The origins of the mean-variance approach in finance: revisiting de Finetti 65 years later

published in Decisions in Economics and Finance, **30**, 2007, p. 19-49 DOI: 10.1007/s10203-007-0067-7

Flavio Pressacco and Paolo Serafini, Università di Udine

Abstract: In a recent critical review of de Finetti's paper "Il problema dei pieni", the Nobel Prize H. Markowitz recognized the primacy of de Finetti in applying the mean-variance approach to finance, but pointed out that de Finetti did not solve the problem for the general case of correlated risks. We argue in this paper that a more fair sentence should be: de Finetti did solve the general problem only under an implicit hypothesis of regularity which is not always satisfied. Moreover, as we shall see, a natural extension of de Finetti's procedure to non regular cases offers a general solution for the correlation case and shows that de Finetti anticipated a modern mathematical programming approach to mean-variance problems.

1. Introduction

In a paper appeared in 1940 (de Finetti (1940)) and concerning proportional reinsurance decisions, Bruno de Finetti applied the mean-variance approach to solve a financial problem under uncertainty. The paper written in Italian on an actuarial journal and at the beginning of the Second World War went unobserved to researchers in financial economics and his knowledge remained for a very long time restricted to European actuarial circles.

In the following decade, Markowitz published his milestone papers on the mean-variance approach to portfolio selection, which brought him the Nobel Prize in Economics and the name of founder of modern finance (Markowitz (1952, 1956)). On the contrary de Finetti's paper fell into oblivion; he himself did not insert it into the list of his own papers bearing some, even mild, connections with economic problems (de Finetti (1969)).

Only recently, alerted by signaling from late (or indirect) de Finetti's scholars, leading economists of the English speaking world recognized the importance of de Finetti's paper. In particular Rubinstein (2006 a) expressed the opinion that de Finetti anticipated much of mean-variance portfolio theory and had quietly laid the foundations of modern finance theory and practice twelve years earlier than Markowitz. Markowitz (2006) himself, involved by Rubinstein, offered a kind critical review of de Finetti's paper.

Among other considerations Markowitz found a shortcoming in de Finetti's treatment of the general case of correlated risks. It turns out that the procedure suggested by de Finetti (as a plain extension of the one correctly applied to the no correlation case) to solve the problem under correlation among risks works only under a special hypothesis (which we define regularity hypothesis) which may not be verified. Hence Markowitz concluded that de Finetti did not solve the general problem of correlated risks and also that advanced mathematical programming techniques, not known yet at the time of de Finetti's paper, must be applied to solve the proportional reinsurance problem in the general case. In turn Rubinstein (2006 b) writes that de Finetti 'fully worked out only an algorithm for solving the problem in the special case of uncorrelated returns'.

However, in our opinion these sentences, although literally correct, are rather drastic. It is our opinion that, as we will show in this paper, the tools proposed by de Finetti, along with his very smart ideas, deserve

and should obtain a double reevaluation. On one side a natural adjustment of his sequential procedure offers an easy solution to the general proportional reinsurance problem even in case where the regularity hypothesis is not verified. On the other side these tools may be successfully embedded in the modern environment of mathematical programming to pave the way for a meaningful sequential procedure to define the optimum path.

Quite likely de Finetti neglected the lack of regularity, thinking that it was an unlike circumstance in the reinsurance market. To this aim, we also discuss conditions granting regularity in case of constant positive correlation among risks.

The paper is organized as follows. Section 2 recalls the essentials of de Finetti's approach to the meanvariance approach to optimum proportional reinsurance decisions. Section 3 describes de Finetti's correct treatment of the case of no correlation. Section 4 is devoted to the extension of the procedure to the case of correlated risks and to a discussion of the shortcomings of the regularity hypothesis implicit in de Finetti's treatment. Section 5 offers a correct extension of de Finetti's procedure to find the optimum path in the general correlation case. Sections 6 and 7 show how to exploit de Finetti's ideas in the up-to-date mathematical programming scenario. In Section 8 we provide an enlightening example of a three dimensional non regular case. Sufficient conditions for regularity are shown in Section 9. In Section 10 we shortly discuss the connections with portfolio theory and CAPM in actuarial applications. A final conclusion is offered in Section 11. A technical appendix devoted to prove the results of Section 9 concludes the paper.

Before entering into the core of the paper we would say something about our choice of symbols and terminology. As to the symbols we choose to adopt the ones typical in the optimization literature, but giving in a while a conversion table so as to allow the reader a better comparison with the papers de Finetti (1940) and Markowitz (2006). As to the terminology, a more detailed explanation of our approach is in order. We tried to be as much as possible close to de Finetti's style to make large recourse to geometric ideas. Indeed it was typical of de Finetti to use a geometric approach to analyze problems, to describe the steps bringing logically to the solution and to explain the meaning and the properties of the solution. This style characterizes strongly also his 1940 paper. It may be perceived today as rather unusual, but we think that it is a simple and powerful approach and at the same time a tribute to de Finetti's spirit.

2. The pillars of de Finetti's approach

Let us shortly recall the essentials of de Finetti's approach. An insurance company is faced with n risks (policies). The net profit of these risks is represented by a vector of random variables with expected value $m := \{m_i > 0 : i = 1, ..., n\}$ and non singular covariance matrix $\mathbf{V} := \{\sigma_{ij} : i, j = 1, ..., n\}$. The company has to choose a proportional reinsurance or retention strategy specified by a retention vector x. The retention strategy is feasible if $0 \le x_i \le 1$ for all i. A retention x induces a random profit with expected value $E = x^{\top}m$ and variance $V = x^{\top}\mathbf{V}x$.

