2 we show that this problem can be circumvented if it is assumed that variances and covariances are constant, while expected returns are allowed to vary over time, although this assumption is not in accordance with most equilibrium models and with the empirical evidence regarding time-varying second moments. Using this simplifying assumption however, it is shown that the conditioning variables can easily be accounted for by using them as additional regressors. The restrictions for the intersection and spanning hypotheses then become similar to the restrictions in the i.i.d. case. This way of incorporating conditional variables also has the additional advantage that the regression estimates indicate under what economic circumstances, i.e., for what values of the conditioning variables, intersection and spanning can or can not be rejected. Finally, in Section 4.3 we will discuss the use of conditioning variables as, e.g., in Shanken (1990) and Ferson and Schadt (1996).

4.1 Incorporating conditional information using scaled returns

Suppose that z_t is an $(L-1)$-dimensional vector of instruments that has predictive power for R_{t+1} and r_{t+1}, and define the L-dimensional vector Z_t as $Z_t \equiv (1\ z_t')'$. A common way to use these instruments is to look at scaled returns: $Z_t \otimes R_{t+1}$. If M_{t+1} is a valid stochastic discount factor, then from (1) we have:

$$E[M_{t+1}(Z_t \otimes R_{t+1}) \mid I_t] = Z_t \otimes \iota_K.$$

Taking unconditional expectations, this yields

$$E[M_{t+1}(Z_t \otimes R_{t+1})] = E[Z_t \otimes \iota_K]. \tag{31}$$

Thus, the scaled return $Z_{i,t}R_{j,t+1}$ has an average price equal to $E[Z_{i,t}]$. The scaled returns can be interpreted as the payoffs of a strategy where each period an amount equal to $Z_{i,t}$ dollars is invested in a security, yielding a payoff equal to $Z_{i,t}R_{j,t+1}$. Therefore, we can also think of $Z_t \otimes R_{t+1}$ as the returns on managed portfolios (see, e.g., Cochrane (1997)). By allowing for such managed portfolios, we take into account that investors may use dynamic strategies, based on the realized values of Z_t. In effect this increases the set of available assets by a factor L (i.e., from K to $K \times L$).

To simplify notation, denote the $(L \times K)$-dimensional vector $Z_t \otimes R_{t+1}$ by R^Z_{t+1}. Also, denote the $(L \times K)$-dimensional vector $E[Z_t \otimes \iota_K]$ by q_K. For further reference, r^Z_{t+1} and q_N are defined in a completely analogous way and we use a superscript Z for all variables and parameters that correspond to R^Z_{t+1} and r^Z_{t+1}. Valid stochastic discount factors M^Z_{t+1} now have to satisfy

$$E[M^Z_{t+1}R^Z_{t+1}] = q_K. \tag{32}$$

As shown by Bekaert and Urias (1996), following the same line of reasoning as in Sections 2.1 and 2.2, it is straightforward to show that the minimum variance stochastic discount factor with expectation v is given by

$$\begin{aligned} m^Z_R(v)_{t+1} &= v + \alpha^Z(v)'(R^Z_{t+1} - \mu^Z_R); \\ \alpha^Z(v) &= (\Sigma^Z_{RR})^{-1}(q_K - v\mu^Z_R). \end{aligned} \tag{33}$$

This expression for the volatility bound is a straightforward generalization of the one given in (4) and (5). The restrictions imposed by the hypotheses of intersection and spanning also turn out to be very similar to the ones given in previous sections, as we will see below.

Thus, conditioning information can be incorporated by including managed portfolios, the returns of which depend on the conditioning variables. If there is to be conditional intersection or spanning of r_{t+1} by R_{t+1}, the unconditional volatility bound (or mean-variance frontier) of R^Z_{t+1} must intersect or span the volatility bound (or mean-variance frontier) of (R^Z_{t+1}, r^Z_{t+1}). The interest is therefore in the returns R_{t+1} and r_{t+1} themselves plus the returns on all the managed portfolios. Intersection or spanning is equivalent to

$$E[r^Z_{t+1}m^Z_R(v)_{t+1}] = q_N; \tag{34}$$

for one value of v or for all values of v respectively. To see which restrictions these hypotheses imply, substitute (33) into (34) to obtain

$$(\mu_r^Z - \beta^Z \mu_R^Z)v + (\beta^Z q_K - q_N) = 0; \tag{35}$$

for intersection, and

$$(\mu_r^Z - \beta^Z \mu_R^Z) = 0; \quad \text{and} \quad (\beta^Z q_K - q_N) = 0; \tag{36}$$

for spanning. Here $\beta^Z = \Sigma_{rR}^Z \left(\Sigma_{RR}^Z\right)^{-1}$ is a $(L \times N) \times (L \times K)$ matrix with slope coefficients from a regression of r_{t+1}^Z on R_{t+1}^Z plus a constant. These restrictions are also given in Bekaert and Urias (1996). Regressing r_t^Z on R_t^Z to incorporate conditioning information is very similar to the approach to be discussed in Section 4.3, where the regression parameters α and β are time varying. In that section we will assume that the mean returns and the (co)variances are functions of the instruments that can be linearized using a Taylor series approximation, leading to a similar regresssion as in the case discussed here. Therefore, the use of scaled returns can also be motivated as a convenient way of dealing with time-varying means and variances.

The similarity with the case in which there was no conditioning information is obvious. The only difference in the restrictions is that in (35) and (36) we have $(\beta^Z q_K - q_N)$ instead of $(\beta \iota_K - \iota_N)$. The fact that q_K and q_N enter the restrictions reflects the fact that R_{t+1}^Z and r_{t+1}^Z are not really returns, in the sense that their current prices are not necessarily equal to one. The average prices of R_{t+1}^Z and r_{t+1}^Z are instead given by q_K and q_N. The average cost of the managed portfolios with payoff vector r_{t+1}^Z is given by the vector q_N, and the cost of the mimicking portfolios from R_{t+1}^Z is given by $\beta^Z q_K$. The interpretation of the restrictions given in Section 3.4 is therefore still valid.

The main disadvantage of this way of incorporating conditioning information is that the number of parameters to be estimated as well as the number of restrictions to be tested grows rapidly with the number of instruments L. The number of exogenous variables equals $K \times L$ and the number of restrictions to be tested equals $N \times L$ for the hypothesis of intersection, and $2N \times L$ for the hypothesis of spanning. This is the case because for each new instrument there are K new managed portfolios to be considered for the assets in R_{t+1} and N additional managed portfolios for the assets in r_{t+1}.

This problem can at least partially be circumvented if we are willing to assume a more specific form of predictability. Specifically, in the next section we make the assumption that only the expected returns of R_{t+1} and r_{t+1} depend linearly on the instruments z_t, whereas all variances and covariances are constants. In Section 4.3 the slope coefficients β are assumed to depend linearly on the instruments, which also allows for a straightforward way of incorporating conditional information in the regression framework to test for intersection and spanning.

4.2 Expected returns linear in the conditional variables

In this section we assume that there is a specific form of predictability, which allows us to incorporate conditioning information in a straightforward way in the regression framework for spanning and intersection. The assumption made is that expected returns are linear in the conditional variables and that returns are conditionally homoskedastic. This way of incorporating conditioning information is used in Harvey (1989), as well as, for instance, in Campbell and Viceira (1998) and DeRoon, Nijman and Werker (1998). The assumption we make is that

$$\begin{aligned} E_t[R_{t+1}] &= c_R' Z_t; \\ E_t[r_{t+1}] &= c_r' Z_t; \end{aligned} \tag{37}$$

and the variances and covariances of R_{t+1} and r_{t+1} conditional on Z_t are given by $Var[R_{t+1} \mid Z_t] = \Sigma_{RR}$, $Var[r_{t+1} \mid Z_t] = \Sigma_{rr}$, and $Cov[r_{t+1}, R_{t+1} \mid Z_t] = \Sigma_{rR}$. Starting from (1), the minimum variance stochastic discount factor, conditional on Z_t, is given by

$$\begin{aligned} m_R(v_t)_{t+1} &= v_t + \varphi(v_t)_t'(R_{t+1} - E_t[R_{t+1}]); \\ \varphi(v_t)_t &= \Sigma_{RR}^{-1}(\mu_K - v_t E_t[R_{t+1}]). \end{aligned} \tag{38}$$

Notice that since the projection of the kernel on the asset returns is now conditional on Z_t, we explicitly allow for time variation in the coefficients $\varphi(v_t)_t$, as well as in v_t, the conditional expectation of the stochastic discount factor. Also note that in describing the conditional mean-variance frontier or volatility bound we still can use v_t as a free parameter.

If there is intersection, $m_R(v_t)_{t+1}$ must price r_{t+1} correctly conditional on Z_t, which results in

$$\begin{aligned} \mu_N &= E_t[r_{t+1} m_R(v_t)_{t+1}] = v_t c_r' Z_t + \Sigma_{rR}\Sigma_{RR}^{-1}(\mu_K - c_R' Z_t) \\ &\Leftrightarrow \quad (c_r' - \Sigma_{rR}\Sigma_{RR}^{-1} c_R') Z_t v_t + (\Sigma_{rR}\Sigma_{RR}^{-1}\mu_K - \mu_N) = 0. \end{aligned} \tag{39}$$

In case there is spanning this condition must again hold for every v_t, implying

$$(c_r' - \Sigma_{rR}\Sigma_{RR}^{-1}c_R')Z_t = 0 \quad \text{and} \quad (\Sigma_{rR}\Sigma_{RR}^{-1}\iota_K - \iota_N) = 0. \tag{40}$$

It turns out that the regression framework that we used to test for spanning and intersection can be modified to test the restrictions in (39) and (40). Straightforward use of the algebra of partitioned matrices shows that in the regression

$$r_{t+1} = cZ_t + dR_{t+1} + u_{t+1}, \tag{41}$$

with $E[u_{t+1}Z_t] = 0$, and $E[u_{t+1}R_{t+1}] = 0$, the OLS-estimates of c and d are consistent estimates of $(c_r' - \Sigma_{rR}\Sigma_{RR}^{-1}c_R')$ and $(\Sigma_{rR}\Sigma_{RR}^{-1}\iota_K - \iota_N)$ respectively, which are the parameters of interest in the restrictions in (39) and (40) (see DeRoon, Nijman, and Werker (1998)). The hypotheses of intersection and spanning can therefore be based on the OLS-estimates of (41). The hypothesis that there is intersection for a given value of v_t and Z_t can be tested by testing the restrictions

$$cZ_tv_t + (d\iota_K - \iota_N) = 0; \tag{42}$$

and the hypothesis of spanning by testing the restrictions

$$cZ_t = 0 \quad \text{and} \quad (d\iota_K - \iota_N) = 0. \tag{43}$$

These restrictions are very similar to the restrictions implied by intersection and spanning in the unconditional case, except that the intercept α in (20) is replaced by cZ_t.

