

## Chapter 5: Conclusions

This thesis develops new tests for the efficiency of a given portfolio and two-fund separation in a Stochastic Dominance (SD) framework, and illustrates them in a range of empirical settings.

The motivations for this are discussed in Chapter 1; the ‘standard’ Mean-Variance (MV) framework has several drawbacks that open the way for alternatives; most notably the fact that MV analysis can violate the regularity condition of non-satiation (investors prefer more over less), and that variance is only a correct measurement of risk under very restrictive circumstances, such as normal returns (or several classes that can be described by just a scale and location parameter) or quadratic utility. The former is usually empirically refuted, the latter is highly unlikely. Also, the SD framework is much more flexible than the MV one, as SD allows for non-linear marginal utility. To be precise, the customary SD orders can be formulated as a series of restrictions on marginal utility (also called the kernel), so that the utility specification explicitly conforms to these regularity conditions.

Our new tests (following much of the Asset Pricing literature) have a focus on the market portfolio. Though not strictly required, it is logical to test for efficiency of the market, and if two-fund separation holds, this must necessarily involve the market. The models contain a representative investor, meaning that a single investor (acting as the aggregate of all investors) must hold the market portfolio. However, in the SD framework we have not yet succeeded in describing all conditions that allow for such an aggregation. The results so far are not hopeful: completeness of the market is one way to support aggregation, but is empirically difficult to defend. Restrictions on the utility specifications is another, but to achieve sufficient homogeneity of preferences we would have to parameterise the kernel to such an extent that one of the advantages of SD analysis – its flexibility – is gone (see Rubinstein (1974), Brennan & Kraus (1978)). Finally, there may be a less restrictive combination of utility specifications and return distributions that result in convexity of the efficient set, but at this point, we lack conditions that are both necessary and sufficient.

Nevertheless, the empirical question of the SD efficiency of a given portfolio is interesting in its own right. Chapter 2 develops new tests for SSD and TSD efficiency (which could be

extended to higher orders as well). This is not the first test for SD efficiency proposed, but it is more powerful than the comparable earlier tests, and in contrast to the existing linear programming tests (e.g. Post (2003)), our tests consider all pricing errors jointly, rather than the maximum positive error only. The test statistics can be computed by iterating a convex quadratic programming problem that can be solved using standard mathematical programming techniques. We derive bounded, asymptotical statistics for this test, as well a variant of the new test under restrictions on the portfolio weights (which is useful outside the context of market portfolio efficiency).

Futhermore, we show the power and size properties of the efficiency test through a series of simulations, which show a power almost comparable to that of the Gibbons-Ross-Shanken (1989, GRS) test for MV efficiency for medium to large samples under both normal distributions (which gives the home-field advantage to the MV test) and empirical distributions. The small loss in power seems an acceptable price for the superior statistical size for non-normal distributions, where the GRS test does not apply due to the inadequacy of the mean-variance framework.

The third chapter extended this work and proposed a test for two-fund separation (TFS). If TFS is not rejected, all utility specifications that conform to the SD criterion being employed will explain the market. This means that the test needs to have a different focus (maximizing the pricing errors instead of minimizing them) and due to technical considerations, the TFS test is in fact composed of a series of such maximizations; we have to solve a convex quadratic programming problem for every asset. This means that the results are not completely comparable to the efficiency tests, as they use a different error measure.

The size properties reflect a related issue, namely our focus on minimizing Type II error (failure to detect violations of TFS). However, the power properties (again analysed through simulations) are very encouraging.

Chapter 4 shows how the concepts of SD efficiency (Chapter 2) and two-fund separation (Chapter 3) are related and can be combined to give insights in the robustness of the possible utility specifications. If there is SD efficiency, at least one kernel conforming to the criterion used rationalizes the market; if there is two-fund separation, *all* specifications in that class are possible. As we have very little prior information about how to model utility, such a conclusion can be very reassuring. However, not all datasets support two-fund separation; while the beta-sorted, size-sorted and reversal sorted portfolios do so, the book-to-market

portfolios remain dependent on a correct parametrisation, and the momentum sorted portfolios cannot be explained at all using SD utility specifications. The beta sorted portfolios do illustrate the contrast between the MV and SD frameworks, as they are SD efficient, but the MV criterion misses this and rejects efficiency. Also, the sample period plays a non-trivial role, and especially the period 1946-1964 remains hard to explain in almost every dataset.