A retention x is by definition optimal (Pareto efficient) if for no feasible retention y we have both $x^{\top}m \leq y^{\top}m$ and $x^{\top}\mathbf{V}x \geq y^{\top}\mathbf{V}y$, with at least one inequality strict. Let X^* be the set of optimal retentions. The core of de Finetti's approach is represented by the following simple smart ideas, where the geometric intuition plays an important role.

The set of feasible retentions is represented by points of the *n* dimensional unit cube. The set X^* is a path in this cube. It connects the natural starting point, the vertex **1** of full retention (with the largest expectation $E = \sum_i m_i$), to the opposite vertex **0** of full reinsurance (zero retention and hence minimum null variance).

Let us consider a point $x^* \in X^*$ (with $x^* > 0$). De Finetti argues that a measure of the advantage coming from a (small) additional (or initial) reinsurance of the *i*-th risk is given by the value that the

functions (defined on the unit cube for x > 0 in de Finetti (1940) and on the whole unit cube $x \ge 0$ in the extension of his method)

$$F_i(x) := \frac{1}{2} \frac{\frac{\partial V}{\partial x_i}}{\frac{\partial E}{\partial x_i}} = \sum_{j=1}^n \frac{\sigma_{ij}}{m_i} x_j , \qquad i = 1, \dots, n,$$
(1)

assume at x^* . Note that $F_i(x)$ may be seen as (one half) the decrease of a properly normalized variance, i.e., the decrease of the variance measured in units of the expectation numeraire.

As we shall see in the sequel, the functions $F_i(x)$ play a key role in describing the efficient set. Taking into account this fact and that the symbol used by de Finetti was K_i , we will call these functions the key functions of our problem. The connection between the functions $F_i(x)$ and the optimal set X^* comes from some intuitive golden rules embedded in de Finetti's approach.

The first golden rule is that at any $x \in X^*$ an additional (or initial) reinsurance must be provided only for the risk(s) *i* for which $F_i(x)$ is maximum. Intuitively, an additional reinsurance is only for those risks giving the biggest advantage. Let us denote $\lambda(x) := \max_i F_i(x)$.

As we proceed along the optimum path, de Finetti argues that $\lambda(x) : X^* \to R$ is a monotonically decreasing continuous function, so that there is a one-to-one correspondence between $x \in X^*$ and $\lambda(1) \geq \lambda(x) \geq \lambda(0)$. The notation used by de Finetti for our λ is A. As we shall see later, the parameter λ is nothing but the counterpart of the Lagrange multiplier λ in constrained optimization as well as in Markowitz (2006). Given its importance in determining the optimal set X^* , we call λ the key parameter.

Let us consider the set I(x) of those risks *i* such that $F_i(x) = \lambda(x)$. If I(x) is a singleton the direction of movement of the optimum path leaving *x* is unequivocal. What if I(x) has two or more elements? In such a case de Finetti provides a second golden rule.

Second golden rule: if $x \in X^*$ is a point where I(x) has more than one element, then leave x in the direction where all risks of I(x) are additionally (or initially) reinsured in such a way to preserve the equality of the $F_i(x)$ functions for all $i \in I(x)$.

Given the form of the $F_i(x)$ functions, this implies a movement on a segment of the cube characterized by the set of equations $F_i(x) = \lambda(x)$ for all $i \in I(x)$. Note that only the risks belonging to I(x) play an active role in defining the direction of the optimum path at x. They may be expressively named *directional* risks. Note also that on that direction the other non directional risks are fixed at some constant retention level. Moreover, $F_i(x) < \lambda(x)$ for all $i \notin I(x)$, and, as long as the inequality is preserved, the direction does not change and the set I(x) is invariant.

But, argues de Finetti, going on that direction along the optimum path, we will find a point x (or a value $\lambda(x)$) such that one of the non directional risks, say j, attains a value of $F_j(x)$ matching the common value $F_i(x) = \lambda(x)$ of all the previous directional risks and the corresponding inequality $(F_j(x) < \lambda(x))$ is no longer preserved.

This introduces the third golden rule: at that point x there is a change of direction in the optimum path. Indeed, also the risk j is able to obtain at x, by a reinsurance movement, the biggest advantage $\lambda(x)$ so that it becomes coherently a member of the set I(x) by joining the other current members. A directional change is then forced by the fact that, according to the second golden rule, a new set of equations $F_i(x) = \lambda(x)$, with the addition of the equation for the newcomer directional risk, determines the new direction of the optimum path.

We underline the need to be cautious in accepting the procedure previously described. As we shall see, the procedure is correct in the correlation case only under a regularity condition to be introduced and discussed later and which de Finetti gave for granted in any case.

Although the procedure does not extend in its simple form to the general case, the role played by the key functions and the key parameter remains, as we shall see, fundamental in determining the optimum path. This shows the far reaching intuition of de Finetti in a time when the modern powerful mathematical programming techniques were not available yet. Moreover, de Finetti solved the problem in an algorithmic way and this was unusual at the time of his writing, when computers did not exist yet. We stress that it is the algorithmic nature of his method which allows for a simple computation of optimal points for any value of the expected profit even in the general case. We may say that he was a precursor of the celebrated critical line algorithm (CLA) by Markowitz (1956).