It can easily be seen from (42) and (43) that the number of restrictions to be tested for intersection and spanning is the same as in the unconditional case, which makes this method of incorporating conditional information more parsimonious than using scaled returns. Note that the hypotheses underlying (42) and (43) are that there is intersection or spanning for a particular value of Z_t, i.e., for a particular state of the economy. This has the additional advantage that the regression estimates of (41) make it possible to derive confidence intervals for the values of Z_t for which there can be intersection or spanning.

If the hypothesis of interest is whether there is spanning regardless of the state of the economy, the restrictions in (43) should hold for all values of z_t, implying that each element of c should be equal to 0. In that case, with L instruments and N assets in r_{t+1}, there are $L \times N$ restrictions to be tested, which, although smaller than the $2 \times L \times N$ restrictions in (36), can be a large number. Also, as follows readily from (42) and (43), in this case the hypothesis of intersection and the hypothesis of spanning both imply the same restrictions. This latter result is due to the fact that the value of v_t for which we test intersection is constant. Since the tangency point on the mean-variance frontier that corresponds to v_t is a function of Z_t, the only way to have intersection irrespective of the specific value of Z_t is to have spanning.

4.3 Regression coefficients linear in the conditional variables

An alternative way of incorporating conditional information in the regression framework is suggested by Shanken (1990) and Ferson and Schadt (1996) e.g., where the coefficients α and β are assumed to be a linear function of the instruments. In the regression in (20), the ith row can be written as

$$r_{i;t+1} = \alpha_i + \beta_i R_{t+1} + \varepsilon_{t+1}.$$

Shanken (1990) simply assumes that

$$\begin{aligned} \alpha_i &= a_{i0} + z_t' a_{i1}; \\ \beta_i &= b_{i0} + z_t' b_{i1}; \end{aligned} \tag{44}$$

where z_t are now supposed to be L demeaned variables. Here a_{i0} is scalar, a_{i1} is an L-vector, b_{i0} is a K row-vector, and b_{i1} is $L \times K$ matrix. Ferson and Schadt (1996) motivate (44) as a first order Taylor-series expansion for a general dependence of β on $Z_t = (1\ z_t')'$. Let $Cov[r_{t+1}; R_{t+1} \mid Z_t] = \Sigma_{rR}(Z_t)$, and $Var[R_{t+1} \mid Z_t] = \Sigma_{RR}(Z_t)$, where $\Sigma(.)$ indicates some functional form for the covariance matrix. Starting from (13) intersection for a given zero-beta rate $\eta_t = 1/v_t$ conditional on Z_t means

$$\begin{aligned} E[r_{t+1} - \eta_t \iota_N] &= \beta(Z_t) E[R_{t+1} - \eta_t \iota_K]\ , \\ r_{t+1} - \eta_t \iota_N &= \beta(Z_t)(R_{t+1} - \eta_t \iota_K) + u_{t+1}; \end{aligned}$$

with $\beta(Z_t) = \Sigma_{rR}(Z_t)\Sigma_{RR}(Z_t)^{-1}$, $u_{t+1} \equiv (r_{t+1} - \beta(Z_t)R_{t+1}) - (E[r_{t+1}] - \beta(Z_t)E[R_{t+1}])$, and $E[u_{t+1} \mid Z_t] = 0$. Ferson and Schadt (1996) suggest a linear approximation of $\beta_i(Z_t)$:

$$\beta_i(Z_t) \approx b_{i0} + z_t' b_{i1}; \tag{45}$$

from which

$$\begin{aligned} r_{i,t+1} &= a_{i0} + z_t' a_{i1} + b_{i0} R_{t+1} + (z_t' b_{i1})\, R_{t+1} + \varepsilon_{t+1}, \\ a_{i0} &= \eta_t (1 - b_{i0} \iota_K), \\ a_{i1} &= -\eta_t b_{i1} \iota_K, \end{aligned} \tag{46}$$

with $\varepsilon_{i,t+1} = u_{i,t+1} + (\beta_i(Z_t) - b_{i0}' - (z_t' b_{i1}))(R_{t+1} - \eta_t \iota_K)$, for which it is assumed that $E[\varepsilon_{i,t+1} \mid Z_t] = 0$. This yields precisely the regression in (20) where the regression parameters are linear in the instruments as assumed by Shanken (1990).

Intersection for a given value of $\eta_t = 1/v_t$ and z_t can now be tested by testing the restrictions that

$$(a_{i0} + z_t' a_{i1}) + \{(b_0 + z_t' b_{i1}) \iota_K - 1\} \eta_t = 0. \tag{47}$$

As in the previous section, these restrictions have the additional advantage that statements can be made about in which state of the economy, (i.e., values of z_t) there is intersection. If there is intersection for all values of z_t, this implies

$$\begin{aligned} a_{i0} + (b_{i0} \iota_K - 1) \eta_t &= 0, \\ a_{i1} + b_{i1} \iota_K \eta_t &= 0. \end{aligned} \tag{48}$$

The regression in (46) can also be motivated from the scaled returns in Section 4.1. Using the pricing kernel that is linear in R^Z_{t+1} and that is supposed to price the returns r^Z_{t+1} as well, the restrictions implied by intersection are very similar to the ones in (48). Thus, the use of managed returns is similar to the coefficients in the spanning regression being linear in the instruments.[3]

Spanning for a given value of z_t is equivalent to

$$\begin{aligned} a_{i0} + z_t' a_{i1} &= 0, \\ (b_{i0} + z_t' b_{i1}) \iota_K &= 1. \end{aligned} \tag{49}$$

Again, for a specific value of z_t, i.e., for specific economic conditions, these restrictions can easily be tested in the regression framework outlined above. If there is to be spanning under all economic conditions the restrictions are

$$\begin{aligned} a_{10} &= 0, \\ b_{10} \iota_K &= 1, \\ a_{i1} &= 0, \\ b_{i1} &= 0. \end{aligned}$$

[3] We thank the referee for pointing this out to us.

If there are L instruments (including a constant) with K benchmark assets and N new assets, we now have $(K+1) \times N \times L$ restrictions to test, which is even larger than with the scaled returns in Section 4.1. Also, the numbers of parameters to be estimated is $(K+1) \times N \times L$. Thus, in terms of the number of parameters and the number of restrictions, this approach does not offer additional benefits over the use of scaled returns. However, this approach does have the benefit that it shows under what economic circumstances there may or may not be intersection or spanning.

Notice that this way of incorporating conditional information is very similar to the one suggested in the previous section. The restrictions on the regression parameters in (46) are analogous to the ones on the parameters in (41). The main difference arises because the slope coefficients for R_{t+1} also depend on the instruments, implying that the interaction term $z_t R_{t+1}$ should also be included in the regression. It is easy to see that the approach in the previous section can be interpreted as a special case of the approach outlined here, where only the intercepts in (20) are a function of the instruments z_t, whereas the slope coefficients are constant.

Summarizing, we have shown that a number of approaches is available to incorporate conditioning information in tests for intersection and spanning. Using either scaled returns or regression coefficients that are linear functions of the instruments, the regression approach outlined in Section 3 can easily be extended to test for intersection or spanning. The restrictions implied by the hypotheses of intersection and spanning are very similar to the case where there is no conditioning information (i.e., where the only instrument is a constant) and have very similar interpretations as well. Our methods focus on specific functional forms of incorporating conditioning information.

5 The relation between spanning tests, performance evaluation and optimal portfolio weights

So far the focus has been on the restrictions that are implied by the hypotheses of intersection and spanning on the distribution of R_{t+1} and r_{t+1} and on tests of these hypotheses. In this section the interest will be in the deviations from the restrictions. We will show that the test statistics and regression estimates have clear interpretations in terms of performance measures like Jensen's alpha and the Sharpe ratio as well as in terms of the new optimal portfolio weights. Since it is natural to think about these performance measures in terms of mean-variance efficient portfolios, most of the analysis in this section will be in terms of mean-variance frontiers rather than volatility bounds. Nonetheless, the duality between these two frontiers also holds for these performance measures. These interpretations of tests for mean-variance efficiency, intersection, and spanning in terms of performance measures can also be found in Cochrane (1996), Dahlquist and Söderlind (1999), Gibbons, Ross and Shanken (1989), Jobson and Korkie (1982, 1984, 1989), and Kandel and Stambaugh (1989).

5.1 Performance measures

To set the stage, define the vector of Jensen's alphas, or Jensen performance measures, $\alpha_J(\eta)$, as the intercepts in a regression of the N excess returns $(r_{t+1} - \eta\iota_N)$ on the excess returns of the K benchmark assets, $(R_{t+1} - \eta\iota_K)$:

$$r_{t+1} - \eta\iota_N = \alpha_J(\eta) + \beta(R_{t+1} - \eta\iota_K) + \varepsilon_{t+1}, \tag{50}$$

with $E[\varepsilon_{t+1}] = E[\varepsilon_{t+1}R_{t+1}] = 0$. Since it is not assumed that there exists a risk free asset, we define excess returns as the return on an asset or portfolio in excess of a given zero-beta rate η. Alternatively, when regressing r_{t+1} on R_{t+1} as in (20), it follows that Jensen's alpha is equal to

$$\alpha_J(\eta) = \alpha + (\beta\iota_K - \iota_N)\eta, \tag{51}$$

where $\alpha = \mu_r - \beta\mu_R$ and $\beta = \Sigma_{rR}\Sigma_{RR}^{-1}$. Notice from this expression that the hypothesis that there is intersection for a given value of η is equivalent to the hypothesis that the Jensen performance measure is zero, i.e., $\alpha_J(\eta) = 0$. Similarly, the hypothesis of spanning is equivalent to the hypothesis that $\alpha_J(\eta) = 0$, $\forall\eta$. Recall from Section 3.3, that the regression in (50) produces the same intercept $\alpha_J(\eta)$ as a regression of $r_{t+1} - \eta\iota_N$ on the excess return of a portfolio w_R^* that is mean-variance efficient for R_{t+1} and that has η as its zero beta rate, i.e.,

$$r_{t+1} - \eta\iota_N = \alpha_J(\eta) + \beta^*(R_{t+1}^* - \eta) + \varepsilon_{t+1}.$$

Following Jensen (1968), it is common in the literature to define Jensen's alpha as the intercept of a regression of r_{t+1} in excess of the risk free rate on the return of the market portfolio in excess of the risk free rate. The definition in (50) is more general and has this more traditional definition as a special case if there exists a risk free asset ($\eta = R_t^f$) and if the market portfolio is mean-variance efficient ($R_{t+1}^{*} = R_{t+1}^{m}$). The Jensen measure in (50) is also referred to as the generalized Jensen measure. Given the minimum variance stochastic discount factor $m_R(v)_{t+1}$ as defined in (4) and (5), it can easily be seen that the generalized Jensen measure is also equal to $\lambda(v)/v$ as defined in (28). This is also discussed in Cochrane (1996) and in Dahlquist and Söderlind (1999).