We started the introductory section on a rather negative note, namely that the dominant mean-variance methodology generally is not economically meaningful and it may fail to detect inefficiency for portfolios that no non-satiabile risk-averter would select (especially if the portfolios have relatively fat left tails), and may reject efficiency for portfolios that are perfectly good solutions for some non-satiabile risk-averters (especially if the portfolios have relatively thin left tails). While there remains a lot of work to be done, we can at least end on the brighter note that both efficiency and two-fund separation can now be tested in a framework without these limitations.

## Samenvatting (summary in Dutch)

Investeerders dienen een afweging te maken tussen rendement en risico. Het bepalen van het rendement is meestal vrij eenvoudig, maar het modelleren van risico op financiële markten is een intrigerend probleem: de standaard aanpak – het zogenaamde Mean-Variance (MV) raamwerk – beschouwt alle afwijkingen van het verwachte rendement als risico, en neemt de variantie als maatstaf. Deze aanpak heeft meerdere nadelen, onder andere dat afwijkingen naar boven en naar beneden beide als negatief worden gezien, terwijl dit bij rendementen duidelijk een onwaarschijnlijke situatie kan zijn. Dit kan zelfs tot gevolg hebben dat investeerders (in sommige gevallen, met name bij zeer grote positieve rendementen) minder rendement zouden prefereren boven meer.

Stochastische Dominantie (SD) biedt een alternatieve aanpak voor het modelleren van risico: in plaats van een parametrische benadering van de preferenties van beleggers (de nutsfunctie heeft dan een vooraf vastgestelde functionele vorm) wordt een non-parameterische benadering toegepast die de data de (afgeleide van de) nutsfunctie laat bepalen, waarbij juist de condities van ‘non-satiation’ (onverzadigbaarheid) en risico-aversie dwingend worden opgelegd. We spreken dan van SSD (Second order SD); hogere ordes leggen meer restricties op.

Dit proefschrift ontwikkelt nieuwe tests voor de efficiëntie van een gegeven portefeuille in dit SD raamwerk, evenals voor “two-fund separation”, en illustreert deze testen en hun onderlinge relatie met behulp van empirische data.

De motivatie hiervoor wordt gegeven in Hoofdstuk 1; waar behalve de hierboven al genoemde aspecten van het MV-raamwerk Ook wordt ingegaan op de beperkte geldigheid van variantie als risico-maatstaf: normaal verdeelde rendementen (dan wel verdelingen die met uitsluitend een locatie- en schaal-parameter kunnen worden beschreven), of een kwadratische nutsfunctie zijn noodzakelijk. De eerste voorwaarde kan empirisch worden weerlegd, de tweede is onwaarschijnlijk.

In navolging van een groot deel van het vakgebied Asset Pricing neemt de marktportefeuille (waarin alle beleggingsobjecten zijn opgenomen en naar marktwaarde worden gewogen) een centrale plaats in. Efficiëntie betekent dat er een ‘well behaved investor’ – oftewel een belegger die aan de standaard condities voldoet – is die de geanalyseerde portefeuille zou aanhouden. Aangezien alle beleggingsobjecten in de markt dienen te worden aangehouden is

de efficiëntie van de marktportefeuille een interessant vraagstuk. “Two-fund separation” vraagt zich af of *alle* beleggers dezelfde portefeuille zullen aanhouden, waarbij op basis van dezelfde redenering de markt betrokken moet zijn.

Beide modellen gaan uit van een representatieve belegger; een belegger die het geaggregeerde effect van alle individuele beleggers weergeeft. Helaas is er in het SD raamwerk nog geen volledige beschrijving mogelijk van de voorwaarden waaraan voldaan moet zijn om van een dergelijke constructie gebruik te kunnen maken. De resultaten zijn niet bemoedigend: compleetheid van de markt is een mogelijkheid, maar is empirisch moeilijk verdedigbaar. Verdere restricties op de marginale nutsfunctie betekenen een dusdanige parameterisatie dat de flexibiliteit van de SD aanpak verloren gaat. Wellicht is er een minder restrictieve combinatie van restricties op rendementsverdelingen en nutsfuncties mogelijk als deze tot een convexe efficiënte set leidt, maar voorwaarden die noodzakelijke en voldoende zijn ontbreken vooralsnog.