Before proceeding we find useful to provide the reader with a correspondence table between the symbols used in this paper and the ones actually employed by de Finetti (1940) and Markowitz (2006).

	de Finetti	Markowitz	this paper
retentions	a_i	X_i	x_i
key parameter	A	λ	λ
risk exp. values	\bar{k}_i	$\mu_i \text{ or } m_i$	m_i
risk cov. matrix		С	v
risk cov. matrix entries	$r_{ij}\bar{\sigma}_i\bar{\sigma}_j$	σ_{ij}	σ_{ij}
key function	$K_i(a)$		$F_i(x)$
risk state		UP	I^1
risk state		DOWN	I^0
risk state		IN	$I(x)$ and I^*

3. The no correlation case

We briefly describe in some detail how the procedure devised by de Finetti works in the simple case of no correlation, when $F_i(x)$ reduces to $x_i \sigma_{ii}/m_i$. Let us denote $v_i := \sigma_{ii}/m_i$ (the normalized variance). Then $F_i(x) = F_i(x_i) = x_i v_i$. It is convenient to order the risks according to the normalized variance values $v_i = F_i(1)$. Henceforth we suppose there is a strict ordering assumption to be denoted as non-degenerate hypothesis, so that $v_1 > v_2 > \ldots > v_n$.

The correspondence between points $x \in X^*$, the key parameter λ and the key functions $F_i(x)$ is captured by the following sequential procedure.

Let us start from the vertex **1**. According to the first golden rule, we look for $\max F_i(\mathbf{1}) = \max_i v_i = v_1$. The risk labelled 1 is the first to be reinsured. Accordingly $\lambda(\mathbf{1}) = v_1 = \max F_i(\mathbf{1})$. Then de Finetti points out that for any value $0 = \lambda(\mathbf{0}) \leq \lambda \leq \lambda(\mathbf{1})$ of the parameter λ , an optimal retention $x \in X^*$ is given by

$$x_i = \min\left\{\lambda/v_i, 1\right\}, \qquad i = 1, \dots, n \tag{2}$$

This means that if $\lambda < v_i$ then $x_i = \lambda/v_i$ so that $F_i(x) = x_i v_i = \lambda$, while if $\lambda > v_i$ then $x_i = 1$ so that $F_i(x) = v_i < \lambda$. Let $v_h < \lambda < v_{h-1}$, for h = 2, ..., n + 1. Coherently with the second golden rule, the first h - 1 risks obtain the biggest advantage λ at any x of the segment (or at any λ of the interval) and play the role of directional risks along the segment identified by the equations $F_i(x) = \lambda$, i = 1, ..., h - 1, so that $I(x) = \{1, ..., h - 1\} =: I^{h-1}$, while the other risks are frozen fully insured at the level 1 and do not belong to the set I(x).

Let us now check what happens at $\lambda = v_h$, $h \leq n$. It is clear that there $x_h = \lambda/v_h$ and also $F_h(x) = x_h v_h = \lambda$ while still $F_i(x) = x_i v_i = \lambda = v_h$ for i = 1, ..., h-1, and $F_i(x) = v_i < \lambda = v_h$ for i = h+1, ..., n.

But this means exactly that at the value $\lambda = v_h$ and at the corresponding optimal point x the third golden rule plays its role. There the matching happens so that risk h (responsible of the matching event) enters in the set of directional risks joining the other h-1 previously belonging to I^{h-1} . This implies that risk hbegins to be reinsured and stops to be frozen at the level 1 of full retention. We define $\lambda^h := v_h$ a matching value and the corresponding $x^h \in X^*$ a matching point.

Summing up, for any matching value λ^h , we have $\lambda^h = F_h(x^h) = 1 v_h$ (and so $\lambda^1 = \lambda(\mathbf{1}) = v_1$), while in general we have the following one-to-one correspondence between $x \in X^*$ and $\lambda^1 \ge \lambda \ge 0$

$$x_i = \min\left\{\lambda/v_i, 1\right\}, \quad i := 1, \dots, n, \qquad \lambda = \max x_i v_i \tag{3}$$

Geometrically, this means that X^* is described in the *n*-dimensional unit cube by a piecewise linear path, whose corner points x^h are found jointly with the corresponding value of the key parameter λ^h . The corner point x^h lies along the optimum direction given by the first h-1 risks, at the point where the common value of the their key functions is matched from below by the key function of risk h (while the other key functions have smaller values). In turn λ^h is the common value of the key functions involved. Assuming the non degenerate hypothesis, X^* consists of n segments (otherwise it consists of a number of segments equal to the number of different v_i values).

In Markowitz (2006) terminology the risks stay in one of three "states", according to the retention level: UP if the retention level is 1, DOWN if it is 0 and IN if it is strictly between 0 and 1. In the no correlation case the state transitions happen at each corner where risk h changes state from UP to IN. At the final corner point **0** all risks jointly collapse to the retention 0 and change state from IN to DOWN.

4. The general case of correlated risks in de Finetti's treatment

In his treatment of the general case with correlated risks de Finetti suggests that the properties of the optimum path fully mimick the pattern characterizing the no correlation case. The only difference is that we lack an a priori ordering of the risks because of a more complicated expression of the key functions

$$F_i(x) = \sum_{j=1}^n \frac{\sigma_{ij}}{m_i} x_j , \qquad i = 1, \dots, n,$$

replacing $F_i(x) = x_i \sigma_{ii}/m_i$. Once again, we start defining $\lambda^1 = \max_i F_i(\mathbf{1})$. Let k be the risk providing the largest value and $I^1 := \{k\}$. Then a sequence of corner values λ^h , corner points x^h and sets I^h of directional risks is recursively found for h = 2, ..., n. The direction of movement on X^* for $\lambda^h < \lambda \leq \lambda^{h-1}$ is described by the equations

$$\lambda = F_i(x), \qquad i \in I^{h-1}, \qquad \qquad x_i = 1, \qquad i \notin I^{h-1},$$

which define x parameterized by λ . The corner point x^h is still found on this direction, where one of the non directional risks (previously frozen at the level 1 of full retention) matches from below the common value of the key functions of the h-1 directional risks in I^{h-1} . The corner value λ^h is defined as the smallest λ^h such that

$$F_i(x(\lambda)) \le \lambda, \quad i \notin I^{h-1}$$
(4)

There, the new matching risk, providing the equality in (4), joins the set I^{h-1} of the previous directional risks, changing state from UP to IN, and defining the new set I^h .