The Sharpe ratio of a portfolio with return R_{t+1}^p is defined as the expected excess portfolio return, divided by the standard deviation of portfolio return,

$$Sh(R_{t+1}^p;\eta) \equiv \frac{E[R_{t+1}^p] - \eta}{\sigma(R_{t+1}^p)}.$$

By definition, for a given expected portfolio return, or for a given standard deviation of portfolio return, the maximum attainable (absolute) Sharpe ratio is the Sharpe ratio of the minimum-variance efficient portfolio. For a minimum-variance efficient portfolio w_R^* of the K assets R_{t+1} with zero-beta rate η, the Sharpe ratio is equal to the slope of the line tangent to the frontier originating at $(0;\eta)$ in mean-standard deviation space, and is denoted by $\theta_R(\eta)$:

$$\theta_R(\eta) = \frac{E[R_{t+1}^*] - \eta}{\sigma(R_{t+1}^*)}; \tag{52}$$

where $R_{t+1}^* \equiv w^{*\prime}R_{t+1}$.

Although both Jensen's alpha and the Sharpe ratio are used as performance measures, there is an important difference between the two. Whereas the Sharpe ratio is defined in terms of the characteristics of one portfolio (the expected excess portfolio return and its standard deviation), Jensen's alpha is defined in terms of one asset or portfolio relative to another. Sharpe ratios answer the question whether one portfolio is to be preferred over another, whereas Jensen's alpha answers the question whether investors can improve the efficiency of their portfolio by investing in the new asset. However, there is a close relation between the two measures, in that Jensen's alphas together with the covariance matrix of the error terms ε_{t+1} in (20) (and (50)) determine the potential improvement in the maximum attainable Sharpe ratio from adding the new assets r_{t+1}. Recall from Section 2.2 that we defined the variables $A \equiv \mu'\Sigma^{-1}\mu$, $B \equiv \iota'\Sigma^{-1}\mu$, and $C \equiv \iota'\Sigma^{-1}\iota$. For the set R_{t+1} these variables will be denoted as A_R, B_R, and C_R, whereas the absence of subscripts implies that these variables refer to the larger set (R_{t+1}, r_{t+1}). Using partitioned inverses, notice that

$$\Sigma^{-1} = \begin{pmatrix} \Sigma_{RR} & \Sigma_{Rr} \\ \Sigma_{rR} & \Sigma_{rr} \end{pmatrix}^{-1} = \begin{pmatrix} \Sigma_{RR}^{-1} + \beta' \Sigma_{\varepsilon\varepsilon}^{-1} \beta & -\beta' \Sigma_{\varepsilon\varepsilon}^{-1} \\ -\Sigma_{\varepsilon\varepsilon}^{-1} \beta & \Sigma_{\varepsilon\varepsilon}^{-1} \end{pmatrix}. \tag{53}$$

From this, it follows that

$$\begin{aligned} A &= \iota_K' \Sigma_{RR}^{-1} \iota_K + \iota_K' \beta' \Sigma_{\varepsilon\varepsilon}^{-1} \beta \iota_K - 2 \iota_K' \beta' \Sigma_{\varepsilon\varepsilon}^{-1} \iota_N + \iota_N' \Sigma_{\varepsilon\varepsilon}^{-1} \iota_N \\ &= A_R + (\beta \iota_K - \iota_N)' \Sigma_{\varepsilon\varepsilon}^{-1} (\beta \iota_K - \iota_N), \end{aligned} \tag{54}$$

where $\beta = \Sigma_{rR} \Sigma_{RR}^{-1}$ and $\Sigma_{\varepsilon\varepsilon}$ is the covariance matrix of ε_{t+1}, the error term in the regression in (20). In a similar way it can easily be shown that

$$B = B_R + \alpha' \Sigma_{\varepsilon\varepsilon}^{-1} (\iota_N - \beta \iota_K), \tag{55a}$$

$$C = C_R + \alpha' \Sigma_{\varepsilon\varepsilon}^{-1} \alpha, \tag{55b}$$

where $\alpha = \mu_r - \beta \mu_R$, the intercept in the regression in (20).

It is easy to show that for a given η, the Sharpe ratio of a mean-variance efficient portfolio w_R^* can be written as

$$\theta_R(\eta) = (C_R - 2B_R \eta + A_R \eta^2)^{1/2}. \tag{56}$$

A similar expression holds of course for $\theta(\eta)$, the maximum attainable Sharpe ratio of the larger set (R_{t+1}, r_{t+1}). Using (54) and (55), we derive

$$\begin{aligned} \theta(\eta)^2 &= C - 2B\eta + A\eta^2 \\ &= (C_R - 2B_R \eta + A_R \eta^2) \\ &\quad + (\alpha' \Sigma_{\varepsilon\varepsilon}^{-1} \alpha - 2\alpha' \Sigma_{\varepsilon\varepsilon}^{-1} (\iota_N - \beta \iota_K)\eta + (\iota_N - \beta \iota_K)' \Sigma_{\varepsilon\varepsilon}^{-1} (\iota_N - \beta \iota_K) \eta^2) \\ &= \theta_R(\eta)^2 + \alpha_J(\eta)' \Sigma_{\varepsilon\varepsilon}^{-1} \alpha_J(\eta). \end{aligned} \tag{57}$$

Thus, the change in maximum attainable squared Sharpe ratios equals the inner product of the vector of Jensen's alphas, $\alpha_J(\eta)$, weighted by the inverse of the covariance matrix of ε_{t+1}.[4] If there is only one new asset, $N = 1$, the term $\alpha_J(\eta)/\sigma(\varepsilon)$ is known as the adjusted Jensen measure or the appraisal ratio (Treynor and Black (1973)). Notice once more that $\theta_R(\eta)$ and $\theta(\eta)$

[4] This result can be found in Jobson and Korkie (1984) for instance.

characterize portfolios of R_{t+1} and $(R'_{t+1}, r'_{t+1})'$, respectively, whereas $\alpha_J(\eta)$ and $\Sigma_{\varepsilon\varepsilon}$ follow from a regression of r_{t+1} on R_{t+1}, and measure the performance of r_{t+1} relative to R_{t+1}. Stated differently, whereas Sharpe ratios can be used to compare the performance of different portfolios, Jensen's alpha gives the potential improvement in performance when the additional assets are included in the portfolio. The hypotheses of intersection and spanning imply that Jensen's alpha, $\alpha_J(\eta)$, is zero for one or for all values of η respectively. Therefore, if there is intersection (spanning) then there is no improvement in the Sharpe measure possible by including the additional assets r_{t+1} in the investors portfolio.

Cochrane and Saá-Requejo (1996) show how a bound on the maximum Sharpe ratio can be used to price new assets in incomplete markets, which is referred to as "good deal" pricing. In the context of (57) this essentially comes down to putting a bound on the maximum appraisal ratios of the new asset. This kind of analysis is extended by Bernardo and Ledoit (1996), who introduce the gain-loss ratio as an alternative performance measure by which new assets can be priced if restrictions on the maximum gain-loss ratio are imposed. This is similar to a bound on the maximum Sharpe ratio as suggested by Cochrane and Saá-Requejo (1996), but the approach in Bernardo and Ledoit (1996) is more general and allows for non-mean variance utility functions as well.

5.2 Changes in optimal portfolio weights

The performance measures and the intersection regressions discussed above can also be used to infer the changes in optimal portfolio holdings when adding the assets r_{t+1}. In this section we will show that given the initial mean-variance efficient portfolio of the benchmark assets and the OLS-estimates of the regression parameters in (20), it is straightforward to determine the new optimal portfolio weights. Some of the results presented in this section are also presented in Stevens (1998). In order to derive the optimal portfolio weights from the regression results, consider the mean-variance efficient portfolio for the extended set (R_{t+1}, r_{t+1}) for a given value of η:

$$w^* = \gamma^{-1}\Sigma^{-1}(\mu - \eta\iota).$$

Substituting the partitioned inverse as given in (53) in the expression for w^* gives that the optimal portfolio weights for the new assets, w^*_r, can be written as

$$\begin{aligned} w_r^* &= \gamma^{-1}\Sigma_{\varepsilon\varepsilon}^{-1}((\mu_r - \beta\mu_R) - (\iota_N - \beta\iota_K)\eta) \\ &= \gamma^{-1}\Sigma_{\varepsilon\varepsilon}^{-1}\alpha_J(\eta). \end{aligned} \tag{58}$$

Thus, the optimal portfolio weights w_r^* are determined by the vector of Jensen's alphas and the covariance matrix of the residuals of the OLS-regression of r_{t+1} on R_{t+1}.[5] This result is simply a generalization of the well known result in Treynor and Black (1973) regarding the appraisal ratio. The difference with Treynor and Black is that these authors assume that the error terms $\varepsilon_{i;t+1}$ for different securities are uncorrelated, i.e., they assume the diagonal model (Sharpe (1963)), whereas the result in (58) allows for any correlation structure between the securities.

In deriving the new optimal portfolio weights, a problem in (58) is that the coefficient of risk aversion γ is present. Notice that this is a different coefficient than the one that appears in the optimal portfolio $\tilde{w}_R^*$ of the smaller set R_{t+1}:

$$\tilde{w}_R^* = \tilde{\gamma}_R^{-1}\Sigma_{RR}^{-1}(\mu_R - \eta\iota_K);$$

where we now also add a ~ to indicate that a variable refers to the set of benchmark assets R_{t+1} only. It is only the zero-beta return η that is the same in both problems, since we test whether there is intersection for a fixed value of η. Similarly, the expected returns on the portfolios $\tilde{w}_R^*$ and w^* are different, and we indicate these with $\tilde{m}_R$ and m respectively, i.e., $\tilde{m}_R \equiv \tilde{w}_R^{*\prime}\mu_R$, and $m \equiv w^{*\prime}\mu$. In order to substitute out the risk aversion parameter γ, note that

$$\begin{aligned} \gamma &= B - \eta A = B_R - \eta A_R + \alpha_J(\eta)'\Sigma_{\varepsilon\varepsilon}^{-1}(\iota_N - \beta\iota_K) \\ &= \tilde{\gamma}_R + \alpha_J(\eta)'\Sigma_{\varepsilon\varepsilon}^{-1}(\iota_N - \beta\iota_K); \end{aligned}$$

and that

$$\tilde{\gamma}_R = \frac{\tilde{m}_R - \eta}{\tilde{w}_R^{*\prime}\Sigma_{RR}\tilde{w}_R^*} = \frac{\theta_R(\eta)^2}{\tilde{m}_R - \eta}.$$

[5] As an aside, in terms of volatility bounds, notice that $w_r^*\gamma = -\varphi_r(1/\eta)$, i.e., the elements of $\varphi(v)$ in (5) that correspond to r_{t+1}. Thus if we want to know the minimum variance stochastic discount factor from $(R_{t+1}; r_{t+1})$, rather than from R_{t+1}, the projection coefficients corresponding to the additional assets r_{t+1} are given by $-\Sigma_{\varepsilon\varepsilon}^{-1}\alpha_J(\eta)$.