Niettemin, de empirische vraag of een gegeven portefeuille SD efficiënt is blijft interessant. Hoofdstuk 2 ontwikkelt nieuwe tests voor SSD en TSD (tweede en derde orde; uitbreiding naar hogere ordes is mogelijk) efficiëntie. Deze tests hebben een hogere statistische ‘power’ (de mate waarin een nulhypothese verworpen wordt als deze niet correct is) dan bestaande tests, onder andere doordat in plaats van de maximale positieve fout (‘pricing error’) alle fouten in de analyse betrokken worden. De test statistieken kunnen berekend worden door een convex kwadratisch probleem op te lossen, hetgeen slechts relatief eenvoudige optimalisatieprocedures vereist. Het hoofdstuk bevat ook de statistische onderbouwing en een versie van de test die restricties op de portefeuilles toestaat.

De eigenschappen met betrekking tot power en size worden onderzocht door middel van een serie simulaties; de power blijkt bijna vergelijkbaar te zijn met die van de Gibbons-Ross-Shanken (1989, GRS) test voor MV efficiëntie; dit geldt voor zowel middelgrote als grote steekproeven onder zowel de normale verdeling (wat een thuisvoordeel oplevert voor het MV raamwerk) als empirische verdelingen. Het geringe verlies in power lijkt een acceptabele prijs te zijn voor de superieure statistische ‘size’ voor niet-normale verdelingen, waar de GRS test niet toepasbaar is wegens de tekortkomingen van het MV raamwerk.

Het derde hoofdstuk is een uitbreiding van deze resultaten en ontwikkelt een test voor two-fund separation (TFS). Als TFS niet verworpen kan worden zullen alle nutsspecificaties die aan de regulariteitscondities (behorende bij het SD criterium) voldoen, de markt verklaren. Dit betekent dat de test een andere focus dient te hebben dan de efficiëntie test; de 'pricing errors' worden gemaximaliseerd in plaats van geminimaliseerd. Om technische redenen is de TFS test in feite een serie van dergelijke maximalisaties; dit heeft als gevolg dat de maatstaven van de TFS en efficiëntietesten verschillen en de resultaten daardoor niet volledig vergelijkbaar zijn. De eigenschappen met betrekking tot de size reflecteren onze focus op het minimaliseren van de zogenaamde fouten van de tweede soort (het niet detecteren van schendingen van TFS), maar de power (opnieuw geanalyseerd door simulatie) is hoog.

Hoofdstuk 4 behandelt de relatie tussen SD efficiëntie en TFS, en laat zien hoe de twee testen gezamenlijk informatie kunnen geven over de mate waarin de specificatie van de nutsfunctie een rol speelt. Als er SD efficiëntie is, is er tenminste een kernel die voldoet aan de voorwaarden en die de markt kan verklaren. Onder TFS zijn alle toegestane kernels acceptabel. Aangezien we bijzonder weinig informatie hebben over de vorm van de nutsfunctie, kan een dergelijke conclusie een bijdrage leveren. Echter, niet alle onderzochte datasets ondersteunen TFS; bij de sets gesorteerd op beta, marktwaarde en 'reversal' kan TFS niet verworpen worden, maar de 'book-to-market' gesorteerde portefeuilles staan alleen efficiëntie toe (de beta portefeuilles illustreren de manco's van het MV raamwerk). De momentum portefeuilles maken een op risico gebaseerde verklaring van de markt niet mogelijk; ook efficiëntie wordt verworpen. Ook moet geconstateerd worden dat de steekproefperiode een rol speelt; met name de period 1946-1964 is in de meeste datasets problematisch.

De introductie van het proefschrift begint met een vrij negatieve klank, namelijk dat de dominante MV methodologie in zijn algemeenheid niet economisch verdedigbaar is, en er toe kan leiden dat inefficiënte portefeuilles, die geen enkele onverzadigbare en risico-averse belegger zou kiezen, toch geselecteerd worden, en dat efficiënte portefeuilles ten onrechte geëlimineerd worden uit de efficiënte set. Hoewel er nog veel werk dient te gebeuren, kunnen wij tenminste eindigen met de positievere boodschap dat zowel efficiëntie als two-fund separation nu getest kunnen worden in een raamwerk zonder deze beperkingen.