As before, all corner points are of the same matching type and the set X^* is a piecewise linear path consisting of *n* segments (under a non degeneracy assumption that the equality in (4) is due to one risk only) in a one-to-one correspondence with the values of λ . Once again, all risks collapse jointly to the retention 0. Hence it could be said that the general case with correlation differs from that of no correlation only for the computational effort required to find X^* through a truly sequential procedure, whereas in the no correlation case the elegance and simplicity of the closed formulas (3) allows for an immediate computation.

We remark that de Finetti's approach assumes for granted that at any stage a matching event occurs, that is, a feasible corner point is found along the current optimum direction. It could be shown by counterexamples and has been pointed out by Markowitz that this does not always happen. Nevertheless, we can say that de Finetti worked with an implicit regularity assumption and gave a correct procedure to find the solution to the problem with correlated risks under this hypothesis of regularity. We repeat that the hypothesis consists in the existence of a feasible corner point, induced by a matching event, at any stage of his procedure. What if the assumption is not verified? We will treat the point in the next Section 5.

5. An extension of de Finetti's procedure to the general case

We start from the question left open at the end of the previous section. It is clear that the regularity is violated iff along the optimum path one of the currently directional risks reaches a boundary value of the retentions before a matching event happens. A boundary value is typically 0, even if 1 cannot a priori be excluded once we accept that the monotony of the retentions as functions of λ is no more granted under correlation.

Let us consider a point x^* , or the corresponding value λ^* , where the non regularity happens for the first time along the optimum path. Let us call *critical* such a point. This means that previously the conditions of regularity were holding. Then, without loss of generality, there were just h - 1 previous directional risks for which $F_i(x) = \lambda$, while for the other non directional risks, still at the level 1 of full retention, $F_i(x) < \lambda$. As a matching event on this current direction has not happened yet, these relations still hold at the critical point x^* (or critical value λ^*).

At this critical point the golden rules of de Finetti's procedure do not hold anymore. Indeed, while at this point the risks maximizing $F_i(x^*)$ are the current directional ones, we cannot move away from x^* along this direction, as dictated by the first and the second golden rule. Indeed advancing means going out of the unit cube because the critical risk becomes less than 0 or bigger than 1.

We are then faced with the need of defining a new procedure to leave x^* or more generally to deal with new situations not found in the regular case. As we already said, de Finetti did not treat the point, but, in our opinion, his tools (the key functions) and his logic (the golden rules), once properly adjusted, may nevertheless be successfully applied. In this section we present in some detail, but informally, and making large recourse to geometric intuitions, the adjusted procedure for the non regular case.

Let us first note that the critical point x^* is surely a corner point of the optimum path because there a compulsory change of direction happens. But this corner point is of a different type with respect to the ones found in the regular case. Those were driven by matching events. This was driven by what could be defined a *boundary event*. As the ordinary rule to leave x^* (or any critical point generated by a boundary event) cannot hold, we suggest, in the spirit of de Finetti's approach, to follow a new rule, as much as possible coherent with the ordinary golden rules of maximization of the advantage.

In order to find a new direction as close as possible to the old logic, the only possibility is that one of the old directional risks leave the previous directional set I, so that we remain with a set $I(x^*) \subset I$ of h-2directional risks (provided $I(x^*)$ is not empty; this case will be discussed later). Or, more formally, the new optimum direction from x^* is obtained by the solution of the set of h-2 equations $F_i(x) = \lambda$ for $i \in I(x^*)$, together with $x_i = 1$, for $i \notin I$. It remains to choose the risk to be excluded from I to obtain $I(x^*)$. Since we need one more equation to determine the new direction, it seems quite obvious (at least in case the boundary value of the critical risk is 0) that this must be just the critical risk, responsible of the boundary event, and that the *n*-th equation is just setting this variable to 0 or to 1, i.e., fixed to its boundary value.

Applying this adjustment of the golden rules, we find a nice symmetry between the rules holding in a corner point of a matching type and boundary type respectively. In the former case the critical (matching) risk joins, while in the latter the critical (boundary) risk leaves, the set of the previous directional risks. Now it is clear how to leave any type of corner point x^* (unless $I(x^*)$ is empty).

Let us discuss what happens to the excluded critical risk along the newly defined optimum path. If it is excluded at a level 1, of course its exclusion must be temporary. Waiting to regain an active role as a critical risk of a matching event for some lower value of λ , it is obvious that in the meantime its key function must satisfy $F_i(x) < \lambda$, coherently with the rule holding in the regular case for the risks not entered yet in reinsurance.

If a critical risk is excluded at a level 0, the temptation would be, at first glance, to consider this as a final exclusion, and to neglect the risk with no need to further keep into account the corresponding $F_i(x)$. Maybe surprisingly, it is not so. As shown in the example in Section 8, the exclusion at a level 0 may be temporary too. What happens in the meantime (or definitely if this is the case) to the excluded risk? Its key function must satisfy $F_i(x) > \lambda$. Indeed the exclusion is a final one iff for all remaining points of the optimum path $F_i(x) > \lambda$ for the critical risk *i*. And it is otherwise temporary up to the first *x* where $F_i(x) = \lambda$.