Using these latter two expressions, the optimal portfolio weights w_r^* can be expressed as

$$w_r^* = \left(\frac{\tilde{m}_R - \eta}{\theta_R(\eta)^2 + (\tilde{m}_R - \eta)\alpha_J(\eta)'\Sigma_{\varepsilon}^{-1}(\iota_N - \beta\iota_K)} \right) \Sigma_{\varepsilon}^{-1}\alpha_J(\eta). \tag{59}$$

The advantage of (59) is that it contains only variables that either result from the initial optimal portfolio $\tilde{w}_R^*$, or from a regression of r_{t+1} on R_{t+1}.

Along the same lines it is straightforward to show that the new optimal weights w_R^* are given by

$$w_R^* = \left(\frac{\theta_R(\eta)^2}{\theta_R(\eta)^2 + (\tilde{m}_R - \eta)\alpha_J(\eta)'\Sigma_{\varepsilon}^{-1}(\iota_N - \beta\iota_K)} \right) \tilde{w}_R^* - \beta' w_r^*. \tag{60}$$

Again, this expression only depends on characteristics of the old portfolio, $\tilde{w}_R^*$, and the regression output. Therefore, given the initial mean-variance efficient portfolio $\tilde{w}_R^*$ of the benchmark assets and the OLS-estimates of the regression in (20), equations (59) and (60) answer the question how to adjust the portfolio in order to obtain the new mean-variance efficient portfolio w^*.

In order to give an interpretation of the new portfolio weights in (59) and (60), it is useful to rewrite them in the following way:[6]

$$w_r^* = \frac{m - \eta}{\theta(\eta)^2}\Sigma_{\varepsilon}^{-1}\alpha_J(\eta), \tag{61}$$

and

$$w_R^* = \frac{\theta_R(\eta)^2}{\theta(\eta)^2}\frac{m - \eta}{\tilde{m}_R - \eta}\tilde{w}_R - \beta' w_r^*. \tag{62}$$

If there is only one new asset, i.e., $N = 1$, Equation (61) first of all shows that $\alpha_J(\eta)$ determines the sign of the new portfolio weight w_r^* (given that $m - \eta > 0$): if Jensen's alpha is positive (negative) the investor can improve the performance of his portfolio by taking long (short) positions in the new asset. When there is more than one new asset, the sign of the portfolio weights is not only determined by the sign of Jensen's alpha, but also by the inverse of the covariance matrix of ε_{t+1}. If the mean-variance frontier is not strongly affected by the introduction of the new assets, then

[6] Here we use the fact that $\theta_R(\eta)^2/(\tilde{m}_R - \eta) = A_R - \eta B_R$, and that $A_R - \eta B_R + \alpha_J(\eta)'\Sigma_{\varepsilon}^{-1}(\iota_N - \beta\iota_K) = A - \eta B$.

$(\sigma_R(\eta)^2/\sigma(\eta)^2)(m-\eta)/(\mu_R-\eta) \approx 1$, and the coefficients β show which of the old assets are replaced by the new assets.

Finally, notice that we did not consider a risk free asset. The portfolio weights given above are for the tangency portfolio when the zero-beta rate is η. If a risk free asset is available, we can replace η with R^f in (61) and (62) and these equations still give the portfolio weights for the tangency portfolio. The new tangency portfolio has an expected return equal to m, whereas the old tangency portfolio has an expected return μ_R. Notice though, that in case a risk free asset is available it is easy to shift funds between the tangency portfolio and the risk free asset and to let the expected portfolio return vary. For practical purposes, the interest may be in the new portfolio w^* that has the same expected return as the old portfolio. Given that there is a risk free asset available, this is easily achieved by letting $m - R^f = \mu_R - R^f$. In this case Equations (61) and (62) simplify to

$$w_r^* = \frac{m - R^f}{\sigma^2}\Sigma_{\varepsilon\varepsilon}^{-1}\alpha_J \tag{63}$$

and

$$w_R^* = \frac{\sigma_R^2}{\sigma^2}\mathbf{w}_R - \beta' w_r^*. \tag{64}$$

Notice that here it is not necessarily the case that the weights in w_r^* and w_R^* sum to one. The investor will have to borrow or lend a fraction $(1 - \iota_K' w_R^* - \iota_N' w_r^*)$ to achieve an expected portfolio return equal to m.

5.3 Interpretation of spanning and intersection tests in terms of performance measures

Finally, we want to relate the Wald test-statistics presented in Section 3 to the performance measures discussed above. It will be shown that these test-statistics can be expressed as changes in maximum Sharpe ratios of R_{t+1} and (R_{t+1}, r_{t+1}) respectively. Therefore, they have a clear economic interpretation. In order to interpret the test-statistics for intersection and spanning in terms of performance measures, recall the basic regression model in (20):

$$r_{t+1} = \alpha + \beta R_{t+1} + \varepsilon_{t+1},$$

where intersection for a given value of η means that

$$\alpha_J(\eta) = \alpha + (\beta\iota_K - \iota_N)\eta = 0.$$

Thus, the restrictions on the regression coefficients that are imposed by the hypothesis of intersection have a natural interpretation in terms of Jensen's alphas, and - as noted before - testing whether there is intersection for η, is equivalent to testing whether Jensen's alpha is zero. Testing for spanning is of course equivalent to testing whether the Jensen's alphas are zero for all values of η.

It can be shown that the test statistics for intersection and spanning, ξ_W^{int} and ξ_W^{span}, presented in Section 3.4, can also be interpreted in terms of Jensen's alphas and Sharpe ratios. To see this, start again from the specification of the regression equation in (23):

$$r_{t+1} = \left(I_N \otimes \left(1 \quad R_{t+1}'\right)\right) b + \varepsilon_{t+1}.$$

Note that (using partitioned inverses) the asymptotic covariance matrix of the OLS-estimates of b, $\hat{b}$ in (23) is given by

$$\begin{aligned} &\Sigma_{\varepsilon\varepsilon} \otimes \begin{pmatrix} 1 & \mu_R' \\ \mu_R & E[R_t R_t'] \end{pmatrix}^{-1} \\ = \quad &\Sigma_{\varepsilon\varepsilon} \otimes \begin{pmatrix} 1 + \mu_R' \Sigma_{RR}^{-1} \mu_R & -\mu_R' \Sigma_{RR}^{-1} \\ -\Sigma_{RR}^{-1} \mu_R & \Sigma_{RR}^{-1} \end{pmatrix}. \end{aligned} \tag{65}$$

Straightforward algebra shows that premultiplying (65) with $H(\eta)_{int}$ and postmultiplying with $H(\eta)_{int}'$ as defined in (25), yields

$$Var[\hat{\alpha}_J(\eta)] = \Sigma_{\varepsilon\varepsilon}(1 + \theta_R(\eta)^2), \tag{66}$$

where the Sharpe ratio $\theta_R(\eta)$ was defined in (56). Since from the analysis above we know that the term $h(\eta)_{int}$ as defined in (25) equals $\alpha_J(\eta)$, (57) can be used to rewrite the test statistic for intersection, ξ_W^{int}, as

$$\xi_W^{int} = T \frac{\hat{\alpha}_J(\eta)' \hat{\Sigma}_{\varepsilon\varepsilon}^{-1} \hat{\alpha}_J(\eta)}{1 + \hat{\theta}_R(\eta)^2} = T \left(\frac{1 + \hat{\theta}(\eta)^2}{1 + \hat{\theta}_R(\eta)^2} - 1 \right), \tag{67}$$

where $\hat{\theta}_R(\eta)$, $\hat{\theta}(\eta)$, and $\hat{\alpha}_J(\eta)$ are the sample Sharpe ratios and Jensen's alpha respectively. Equation (67) is a well known result from, e.g., Jobson and Korkie (1982) and Gibbons, Ross and Shanken (1989). It clearly shows that the Wald test statistic for intersection can easily be interpreted as the percentage increase in squared Sharpe ratios scaled by the sample size. Under the null-hypothesis that there is intersection, $\theta(\eta) = \theta_R(\eta)$ and the increase of the sample Sharpe ratios scaled by the sample size T (as in (67)) will asymptotically have a $\chi^2_{(N)}$-distribution.[7]

MacKinlay (1995) uses a similar interpretation of the Wald test-statistic in case returns are normally distributed together with (57) to distinguish between risk-based alternatives for the CAPM and nonrisk-based alternatives. His analysis suggests that for reasonable values of the maximum attainable Sharpe ratios a multifactor model like the one proposed by Fama and French (1996) can not explain the deviations from the CAPM that are found in the cross section of asset returns.

For the spanning test-statistic, a similar interpretation can be given. Let μ_R^0 denote the expected return on the global minimum variance portfolio of R_{t+1}, i.e., $\mu_R^0 = B_R/A_R$, and let the variance of this portfolio be given by $(\sigma_R^0)^2$. Similarly, let $(\sigma^0)^2$ be the global minimum variance of $(R_{t+1}; r_{t+1})$. It is shown in Appendix B that the Wald test-statistic for spanning, ξ_W^{span}, can be written as

$$\xi_W^{span} = T\left(\frac{1+\hat{\theta}^2(\hat{\mu}_R^0)}{1+\hat{\theta}_R^2(\hat{\mu}_R^0)} - 1\right) + T\left(\frac{(\hat{\sigma}_R^0)^2}{(\hat{\sigma}^0)^2} - 1\right). \qquad (68)$$

This shows that the spanning test-statistic consists of two parts. The first part is similar to the test-statistic for intersection in (67) and is determined by a change in Sharpe ratios. The Sharpe ratios in (68) are for a zero-beta rate equal to the (in-sample) expected return on the global minimum variance portfolio however, and therefore are the slopes of the asymptotes of the mean-variance frontier. Notice that the slope of the upper limb of the frontier is simply the negative of the slope of the lower limb of the frontier, and therefore, the squared Sharpe ratios for those two extremes are the same. The first term of the spanning test-statistic in a sense measures whether there is intersection at the most extreme points of the frontier (i.e., whether there is a limiting form of intersection if we go sufficiently far up or down the frontier). The second term of the statistic in (68) is determined by the change in the global minimum variance of the portfolios, and measures whether the point most to the left on the frontier changes or not. Put differently, the first term measures whether there is intersection for a mean-variance investor

[7] Gibbons, Ross, and Shanken (1989) study the small sample properties of this test statistic in case there is a risk free asset, as well as the distribution under the alternative hypothesis. Kandel and Stambaugh (1987) and Shanken (1987) extend their results to the case where the mean-variance efficient benchmark portfolio (or intersection portfolio) can not be observed but has a given correlation with the observed proxy portfolio.

with a very small risk aversion ($\gamma = 0$), while the second term measures whether there is intersection for a mean-variance investor with a very high risk aversion ($\gamma \to \infty$). Note that in the second term the old global minimum variance appears in the numerator and the new global minimum variance in the denominator, since this variance can only decrease as assets are added to the portfolio. Therefore, both terms in (68) are always larger than or equal to one. Jobson and Korkie (1989) derive a similar result for a likelihood ratio test for spanning.