A comment is in order to explain the logic behind this statements. Clearly, for a risk i at the retention level 0 there is no intuitive reinsurance meaning for $F_i(x)$. We cannot have an additional reinsurance for this risks. But in a portfolio of reinsurance movements under correlation of risks, also risks currently at 0 may play a role. For risks of this type, coming back from full reinsurance surely increases the variance but also the expectation and we should look at the global effect of the combination with reinsurance movements of the other risks. Suppose ad absurdum it were $F_i(x) < \lambda$. Then a small step back from full reinsurance combined with reinsurance movements of the other directional risks would open the way for a diminution of the global normalized variance bigger than λ , a logical contradiction. And when $F_i(x) = \lambda$ for a lower value λ of the parameter, it is just the point to come back from full reinsurance realizing a matching event of an unusual type. This is a matching event where the critical matching risk is not coming from level 1 as in the regular case, but from level 0 and the corresponding key function now matches from above and not from below the key functions of the currently directional risks.

It is obvious that this adjusted procedure should be followed for any sequence of critical matching and boundary points. Before summarizing the results, we discuss the case when at a corner point of the boundary type the set $I(x^*)$ is empty. In this case the adjusted procedure cannot immediately help us in escaping from the critical point x^* . Indeed there are no available directional risks. Note that such a point corresponds to a vertex of the unit cube as $x_i^* \in \{0, 1\}$ for any *i* as a consequence of the vanishing of the directional risks and in particular the critical risk is surely of type 0 to avoid logical contradictions.

Consider the current value of λ at $x^* = x^h$. We label λ as $\overline{\lambda}^h$, instead of simply λ^h , for a reason which will be soon clear. According to our rules, we have for the set of risks at level 1 (surely non empty otherwise we had reached the final vertex **0**) $F_i(x^*) < \lambda$. For the set of risks at level 0, except for the critical one, $F_i(x^*) > \lambda$. And of course $F_i(x^*) = \lambda$ for the critical one at level 0. Now we must adjust the procedure by decreasing λ ; note that in the absence of directional risks the key functions stay all fixed at their level $F_i(x^*)$, so that decreasing λ has no effect until the decreasing parameter λ matches the biggest value among the $F_i(x^*)$ of the risks at level 1. This is to be seen as nothing but a new matching event. Accordingly the matching risk becomes the new singleton directional risk and things go on.

We may view this as a resetting of the procedure. Indeed the situation at this vertex looks like the one at the initial starting point, vertex 1, where we looked for $\max_i F_i(1)$ over the risks at level 1, i.e., among all risks. Here on the contrary we look for $\max_i F_i(x^*)$ only (once again) on those risks *i* currently at level 1 and

define $\underline{\lambda}^h$ that maximum. In both cases the maximizing risk is the first to be reinsured in the path starting from the vertex. It may be said that a boundary point of this vertex type reveals a stronger non regularity than the one characterizing other non vertex boundary points. Indeed in the first case only, the one-to-one correspondence between points of X^* and values of $0 \leq \lambda \leq 1$ is lost: the critical vertex x^h is associated to an interval $\overline{\lambda}^h \geq \lambda \geq \underline{\lambda}^h$.

Let us summarize the main results:

a) The procedure starts with the search for $\max_i F_i(1)$ which defines the first risk to be reinsured.

b) The optimum path is a piecewise linear path with possibly more than n segments. The corner points of the path play a key role. They may be generated by two types of events: matching events and boundary events.

c) In a matching event the key function of a critical matching risk, whose retention value is previously frozen at a fixed level (1 or 0), matches from below (for critical risks at level 1) or from above (for critical risks at level 0) the common value of the key function of the currently directional risks, whose retention value was along the previous segment variable. As a matching event happens, the matching risk (whatever its previous level 1 or 0) joins the set of previous directional risks. Its retention value starts changing, that is it begins to be reinsured if the previous level was 1, and on the contrary comes back from full reinsurance if its previous level was 0. The other risks do not change their state (directional or not) at the corner point.

d) In a boundary event the variable retention level of a risk belonging to the set of current directional risks reaches a boundary value (that is the retention 0 or 1) before a matching event happens. As a boundary event happens, the boundary risk leaves the set of directional risks. Its retention level becomes fixed at the reached boundary level. The other risks do not change their state at the corner point.

e) In case the set of remaining directional risks is not empty the way of leaving a corner point is the same for both types of corner points: move along the direction given by the set of equations $F_i(x) = \lambda$ for all *i* of the new set of directional risks (the other non directional being fixed at their respective level). In the internal points of each segment of the optimum path, besides being $F_i(x) = \lambda$ for all *i* of the set of directional risks it is also $F_i(x) < \lambda$ for all *i* of the set of risks fixed at the level 1 and $F_i(x) > \lambda$ for all risks fixed at level 0.

f) In case the set of remaining directional risks is empty, the corner point is a vertex of the unit cube. The one-to-one correspondence between $x \in X^*$ and λ is lost and in order to leave the vertex we need to decrease λ to the smallest value of the interval associated to the vertex, where one of the risks at level 1 realizes a new matching event.

Markowitz (2006) underlines that there are four types of corners according to the changes of state of the critical risk: a) from UP to IN, b) from IN to DOWN, c) from IN to UP, d) from DOWN to IN. By using this terminology our regularity condition corresponds to allowing transitions of type a) only. We may also characterize matching corners as transitions of type a) or d) and boundary corners as transitions of type b) or c).

A complex hypothesis of non degeneracy has been implicitly assumed overall the paper. It may be resumed by the condition that at any value of the parameter λ , where a transition of state happens, only one risk is in transition.

It is interesting to note that Markowitz qualifies the "backward transitions" (from IN to UP and from DOWN to IN) as unusual in de Finetti's model (Markowitz (2006) p. 14). He mentions them more in the form of a dubitative warning. He says that the possibility for a risk i to move from DOWN to IN or from IN to UP "may seem like unusual events in the de Finetti model; but we have no theorem that says it cannot happen, so the algorithm should be prepared to handle this cases". In particolar he exhibits a two dimensional example in which the case DOWN-IN occurs but with a negative expectation of the risk involved. We provide in Section 8 a three-dimensional example with a standard positive expectation for all risks, where a DOWN-IN

transition occurs. This reveals that there is no monotonic property of the reinsurance retentions along the optimum path given for granted in the proportional reinsurance problem. Note also that the presence of backward transitions is a sufficient but not necessary condition to reject monotony of the retentions.

A further terminology remark is in order, when we deal with monotony. We have often used a dynamic point of view in our discussion. This is not to be intended in a temporal sense. For example when we speak of further reinsurance starting from x we do not intend that a further reinsurance treaty is added to a previous already signed one. Also when we speak of coming back from full reinsurance we do not mean to retract part of a reinsurance treaty already signed. This is simply to understand the effect of modifying the reinsurance conditions holding at x. Of course it may well happen that an insurance company after having concluded a treaty may decide to arrange further reinsurance, perhaps because a new more risk averse manager has been appointed. And in this case a new problem, with initial situation the one reached through the first treaty should be analyzed, implying monotony at least with respect to the final position coming from the first treaty. But this has nothing to do with the problem studied by de Finetti.

6. A mathematical programming formulation

We now analyze the problem by mathematical programming techniques and show how the tools introduced by de Finetti, like the key functions and the key parameter, play a meaningful role and are the most natural way to build up a computational procedure to solve the optimum retention problem. This can be restated as the following quadratic problem

min
$$\frac{1}{2} x^T \mathbf{V} x$$

 $m^T x = E$
 $\mathbf{0} \le x \le \mathbf{1}$
(5)

for every attainable E, i.e. $0 \le E \le \sum_i m_i$. The strict convexity of the objective function guarantees that there is a one-to-one correspondence between points in X^* and optimal solutions of (5) for all attainable E. The Karush-Kuhn-Tucker conditions are necessary and sufficient for optimality of (5), since the constraints are regular and the objective function is strictly convex (see Shapiro (1979), Karush (1939), Kuhn and Tucker (1951)). The conditions are expressed through the Lagrangean function

$$L(x,\lambda,u,v) = \frac{1}{2} x^T \mathbf{V} x + \lambda \left(E - m^T x\right) + u \left(x - \mathbf{1}\right) - v x$$

and state that \hat{x} is optimal if and only if there exist Lagrange multipliers $(\hat{\lambda}, \hat{u}, \hat{v}), \hat{u} \ge 0, \hat{v} \ge 0$ such that:

According to the third condition, called *complementarity*, we have that $0 < x_j < 1$ implies $v_j = 0$ and $u_j = 0$. It is convenient to introduce a new variable defined as w := u - v and unconstrained in sign. Then the complementarity condition is equivalent to

 $\begin{array}{rrrr} x_j = 0 & \Longrightarrow & w_j \leq 0 \\ 0 < x_j < 1 & \Longrightarrow & w_j = 0 \\ x_j = 1 & \Longrightarrow & w_j \geq 0 \end{array}$

In order to verify 1) of (6), since x is unconstrained (in the Lagrangean minimization), it is enough to compute:

$$\frac{\partial L}{\partial x} = \mathbf{V} \, x - \lambda \, m + u - v = \mathbf{V} \, x - \lambda \, m + w = 0 \tag{7}$$

i.e., componentwise

$$\sum_{j} \sigma_{ij} x_j - \lambda m_i + w_i = 0 \qquad i = 1, \dots, m$$

or equivalently

$$\sum_{j} \frac{\sigma_{ij}}{m_i} x_j + \frac{w_i}{m_i} = \lambda \qquad i = 1, \dots, n$$

so that, by using the key functions, we need solving

$$F_i(x) + \frac{w_i}{m_i} = \lambda$$
 $i = 1, \dots, n$

It is convenient to partition the policies into three sets (possibly empty) as

$$I^0 := \{j : x_j = 0\}, I^1 := \{j : x_j = 1\}, I^* := \{j : 0 < x_j < 1\}$$

For the sake of notational simplicity, we omit to denote that these subsets actually depend on x.

Then the complementarity condition can be restated through the key functions in the following simple form:

Optimality condition: x is optimal if and only if there exists λ such that

$$F_i(x) = \lambda, \ i \in I^*, \quad F_i(x) \ge \lambda, \ i \in I^0, \quad F_i(x) \le \lambda, \ i \in I^1$$
(8)

In case of uncorrelated risks (8) simplifies

$$\frac{\sigma_{ii}}{m_i} x_i = \lambda, \ i \in I^*, \quad 0 \ge \lambda \ \text{if} \ I^0 \neq \emptyset, \quad \frac{\sigma_{ii}}{m_i} \le \lambda, \ i \in I^1$$
(9)

which implies $I^0 = \emptyset$ for E > 0. In this case de Finetti's analysis is always correct.