6 Specification error bounds and intersection

As in the previous section, in this section the focus will be on deviations from intersection rather than on intersection itself. In a recent paper Hansen and Jagannathan (1997) analyze specification errors in stochastic discount factor models which, in some special cases, can be interpreted as deviations from intersection. They derive bounds on the magnitude of these specification errors.

Recall from the discussion in Section 2.1 that each asset pricing model assigns a particular function to the pricing kernel M_{t+1}. Hansen and Jagannathan (1997) note that the pricing kernels implied by most asset pricing models do not yield correct asset prices, either because the asset pricing model can only be viewed as an approximation, or because of measurement error. Measurement errors are for instance often considered to be an important problem in measuring consumption and testing consumption based asset pricing models. Therefore, the pricing kernel implied by an asset pricing model will in general only serve as a proxy stochastic discount factor, that will not yield the correct prices or expected payoffs of the assets under consideration. In a related paper Balduzzi and Robotti (2000) focus on the estimation of risk premia as a separate problem from the testing of asset pricing models. They estimate risk premia by looking at the prices assigned by the minimum variance kernel to risk variables, or by the prices of hedge portfolios that are the linear projections of risk variables on asset returns.

The interest of Hansen and Jagannathan is in the least squares distance between a proxy stochastic discount factor and the set of valid stochastic discount factors. They derive a lower bound on this distance, the specification error bound, as a measure of how well the model performs. These specification error bounds will be derived formally below and it will also be shown that these bounds have a clear economic interpretation in terms of maximum pricing errors or maximum expected payoff errors implied by the asset pricing model. Hansen, Heaton, and Luttmer (1995) derive the limiting distribution for the corresponding estimator of the specification error bounds.

It turns out that if we take the minimum variance stochastic discount factor for the subset R_{t+1} as a proxy stochastic discount factor for the larger set of assets $(R_{t+1}; r_{t+1})$, we can interpret the specification error bounds in terms of mean-variance intersection and the performance measures discussed in the previous section. In particular, provided that both the proxy stochastic discount factor and the discount factors that price R_{t+1} and r_{t+1} correctly have the same expectation v, the squared specification error bound scaled by v turns out to be equal to the difference between the maximum squared Sharpe ratio implied by the set R_{t+1} and the maximum squared Sharpe ratio implied by $(R_{t+1}; r_{t+1})$. This also allows us to interpret the specification errors in terms of mean-variance portfolio choice again. Given that a mean-variance investor is aware of the fact that a portfolio chosen from the subset R_{t+1} is suboptimal relative to a portfolio chosen from the larger set $(R_{t+1}; r_{t+1})$, the specification error bound gives an estimate of the extent to which the portfolio is suboptimal in terms of Sharpe ratios.

6.1 Specification error bounds

As noted above, in Hansen and Jagannathan (1997) the interest is in proxy stochastic discount factors, denoted by y_{t+1}, that assign approximate prices to portfolio payoffs. For instance, the CAPM implies that the proxy is of the form $a + bR^m_{t+1}$, with R^m_{t+1} the return on the market portfolio. As before, let R^p_{t+1} be the return on some portfolio, not necessarily mean-variance efficient, such that $w^{p\prime}\iota_K = 1$. The expected price assigned to such a portfolio by a proxy stochastic discount factor will be denoted by $\pi^a(R^p_{t+1})$:

$$E[y_{t+1}R^p_{t+1}] = \pi^a(R^p_{t+1}). \tag{69}$$

Of course, valid stochastic discount factors M_{t+1} would assign a price $\pi(R^p_{t+1}) = 1$ to any portfolio w^p that satisfies $w^{p\prime}\iota_K = 1$. Because the proxy y_{t+1} may be derived from an asset pricing model that is strictly speaking not valid, or because the proxy may be measured with error, the prices assigned by the proxy, $\pi^a(R^p_{t+1})$, will in general not be equal to one. We only consider payoffs that are returns, i.e., payoffs with (correct) prices equal to one. Hansen and

Jagannathan (1997) take more general payoffs x_{t+1} with current prices q_t. Given that we want to establish the relation between specification errors and mean-variance intersection, the use of returns is not very restrictive however. Moreover, the results derived below can easily be adjusted to the results of Hansen and Jagannathan along the lines of Section 4.1, where we incorporated conditioning information by allowing for payoffs $z_t \otimes R_{t+1}$ with current prices q_t.

A second way in which the results here are somewhat more restrictive than the ones in Hansen and Jagannathan (1997) is that we will always consider valid stochastic discount factors $M(v)_{t+1}$ that have the same expectation as the proxy y_{t+1}, i.e., $v = E[y_{t+1}]$. This may be considered as restrictive, since this assumption in fact requires that the proxy assigns the correct price to the risk free payoff, if it exists. Once more, given that the interest here is in the relation with mean-variance intersection in the absence of a risk free asset, and given that we always defined intersection for a known value of v, this is not restrictive for our purposes.

The problem addressed in Hansen and Jagannathan (1997) is to derive a lower bound δ on the distance between y_{t+1} and the set of stochastic discount factors that price R_{t+1} correctly, which we denote as $\mathcal{M}$:

$$\delta = \min_{\{M_R(v)_{t+1} \in \mathcal{M}\}} \| y_{t+1} - M_R(v)_{t+1} \| ; \tag{70}$$

where $\| x_{t+1} \| \equiv E[x_{t+1}^2]^{1/2}$. Because y_{t+1} and $M_R(v)_{t+1}$ have the same expectation, the distance between y_{t+1} and $M_R(v)_{t+1}$ in (70) is equal to the standard deviation of $y_{t+1} - M_R(v)_{t+1}$, i.e., $\| y_{t+1} - M_R(v)_{t+1} \| = \sigma(y_{t+1} - M_R(v)_{t+1})$. We will denote the stochastic discount factor that solves (70) by $\tilde{M}_R(v)_{t+1}$. Thus, $\tilde{M}_R(v)_{t+1}$ is the stochastic discount factor that prices R_{t+1} correctly and that is closest to y_{t+1} in a least squares sense.

To solve the problem in (70), consider the least squares projections of y_{t+1} and $M_R(v)_{t+1}$ on R_{t+1} and a constant:

$$\begin{aligned} \hat{y}_{t+1} &= \mathrm{Proj}(y_{t+1} \mid 1; R_{t+1}) = v + \zeta(v)'(R_{t+1} - \mu_R); \\ y_{t+1} &= \hat{y}_{t+1} + u_{t+1}; \end{aligned} \tag{71}$$

and

$$\begin{aligned} m_R(v)_{t+1} &= \mathrm{Proj}(M_R(v)_{t+1} \mid 1; R_{t+1}) = v + \phi(v)'(R_{t+1} - \mu_R); \\ M_R(v)_{t+1} &= m_R(v)_{t+1} + w_{t+1}; \end{aligned} \tag{72}$$

where $m_R(v)_{t+1}$ is the minimum variance stochastic discount factor derived in Section 2.1, and $\varphi(v)$ is defined in (5). The projection coefficients in (71) are given by $\Sigma_{RR}^{-1}\Sigma_{Ry}$, with Σ_{Ry} the $K \times 1$-vector of covariances between R_{t+1} and y_{t+1}. Noting that $\| y_{t+1} - M_R(v)_{t+1} \|^2 = Var[y_{t+1} - M_R(v)_{t+1}]$, it easily follows that

$$
\begin{aligned}
Var[y_{t+1} - M_R(v)_{t+1}] &= Var[\tilde{y}_{t+1} - m_R(v)_{t+1}] + Var[u_{t+1} - w_{t+1}] \\
&\geq Var[\tilde{y}_{t+1} - m_R(v)_{t+1}]:
\end{aligned}
$$

Because $\tilde{y}_{t+1} - m_R(v)_{t+1} = y_{t+1} - (m_R(v)_{t+1} + u_{t+1})$ and u_{t+1} is orthogonal to R_{t+1}, this lower bound on the variance of $y_{t+1} - M_R(v)_{t+1}$ is attainable for the stochastic discount factor

$$
\tilde{m}_R(v)_{t+1} = m_R(v)_{t+1} + u_{t+1}; \tag{73}
$$

and we have that

$$
\delta^2 = Var[y_{t+1} - \tilde{m}_R(v)_{t+1}]: \tag{74}
$$

A more detailed characterization of $\tilde{m}_R(v)_{t+1}$ and δ will be given in the following section. For this moment, note that subtracting the variable $y_{t+1} - \tilde{m}_R(v)_{t+1}$ from the proxy y_{t+1} yields a valid stochastic discount factor. Therefore, as noted by Hansen and Jagannathan (1997), $y_{t+1} - \tilde{m}_R(v)_{t+1}$ is the smallest adjustment in a least squares sense that is necessary to make y_{t+1} a valid stochastic discount factor, and δ is a measure of the magnitude of this adjustment.

Hansen and Jagannathan also show that δ can be interpreted as a maximum pricing error. In order to do so, let ω denote a position in R_{t+1} that does not necessarily satisfy the requirement $\omega'\iota_K = 1$, i.e., ω is in general not a portfolio. Denote the payoff of such a position as $R(\omega)_{t+1} = \omega' R_{t+1}$ and note that the correct price of such a position is

$$
E[\omega' R_{t+1} M_R(v)] = \pi(R(\omega)_{t+1}) = \omega'\iota_K;
$$

whereas the price assigned by the proxy y_{t+1} is $\pi^a(R(\omega)_{t+1})$. The pricing error of the proxy y_{t+1} is therefore $\pi^a(R(\omega)_{t+1}) - \pi(R(\omega)_{t+1})$, and Hansen and Jagannathan show that δ provides an upper bound on the absolute value of this pricing error, for positions that have a unit norm:

$$
\delta = \max_{R(\omega)_{t+1};\|R(\omega)_{t+1}\|=1} | \pi^a(R(\omega)_{t+1}) - \pi(R(\omega)_{t+1}) | :
$$

Thus, by looking at a particular class of positions, i.e., positions with a unit norm, δ can be interpreted as the maximum pricing error assigned by the proxy to the payoffs of those unit norm positions.