Let us now investigate the relation between E and λ . It is a known fact that the optimal value of (5), viewed as a function of E, is a non decreasing convex function (see Shapiro (1979)). Indeed, let f(E) be the optimal value of (5) as a function of E. Let x and x' be optima in (5) with r.h.s. E and E' respectively and let $0 < \alpha < 1$. Then we have by definition

$$\alpha f(E) + (1 - \alpha) f(E') = \frac{1}{2} (\alpha x^T \mathbf{V} x + (1 - \alpha) x'^T \mathbf{V} x') .$$
(10)

Moreover,

$$m^{T} \left(\alpha x + (1 - \alpha) x' \right) = \alpha E + (1 - \alpha) E'$$

so that the solution $(\alpha x + (1-\alpha) x')$ is feasible (but not necessarily optimal) in (5) with r.h.s $(\alpha E + (1-\alpha) E')$. Then

$$f(\alpha E + (1 - \alpha) E') \le \frac{1}{2} (\alpha x + (1 - \alpha) x')^T \mathbf{V} (\alpha x + (1 - \alpha) x') < \frac{1}{2} (\alpha x^T \mathbf{V} x + (1 - \alpha) x'^T \mathbf{V} x'),$$

where the second inequality follows from the strict convexity of the quadratic form. By combining this inequality and (10) we have the strict convexity of f(E). Moreover, the function f(E) is strictly increasing. Indeed, let x be the optimum with r.h.s. E and consider the solution αx with $0 < \alpha < 1$, so that αx is feasible with respect to the constraints $0 \le x \le 1$ and the constraint $\alpha m^T x = \alpha E$. Then we have

$$f(E) = \frac{1}{2} x^T \mathbf{V} x > \frac{1}{2} (\alpha x)^T \mathbf{V} (\alpha x) \ge f(\alpha E) ,$$

where the first inequality derives from the positive definiteness of the quadratic form together with $0 \le \alpha < 1$ and the second from the fact the $f(\alpha E)$ is the optimal value.

It is a known fact that the Lagrange multiplier λ is a subgradient of f(E) and is in fact its derivative if f(E) is differentiable at E (Shapiro (1979)). Hence $\lambda > 0$ for any E > 0 and also λ is a strictly increasing function of E (by strict convexity), although not necessarily continuous.

This implies that we may vary indirectly E by varying instead λ in (8). This justifies the previous introduction of λ as a key parameter.

7. A computational procedure

The conditions of Section 6 suggest the following extension of de Finetti's approach to the general case: the starting point is the optimum $x = \mathbf{1}$ (and therefore $E = m^T \mathbf{1}$, $I^0 = I^* = \emptyset$). Then (8) is

$$F_i(\mathbf{1}) \le \lambda \qquad \forall i \implies \lambda = \max_i F_i(\mathbf{1}).$$
 (11)

This is the largest value for λ , corresponding to $E = m^T \mathbf{1}$. Let k be an index for which the maximum in (11) is attained. Then we update the sets I^1 and I^* by moving k from I^1 to I^* .

Then we recursively carry out the following computations until $\lambda = 0$.

If $I^* \neq \emptyset$, the optimal x as function of λ , given by $F_i(x) = \lambda$, $i \in I^*$, is computed by first solving the two linear systems (in the variables y and z respectively)

$$\sum_{j \in I^*} \frac{\sigma_{i,j}}{m_i} y_j = 1, \ i \in I^*, \quad \text{and} \quad \sum_{j \in I^*} \frac{\sigma_{i,j}}{m_i} z_j = \sum_{j \in I^1} \frac{\sigma_{i,j}}{m_i}, \ i \in I^*,$$
(12)

and then combining

$$x_j = \lambda y_j - z_j, \qquad j \in I^* \tag{13}$$

Now we check which condition is violated first by decreasing λ ,

$$F_i(x) \ge \lambda, \ i \in I^0, \quad 0 \le x_i \le 1, \ i \in I^*, \quad F_i(x) \le \lambda, \ i \in I^1,$$

$$(14)$$

and then recompute the sets I^0 , I^* , I^1 . For the sake of notational simplicity let, for $i \in I^0$,

$$\psi_i := \sum_{j \in I^*} \frac{\sigma_{i,j}}{m_i} \, y_j, \qquad \varphi_i := \sum_{j \in I^*} \frac{\sigma_{i,j}}{m_i} \, z_j, \qquad \rho_i := \sum_{j \in I^1} \frac{\sigma_{i,j}}{m_i}$$

Then the conditions are checked by using (13) as follows: for $i \in I^0$ we have

$$F_i(x(\lambda)) = \sum_{j \in I^*} \frac{\sigma_{i,j}}{m_i} \left(\lambda \, y_j - z_j\right) + \sum_{j \in I^1} \frac{\sigma_{i,j}}{m_i} = \psi_i \, \lambda - \varphi_i + \rho_i \ge \lambda$$

from which

$$(\psi_i - 1) \lambda \ge \varphi_i - \rho_i$$

and so the value of λ for which $F_i(x) = \lambda$, $i \in I^0$, is given by

$$\lambda_0 = \max\left\{\frac{\varphi_i - \rho_i}{\psi_i - 1} : i \in I^0, \ \psi_i - 1 > 0\right\} .$$

Similarly for $i \in I^1$ we have

$$(1 - \psi_i) \lambda \ge \rho_i - \varphi_i$$

and so the value of λ for which $F_i(x) = \lambda$, $i \in I^1$, is given by

$$\lambda_1 = \max\left\{\frac{\varphi_i - \rho_i}{\psi_i - 1} : i \in I^1, \ \psi_i - 1 < 0\right\}.$$

To check $0 \le x_i \le 1$ we simply compute

$$\lambda y_i - z_i = 0 \implies \lambda_{*0} = \max\left\{\frac{z_i}{y_i} : i \in I^*, y_i > 0\right\},$$
$$\lambda y_i - z_i = 1 \implies \lambda_{*1} = \max\left\{\frac{1 + z_i}{y_i} : i \in I^*, y_i < 0\right\}.$$