A more intuitive interpretation can be given if we consider errors in expected payoffs, or expected returns, rather than pricing errors. Recall that a valid stochastic discount factor assigns the correct expected return to a one-dollar investment in portfolio w^p (for which, by definition, $w^{p\prime}\iota = 1$) which, using equation (3), can be written as

$$E[R^p_{t+1}] = \frac{1}{v} - \frac{Cov[M_R(v)_{t+1}; R^p_{t+1}]}{v},$$

i.e., as one over the expectation of the pricing kernel, which equals the risk free rate if it exists, plus a risk term which is determined by the covariance of the portfolio return and the pricing kernel. Observe that use of the proxy, that also has expectation v, would give an approximate expected return $E^a[R^p_{t+1}]$ for a one-dollar investment in w^p that in general differs from $E[R^p_{t+1}]$, because the covariance of the proxy with the portfolio return will be different from the covariance of a valid stochastic discount factor with the portfolio return, i.e.:

$$E^a[R^p_{t+1}] = \frac{1}{v} - \frac{Cov[y_{t+1}; R^p_{t+1}]}{v}.$$

From these relations we define the expected return error

$$E^a[R^p_{t+1}] - E[R^p_{t+1}] = \frac{Cov[M_R(v)_{t+1} - y_{t+1}; R^p_{t+1}]}{v}, \tag{75}$$

for which the Cauchy-Schwarz inequality implies that

$$| E^a[R^p_{t+1}] - E[R^p_{t+1}] | \leq \frac{\sigma(y_{t+1} - M_R(v)_{t+1})\sigma(R^p_{t+1})}{v}.$$

Since this inequality holds for all valid stochastic discount factors $M_R(v)_{t+1}$, it also holds for the stochastic discount factor that solves (70), $\tilde{M}_R(v)_{t+1}$, implying

$$| E^a[R^p_{t+1}] - E[R^p_{t+1}] | \leq \frac{\delta\sigma(R^p_{t+1})}{v}.$$

Since for a given value of v, the Sharpe ratio is defined as $Sh(R^p_{t+1}) \equiv (E[R^p_{t+1}] - 1/v)/\sigma(R^p_{t+1})$, and the approximate Sharpe ratio, i.e., the Sharpe ratio according to the proxy y_{t+1}, as $Sh^a(R^p_{t+1}) \equiv (E^a[R^p_{t+1}] - 1/v)/\sigma(R^p_{t+1})$, this can be rewritten as

$$| Sh^a(R^p_{t+1}) - Sh(R^p_{t+1}) | \le \frac{\delta}{v}. \tag{76}$$

Thus, using errors in expected returns rather than errors in assigned prices, the specification error bound δ scaled by the expectation of the proxy has a very clear interpretation in terms of Sharpe ratios. For any portfolio w^p formed from the assets in R_{t+1}, the absolute difference between the approximate Sharpe ratio assigned to the portfolio returns by y_{t+1} and the actual Sharpe ratio of the portfolio can never exceed the scaled specification error bound δ/v. This interpretation is also somewhat easier than the one given for the expected payoff error in Hansen and Jagannathan (1997), where they focus on the maximum error in expected payoffs for positions ω with unit standard deviation.

6.2 The relation between specification error bounds and intersection

The purpose of this section is to show that there is a close relation between intersection and a special case of the specification error bounds. In particular, if the interest is in stochastic discount factors that price the returns (R_{t+1}, r_{t+1}) correctly and we choose for the proxy y_{t+1} the minimum variance stochastic discount factor based on the subset R_{t+1}, $m_R(v)_{t+1}$, the specification error bound can simply be expressed as a deviation from intersection, as was the case with the performance measures discussed in Section 5. To show this, let us first give a more precise characterization of $\tilde{m}(v)_{t+1}$ and δ than given in (73) and (74).

Recall that $\tilde{m}_R(v)_{t+1}$ is given by $m_R(v)_{t+1} + u_{t+1}$, where $u_{t+1} = y_{t+1} - \hat{y}_{t+1}$. Using (72) and (71), this implies for $\tilde{m}_R(v)_{t+1}$:

$$\begin{aligned}
\tilde{m}_R(v)_{t+1} &= v + \varphi(v)'(R_{t+1} - \mu_R) + y_{t+1} - \{v + \gamma(v)'(R_{t+1} - \mu_R)\} \\
&= y_{t+1} + (\varphi(v) - \gamma(v))'(R_{t+1} - \mu_R) \\
&= y_{t+1} + \{(\pi_K - v\mu_R) - \Sigma_{Ry}\}'\Sigma_{RR}^{-1}(R_{t+1} - \mu_R);
\end{aligned} \tag{77}$$

and for δ^2:

$$\delta^2 = \{(\pi_K - v\mu_R) - \Sigma_{Ry}\}'\Sigma_{RR}^{-1}\{(\pi_K - v\mu_R) - \Sigma_{Ry}\}. \tag{78}$$

For further reference it is useful to define the vector $\psi(v)$ as

$$\psi(v) = \varphi(v) - \zeta(v) = \Sigma_{RR}^{-1}\{(\iota_K - v\mu_R) - \Sigma_{Ry}\}. \tag{79}$$

Notice that the expressions for $\psi(v)$ and δ^2 given here differ slightly from the ones given in Hansen and Jagannathan (1997) because we explicitly included a constant in the projections of $M(v)_{t+1}$ and y_{t+1} on R_{t+1}.

The expressions for $\tilde{m}_R(v)_{t+1}$ and δ^2 in (77) and (78) provide a basis to relate the specification error bounds to intersection. In case of intersection the interest is in stochastic discount factors that price both R_{t+1} and r_{t+1}, i.e., in $M(v)_{t+1}$. Therefore, in the expressions (77) and (78) we should leave out all the R-subscripts, replace R_{t+1} with the vector $(R'_{t+1}\ r'_{t+1})'$, and note that all vectors and matrices have dimension $K + N$ rather than K. As before, with intersection we want to know if the minimum variance stochastic discount factor based on R_{t+1} only, $m_R(v)_{t+1}$ can be used to price both R_{t+1} and r_{t+1}. In terms of specification errors this means that we want to use $m_R(v)_{t+1}$ as a proxy y_{t+1} for the stochastic discount factors $M(v)_{t+1}$. Also, in the spirit of the previous section, when using $m_R(v)_{t+1}$ as a proxy, we recognize beforehand that $m_R(v)_{t+1}$ will not assign the correct prices to r_{t+1}, but the interest is in the extent to which the assigned prices are wrong, i.e., the extent to which there are deviations from intersection, as measured by δ.

Recall that the proxy $y_{t+1} = m_R(v)_{t+1}$ is now given by

$$\begin{aligned} y_{t+1} &= m_R(v)_{t+1} = v + \varphi_R(v)'(R_{t+1} - \mu_R), \\ \varphi_R(v) &= \Sigma_{RR}^{-1}(\iota_K - v\mu_R). \end{aligned}$$

Substituting these expressions into (77) and (78), properly adjusted for the fact that the interest is now in stochastic discount factors that price both R_{t+1} and r_{t+1}, straightforward algebra shows that

$$\begin{aligned} \delta^2 &= \{(\iota_N - v\mu_r) - \Sigma_{rR}\Sigma_{RR}^{-1}(\iota_K - v\mu_R)\}'\Sigma_{\varepsilon\varepsilon}^{-1}\{(\iota_N - v\mu_r) - \Sigma_{rR}\Sigma_{RR}^{-1}(\iota_K - v\mu_R)\} \\ &= v^2\alpha_J(1/v)'\Sigma_{\varepsilon\varepsilon}^{-1}\alpha_J(1/v), \end{aligned} \tag{80}$$

or

$$\frac{\delta}{v} = \{\theta(1/v)^2 - \theta_R(1/v)^2\}^{1/2},$$

where $\Sigma_{\varepsilon\varepsilon}$ is the covariance matrix of the residuals ε_{t+1} from a regression of r_{t+1} on R_{t+1} and a constant. Also, the stochastic discount factor closest to y_{t+1} is now given by

$$\tilde{m}(v)_{t+1} = m_R(v)_{t+1} + v\alpha_J(1/v)'\Sigma_{\varepsilon\varepsilon}^{-1}\varepsilon_{t+1} = m(v)_{t+1}. \tag{81}$$

Thus, if we want to use the stochastic discount factor that is on the volatility bound of R_{t+1}, as a proxy stochastic discount factor for the larger set (R_{t+1}, r_{t+1}), then the valid discount factor that is closest to $m_R(v)_{t+1}$ is the discount factor with the same expectation v that is on the volatility bound of (R_{t+1}, r_{t+1}). Therefore, δ is the least squares distance between two stochastic discount factors that are on the volatility bounds of (R_{t+1}, r_{t+1}) and its subset R_{t+1} respectively, and is a straightforward measure of the deviation from intersection, which shows the close relation between this special case of the specification error bound and intersection. This relationship also follows from (80), which shows that δ is directly related to the change in the maximum squared Sharpe ratios that can be attained with R_{t+1} and (R_{t+1}, r_{t+1}) respectively. It also follows that δ measures the difference between the variances of the two minimum variance kernels: $\delta = Var[m(v)_{t+1}] - Var[m_R(v)_{t+1}]$.

An estimate of δ^2 can easily be obtained from the sample equivalent of (78), which we will denote by $\hat{\delta}^2$. If the interest is in whether or not there is intersection, then we want to know whether or not $\delta = 0$, and this hypothesis can easily be tested as outlined in Section 3. From the expression in (80) and the discussion in previous sections, it follows that under the null hypothesis that $\delta = 0$,

$$T\frac{\hat{\delta}^2}{v^2(1+\hat{\mu}_R(1/v)^2)} \sim \chi^2_N. \tag{82}$$

In case of specification errors however, the interest is in the case where δ is strictly positive rather than zero. For that case the limiting distribution of $\hat{\delta}$ is derived in Hansen, Heaton, and Luttmer (1995).

Once we concede that $y_{t+1} = m_R(v)_{t+1}$ is not a valid stochastic discount factor for (R_{t+1}, r_{t+1}), we want to have a measure of the difference between $m_R(v)_{t+1}$ and the valid stochastic discount factor that is closest to it, $m(v)_{t+1}$. The specification error bound δ is one such measure, allowing us to make statements about how good or how bad the proxy performs. The fact that δ^2 is equal to the change in maximum Sharpe ratios, makes the measure δ also useful in terms of the optimal portfolio choice for a mean-variance investor. Recall that a mean-variance investor that initially only invests in R_{t+1} can improve his Sharpe ratio from $\mu_R(1/v)$ to $\mu(1/v)$ by including r_{t+1} in his portfolio. Given that there is no intersection between the mean-variance frontiers of R_{t+1} and (R_{t+1}, r_{t+1}), $\hat{\delta}$ provides an estimate for the potential increase in Sharpe ratios. Notice though that such an estimate can also be derived directly from the Wald test-statistic for intersection.

7 Summary

The purpose of this paper is to analyze and illustrate the concept of mean-variance spanning and intersection. We show that there is a duality between mean-variance frontiers and volatility bounds and that mean-variance spanning and intersection can be understood both in terms of mean-variance frontiers and volatility bounds. The paper shows how regression based tests can be used to test for spanning and intersection and how these regression based tests are related to tests for mean-variance efficiency, performance measurement, optimal portfolio choice and specification error bounds.

A The graphical relationship between mean-variance frontiers and volatility bounds

In this appendix we will show some graphical relations between the volatility bound and the mean-variance frontier for a set of asset returns R_{t+1} with expectation μ and covariance matrix Σ. We will start from a point on the volatility bound where the expectation of the minimum variance pricing kernel is v, i.e.,

$$E[m(v)_{t+1}] = v. \tag{83}$$

Using the efficient set variables A, B, and C, and the variance of $m(v)_{t+1}$ as given in (7), the variance of $m(v)_{t+1}$ can be written as

$$Var[m(v)_{t+1}] = A - 2Bv + Cv^2, \tag{84}$$

which is a simple quadratic function of v that describes the volatility bound. The second panel of Figure 1 gives a plot of $Var[m(v)_{t+1}]$ as a function of v.

As shown in Section 2.2, each minimum variance pricing kernel $m(v)_{t+1}$ corresponds to a mean-variance efficient portfolio that has a zero-beta rate $\eta = 1/v$. Recall that a mean-variance efficient portfolio satisfies

$$w = \gamma^{-1}\Sigma^{-1}(\mu - \eta\iota),$$

for a given risk aversion γ and associated zero-beta rate η. Using $\iota' w = 1$ it follos that

$$\gamma = B - \eta A.$$

Furthermore, the expected portfolio return $\mu' w$ satisfies

$$\mu' w = \gamma^{-1}(C - \eta B) = \frac{C - \eta B}{B - \eta A}.$$

Denote the return on the mean-variance efficient portfolio with zero-beta rate $\eta = 1/v$ as $R(v)_{t+1}$ and define $\mu(v) \equiv E[R(v)_{t+1}]$. From the previous relations $\mu(v)$ can be written as a function of v:

$$\mu(v) = \frac{B - Cv}{A - Bv}. \tag{85}$$

Also, the variance $w' \Sigma w$ for a mean-variance efficient portfolio w can be written as a function of $\mu(v)$:

$$Var[R(v)_{t+1}] = \frac{A\mu(v)^2 - 2B\mu(v) + C}{AC - B^2},$$

or as a function of v:

$$Var[R(v)_{t+1}] = \frac{A - 2Bv + Cv^2}{(A - Bv)^2}. \tag{86}$$

The first panel of Figure 1 shows the standard mean-variance efficient frontier, where the expected portfolio return $\mu(v)$ is plotted as a function of the standard deviation of the portfolio return $stdev[R(v)_{t+1}] = Var[R(v)_{t+1}]^{\frac{1}{2}}$.

In this appendix we will restrict ourselves to characterizing the relation between the volatility bound and the mean-variance frontier in terms of v and $\mu(v)$. Given the relations (84) to (86) above it is straightforward to derive the variances of the pricing kernel and the associated mean-variance efficient portfolio as well.

To see the relation between the two graphs, first of all notice that the expected portfolio return $\mu(v)$ is decreasing in v, since from (85) we have that

$$\frac{\partial \mu(v)}{\partial v} = \frac{B^2 - AC}{(A - vB)^2} < 0,$$

and where the inequality follows from the fact that $AC > B^2$, by the Cauchy-Schwarz inequality (see also Ingersoll (1987, p.85)).

Next, from (85) it also follows that for $v = 0$ we have that $\mu(v) = B/A$, which is the expected return on the Global Minimum Variance portfolio.

Looking at the volatility of the pricing kernel we can of course also distinguish the Global Minimum Variance Pricing Kernel, the expectation of which can be found using (84):

$$
\begin{aligned}
0 &= \frac{\partial Var[m(v)_{t+1}]}{\partial v} = -2B + 2Cv^* \\
&\Leftrightarrow \quad v^* = B/C.
\end{aligned}
$$

The second derivative 2C is always positive, which confirms that this is indeed a minimum. Using (85) again, $v = B/C$ corresponds to $\mu(v) = 0$. Thus, when the expectation of the kernel is zero; $v = 0$, this corresponds to the Global Minimum Variance portfolio on the mean-variance frontier, whereas a zero expected return for the mean-variance efficient portfolio, $\mu(v) = 0$, in turn corresponds to the Global Minimum Variance kernel on the volatility bound.

Having characterized the global minima of the two frontiers, the next step is to look at the other extremes, i.e., where $v \to \pm\infty$ and where $\mu(v) \to \pm\infty$. Taking limits and using (85) we get that

$$
\begin{aligned}
\lim_{v \to -\infty} \frac{B - Cv}{A - Bv} &= \frac{C}{B}; \\
\lim_{v \to +\infty} \frac{B - Cv}{A - Bv} &= \frac{C}{B}.
\end{aligned}
$$

Thus, both extremes of the left and right limb of the volatility bound correspond to the same single point on the mean-variance frontier, where the expected portfolio return is $\mu(v) = C/B$. Since by the Cauchy-Schwarz inequality $C/B > B/A$ if $B > 0$, the point where $\mu(v) = C/B$ will plot on the upper limb of the mean-variance frontier. $B > 0$ is the typical case, since this implies that with positive interest rates or zero-beta returns, efficient portfolios have positive expected returns. It is useful to note that $\mu(v) = C/B$ corresponds to the point where a straight line through the origin is tangent to the mean-variance frontier (since $v \to \pm\infty$ corresponds to $\eta = 0$).

Finally, by rewriting (85) as

$$
v = \frac{B - A\mu(v)}{C - B\mu(v)},
$$

we can find the point(s) on the volatility bound that correspond to the extremes of the mean-variance frontier, i.e., where $\mu(v) \to \pm\infty$. Taking limits again, we get that

$$\lim_{\mu(v)\to-\infty}\frac{B-A\mu(v)}{C-B\mu(v)} = \frac{A}{B};$$
$$\lim_{\mu(v)\to+\infty}\frac{B-A\mu(v)}{C-B\mu(v)} = \frac{A}{B}.$$

Notice that we already discussed this result in Section 2 since $v = A/B$, $\eta = B/A$, i.e. the case where the zero-beta return equals the expected return on the Global Minimum Variance portfolio and where there are no corresponding mean-variance efficient portfolios, since the asymptotes of the mean-variance frontier cross the y-axis at B/A, but there is no line tangent to the frontier starting at this point. Again, if $B > 0$, then the Cauchy-Schwarz inequality implies that $A/B > B/C$, implying that this point will be located on the right limb of the volatility bound. Finally, it is useful to note that if we would plot the volatility bound as the standard deviation of the pricing kernel, $Var[m(v)_{t+1}]^{\frac{1}{2}}$, as a function of v, then $v = A/B$ would correspond to the point where a straight line through the origin is tangent to the volatility bound, similar to the mean-variance frontier when $\mu(v) = C/B$.

B The spanning test-statistic in terms of Sharpe ratios

In this appendix we show how the spanning test statistic can be interpreted in terms of Sharpe ratios, a result that was presented in Section 5.3. Recall from Section 5.3 that the covariance matrix of the OLS-estimates $\hat{\beta}$ equals

$$\Sigma_{\varepsilon\varepsilon}\otimes T^{-1}\begin{pmatrix} 1+\mu_R'\Sigma_{RR}^{-1}\mu_R & -\mu_R'\Sigma_{RR}^{-1} \\ -\Sigma_{RR}^{-1}\mu_R & \Sigma_{RR}^{-1}\end{pmatrix}.$$

Premultiplying with H_{span} and postmultiplying with H_{span} as defined in (55) yields

$$\begin{aligned} & H_{span}\left(\Sigma_{\varepsilon\varepsilon}\otimes T^{-1}\begin{pmatrix} 1+\mu_R'\Sigma_{RR}^{-1}\mu_R & -\mu_R'\Sigma_{RR}^{-1} \\ -\Sigma_{RR}^{-1}\mu_R & \Sigma_{RR}^{-1}\end{pmatrix}\right)H_{span}' \\ = {} & \Sigma_{\varepsilon\varepsilon}\otimes T^{-1}\begin{pmatrix} 1+C_R & -B_R \\ -B_R & A_R\end{pmatrix}; \qquad (87)\end{aligned}$$

the inverse of which is

$$\Sigma^{-1} \otimes \frac{T}{A_R(1+C_R) - B_R^2} \begin{pmatrix} A_R & B_R \\ B_R & 1+C_R \end{pmatrix}. \tag{88}$$

Similarly, for h_{span} in (55) we have

$$\left(I_N \otimes \begin{pmatrix} 1 & 0_K' \\ 0 & \iota_K' \end{pmatrix} \right) b - \left(I_N \otimes \begin{pmatrix} 0 \\ \iota \end{pmatrix} \right) = \begin{pmatrix} \alpha_1 \\ \beta_1 \iota_K - 1 \\ \alpha_2 \\ \beta_2 \iota_K - 1 \\ \vdots \\ \alpha_N \\ \beta_N \iota_K - 1 \end{pmatrix}. \tag{89}$$

Premultiplying (88) with h_{span} and postmultiplying with h_{span}', we get, after replacing population moments by their sample equivalents:

$$\xi_W^{span} = T \frac{\hat{A}_R \hat{\alpha}' \hat{\Sigma}^{-1} \hat{\alpha} - 2\hat{B}_R \hat{\alpha}' \hat{\Sigma}^{-1}(\iota_N - \hat{\beta}\iota_K) + (1+\hat{C}_R)(\iota_N - \hat{\beta}\iota_K)' \hat{\Sigma}^{-1}(\iota_N - \hat{\beta}\iota_K)}{\hat{A}_R(1+\hat{C}_R) - \hat{B}_R^2}. \tag{90}$$

Next note that the maximum attainable Sharpe ratio from R_{t+1}, for $\eta = B_R/A_R$, is equal to

$$\theta_R \left(\frac{B_R}{A_R} \right)^2 = C_R - \frac{B_R^2}{A_R}.$$

For simplicity, write $A = A_R + \Delta A$, $B = B_R + \Delta B$, and $C = C_R + \Delta C$, where the definitions of ΔA, ΔB, and ΔC follow from (54) and (55). Evaluating $\theta(\eta)$ in this same value of η, we get

$$\begin{aligned} \theta \left(\frac{B_R}{A_R} \right)^2 &= C - 2B\frac{B_R}{A_R} + A\frac{B_R^2}{A_R^2} \\ &= C_R + \Delta C - 2(B_R + \Delta B)\frac{B_R}{A_R} + (A_R + \Delta A)\frac{B_R^2}{A_R^2} \\ &= \theta_R \left(\frac{B_R}{A_R} \right)^2 + \frac{1}{A_R} \left(A_R \Delta C - 2B_R \Delta B + \frac{B_R^2}{A_R} \Delta A \right) \end{aligned}$$

Dividing by $(1+C_R) - B_R^2/A_R = 1 + \theta_R \left(\frac{B_R}{A_R} \right)^2$ gives

$$\frac{\theta \left(\frac{B_R}{A_R} \right)^2 - \theta_R \left(\frac{B_R}{A_R} \right)^2}{1 + \theta_R \left(\frac{B_R}{A_R} \right)^2} = \frac{A_R \Delta C - 2B_R \Delta B + \frac{B_R^2}{A_R} \Delta A}{A_R(1+C_R) - B_R^2}$$

$$
\begin{aligned}
&= \frac{A_R \cdot C - 2B_R \cdot B + \left(C_R + 1 - 1 - \theta\left(\frac{B_R}{A_R}\right)^2\right) \cdot A}{A_R(1 + C_R) - B_R^2} \\
&= \frac{A_R \cdot C - 2B_R \cdot B + (1 + C_R) \cdot A}{A_R(1 + C_R) - B_R^2} - \frac{\cdot A}{A_R}.
\end{aligned}
$$

Replacing all population moments with their sample equivalents again and noting that $1/A_R$ is the variance of the global minimum variance portfolio of R_{t+1}, i.e., $1/A_R = (\sigma^0_R)^2$, and similarly, $1/A = (\sigma^0)2$, we finally obtain

$$
\begin{aligned}
\xi^{span}_W &= T\frac{\hat{\theta}\left(\frac{B_R}{A_R}\right)^2 - \hat{\theta}_R\left(\frac{B_R}{A_R}\right)^2}{1 + \hat{\theta}_R\left(\frac{B_R}{A_R}\right)^2} + T\frac{\hat{A} - \hat{A}_R}{\hat{A}_R} \\
&= T\left(\frac{1 + \hat{\theta}(\eta^0_R)^2}{1 + \hat{\theta}_R(\eta^0_R)^2} + \frac{(\hat{\sigma}^0_R)^2}{(\hat{\sigma}^0)^2} - 2\right).
\end{aligned}
$$

References

Balduzzi P., and Kallal H., 1997, "Risk Premia and Variance Bounds", Journal of Finance, (52), 1913-1949.

Balduzzi, P., and Robotti, C., 2000, "Minimum-Variance Kernels and Economic Risk Premia", Working Paper, Boston College.

Bansal, R., and Lehmann, B.N., 1997, "Growth-Optimal Portfolio Restrictions on Asset Pricing Models", Macroeconomic Dynamics, 1, 333-354.

Bernardo, A., and Ledoit, O., 1996, "Gain, Loss and Asset Pricing", Journal of Political Economy, 108, 144-172.

Bekaert, G., and Liu, J.L., 1999, "Conditioning Information and Variance Bounds on Pricing Kernels", Working Paper, Stanford University.

Bekaert, G., and Urias, M.S., 1996, "Diversi…cation, Integration, and Emerging Market Closed-End Funds", Journal of Finance, 51, 835-870.

Campbell, J.Y., and Viceira, L.M., 1996. "Consumption and Portfolio Decisions when Expected Returns are Time-Varying", Quarterly Journal of Economics, 114, 433-495.

Chen, Z., and Knez, P.J., 1996, "Portfolio Performance Measurement: Theory and Applications", Review of Financial Studies, 9, 511-556.

Cochrane, J.H., 1996, "A Cross-Sectional Test of an Investment-Based Asset Pricing Model", Journal of Political Economy, 1996, 104, 572-611.

Cochrane, J.H., 2001, Asset Pricing, Forthcoming, Princeton University Press.

Cochrane, J.H., and Saá-Requejo, J., 1995, "Beyond Arbitrage: Derivatives Pricing in Incomplete Markets", Journal of Political Economy, 108, 79-119.

Cumby, R.E., and Glen, J.D., 1990, "Evaluating the Performance of International Mutual Funds", Journal of Finance, 45, 497-521.

Dahlquist, M., and Söderlind, P., 1999, "Evaluating Portfolio Performance with Stochastic Discount Factors", Journal of Business, 72, 347-383.

DeRoon, F.A., and Nijman, Th.E., 1999, "Testing for Mean-Variance Spanning: A Survey", Working Paper, Tilburg University.

DeRoon, F.A., Nijman, Th.E., and Veld, C., 2000, "Hedging Pressure Effects in Futures Markets", Journal of Finance, 55, 1437-1456.

DeRoon, F.A., Nijman, Th.E., and Werker, B.J.M., 2001, "Testing for Mean-Variance Spanning with Short Sales Constraints and Transaction Costs: The Case of Emerging Markets", Journal of Finance, 56, 723-744.

DeRoon, F.A., Nijman, Th.E., and Werker, B.J.M., 1998, "Testing for Spanning with Futures Contracts and Nontraded Assets: A General Approach", Working Paper, Tilburg University.

DeSantis, G., 1994, "Asset Pricing and Portfolio Diversi…cation: Evidence from Emerging Financial ", in Howell, M, (ed.): Investing in Emerging Markets, Euromoney Books, London.

DeSantis, G., 1995, "Volatility bounds for Stochastic Discount Factors: Tests and Implications from International Financial Markets", Working Paper, University of Southern California.

Errunza, V., Hogan, K., and Hung, M.W., 1999, "Have the Gains from International Diversi…cation Disappeared?", Journal of Finance, 54, 2075-2107.

Fama, E.F., 1984, "Forward and Spot Exchange Rates", Journal of Monetary Economics, 14, 319-338.

Fama, E.F., and French, K.R., 1996, "Multifactor Explanations of Asset Pricing Anomalies", Journal of Finance, 51, 55-84.

Ferson, W.E., 1995, "Theory and Empirical Testing of Asset Pricing Models", in: Jarrow, R.A., Maksimovic, V., and Ziemba, W.T. (eds.), Handbooks in Operations Research and Management Science 9: Finance, Elsevier, North Holland.

Ferson, W.E., Foerster, S.R., and Keim, D.B., 1993, "General Tests of Latent Variable Models and Mean-Variance Spanning", Journal of Finance, 48, 131-156.

Ferson, W.E., and Schadt, R.W., 1996, "Measuring Funds Strategy and Performance in Changing Economic Conditions", Journal of Finance, 51, 25-462.

Ferson, W.E., and Siegel, A., 1997, "The E¢cient Use of Conditioning Information in Portfolios", Working Paper, University of Washington.

Gibbons, M., Ross, S., and Shanken, J., 1989, "A Test of the E¢ciency of a Given Portfolio", Econometrica, 57, 1121-1152.

Glen, J., and Jorion Ph., 1993, "Currency Hedging for International Portfolios", Journal of Finance, 48, 1865-1886.

Hall, A., and Knez, P.J., 1995, "Testing Portfolio Performance", Working Paper, North Carolina State Universtiy.

Hansen, L.P., Heaton, J., and Luttmer, E.G.J.,1995, "Econometric Evaluation of Asset Pricing Models", Review of Financial Studies, 8, 237-274.

Hansen, L.P., and Jagannathan, R., 1991, "Implications of Security Market Data for Models of Dynamic Economies", Journal of Political Economy, 99, 225-262.

Hansen, L.P., and Jagannathan, R., 1997, "Assessing Speci…cation Errors in Stochastic Discount Factor Models", Journal of Finance, 52, 557-590.

Harvey, C.R., 1989, "Time-Varying Conditional Covariances and Tests of Asset Pricing Models", Journal of Financial Economics, 24, 289-317.

Harvey, C.R., 1995, "Predictable Risk and Returns in Emerging Markets", Review of Financial Studies, 8, 773-816.

Huberman, G., and Kandel, S., 1987, "Mean Variance Spanning", Journal of Finance, 42, 873-888.

Ingersoll, J.E., 1987, Theory of Financial Decision Making, Rowman and Little…eld, Maryland.

Jensen, M.C., 1968, "The Performance of Mutual Funds in the Period 1945-1964", Journal of Finance, 23, 389-416.

Jobson, J.D., and Korkie, B., 1982, "Potential Performance and Tests of Portfolio E¢ciency", Journal of Financial Economics, 10, 433-466.

Jobson, J.D., and Korkie, B., 1984, "On the Jensen Measure and Marginal Improvements in Portfolio Performance: A Note", Journal of Finance, 39, 245-251.

Jobson, J.D., and Korkie, B., 1989, "A Performance Interpretation of Multivariate Tests of Intersection, Spanning and Asset Pricing", Journal of Financial and Quantitative Analysis, 24, 185-204.

Kan, R., and Zhou, G., 1999, "A Critique of the Stochastic Discount Factor Methodology", Journal of Finance, 54, 1221-1248.

Kandel, S., and Stambaugh, R.F., 1987, "On Correlations and Inferences about Mean-Variance E¢ciency", Journal of Financial Economics, 18, 61-90.

Kandel, S., and Stambaugh, R.F., 1989, "A Mean-Variance Framework for Tests of Asset Pricing Models", Review of Financial Studies, 2, 125-156.

Kirby, C., 1998, "The restrictions on predictability implied by rational asset pricing models", Review of Financial Studies, 11(2), 343-382.

Luttmer, E.G.J., 1996, "Asset Pricing in Economies with Frictions", Econometrica, 64, 1429-1467.

MacKinlay, A.C., 1995, "Multifactor Models do not Explain Deviations from the CAPM", Journal of Financial Economics, 38, 3-28.

Markowitz, H.M., 1952, "Portfolio Selection", Journal of Finance, 7, 77-

91.

Markowitz, H.M., 1991, Portfolio Selection, Basil Blackwell, Cambridge, Massachusetts.

Newey, W.K., and West, K.D., 1987, "A Simple Positive Semi-de…nite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix", Econometrica, 55, 703-708.

Shanken, J., 1987, "Multivariate Proxies and Asset Pricing Relations: Living with the Roll Critique", Journal of Financial Economics, 18, 91-110

Shanken, J., 1990, "Intertemporal Asset Pricing: An Empirical Investigation", Journal of Econometrics, 45, 99-120.

Sharpe, W., 1963, "A Simpli…ed Model for Portfolio Analysis", Management Science, 9, 277-293.

Snow, K.N., 1991, "Diagnosing Asset Pricing Models Using the Distribution of Asset Returns", Journal of Finance, 46, 955-983.

Stevens, G.V.G., 1998, "On the Inverse of the Covariance Matrix in Portfolio Analysis", Journal of Finance, 53, 1821-1827.

Treynor, J.L., and Black, F., 1973, "How to Use Security Analysis to Improve Portfolio Selection", Journal of Business, 46, 66-86.