Then the corner value λ^h is given by

$$\lambda^h := \max \left\{ \lambda_0, \lambda_1, \lambda_{*0}, \lambda_{*1} \right\}$$

and the corner points x^h are given by $x^h_j = \lambda^h y_j - z_j, \ j \in I^*, \ x^h_j = 0, \ j \in I^0, \ x^h_j = 1, \ j \in I^1.$

We assume that there is no degeneracy in the data. In other words we assume that all maxima in the above expressions are achieved by a unique index. Let k be this index. The sets are recomputed as follows: if $\lambda^h = \lambda_0$ then k moves from I^0 to I^* ; if $\lambda^h = \lambda_1$ then k moves from I^1 to I^* ; if $\lambda^h = \lambda_{*0}$ then k moves from I^* to I^0 ; if $\lambda^h = \lambda_{*1}$ then k moves from I^* to I^1 . In the first two cases x^h is a matching point and in the other two cases x^h is a boundary point.

If $I^* = \emptyset$ (and necessarily x is a vertex of the unit cube) then (8) simplifies to $F_i(x) \ge \lambda$, $i \in I^0$, $F_i(x) \le \lambda$, $i \in I^1$, from which we only have to recompute $\lambda := \max \{F_i(x) : i \in I^1\}$ (with the discontinuity already mentioned). Note that this implies that, whenever $I^* = \emptyset$, the index entering I^* comes from I^1 . As previously remarked, the interval

$$\min\left\{F_i(x^h): i \in I^0\right\} \ge \lambda \ge \max\left\{F_i(x^h): i \in I^1\right\}$$
(15)

is associated to the same vertex corner point x^h . For the two special corner values, extremes of the interval (15), we adopt the following notation:

$$\overline{\lambda}^h := \min\left\{F_i(x^h) : i \in I^0\right\}, \qquad \underline{\lambda}^h := \max\left\{F_i(x^h) : i \in I^1\right\}$$

We briefly comment on the possibility that the index k providing the maximum is not unique. This means that the violation of the complementarity is simultaneously due to different conditions for the same value of λ (so called *degeneracy*). We may take as k any index providing the maximum and continue the procedure as if it were unique. It may happen that there is a sequence of steps such that the sets I^0 , I^* , I^1 change at each step while λ remains constant. This is similar to the degeneracy of the simplex steps in linear programming (see Dantzig (1963)) and therefore we should take care of possible cycling of the method, in the sense that the same pattern of changes of the three sets is repeated for ever while λ remains constant. We do not dwell with this technical problem here.

8. A critical example

Let us consider a case with three risks whose expected profit and covariance matrix are

$$m = \begin{pmatrix} 1\\1\\1 \end{pmatrix}, \qquad \mathbf{V} = \frac{1}{1000} \begin{pmatrix} 160 & 96 & 24\\96 & 90 & 27\\24 & 27 & 40 \end{pmatrix},$$

so the key functions are

$$F_1(x) = \frac{1}{1000} (160 x_1 + 96 x_2 + 24 x_3)$$

$$F_2(x) = \frac{1}{1000} (96 x_1 + 90 x_2 + 27 x_3) ,$$

$$F_3(x) = \frac{1}{1000} (24 x_1 + 27 x_2 + 40 x_3) .$$

For $E = \sum_{j} m_j = 3$, the optimal retention is x = 1. The key function values in the starting point 1 are

$$F_1(\mathbf{1}) = \frac{280}{1000} = \frac{7}{25}, \qquad F_2(\mathbf{1}) = \frac{213}{1000}, \qquad F_3(\mathbf{1}) = \frac{91}{1000}$$

The condition

$$\lambda(\mathbf{1}) = \lambda^1 = \max_i F_i(\mathbf{1})$$

gives

$$\lambda^1 = \frac{7}{25}, \qquad I^* = \{1\}, \quad I^1 = \{2,3\}, \quad I^0 = \emptyset$$

Hence the risk 1 is a directional risk (the only one at the moment). The direction is given by $F_1(x_1, 1, 1) = \lambda$ from which we get

$$x_1 = \frac{1000\,\lambda - 96 - 24}{160} \; .$$

The conditions in (0) correspond to the inequalities:

$$0 \le x_1 \le 1 \qquad \Longrightarrow \qquad 0 \le \frac{1000\,\lambda - 96 - 24}{160} \le 1 \qquad \Longrightarrow \qquad \frac{3}{25} \le \lambda \le \frac{7}{25} \le F_2(x) = \frac{1}{1000} \left(96\,\frac{1000\,\lambda - 96 - 24}{160} + 90 + 27\right) \le \lambda \qquad \Longrightarrow \qquad \lambda \ge \frac{9}{80} \,,$$
$$F_3(x) = \frac{1}{1000} \left(24\,\frac{1000\,\lambda - 96 - 24}{160} + 27 + 40\right) \le \lambda \qquad \Longrightarrow \qquad \lambda \ge \frac{49}{850} \,.$$

To check which condition is violated first by decreasing λ we compute

$$\max\left\{\frac{3}{25} \ ; \ \frac{9}{80} \ ; \ \frac{49}{850}\right\} = \frac{3}{25}$$

This means that we have reached a boundary vertex corner $x^2 = (0, 1, 1)$ and the sets must be changed to $I^* = \emptyset$, $I^0 = \{1\}$, $I^1 = \{2, 3\}$. Then we have the special corner value $\overline{\lambda}^2 = 3/25$. At x^2 the key functions have values

$$F_1(0,1,1) = \frac{120}{1000} = \frac{3}{25}$$
, $F_2(0,1,1) = \frac{117}{1000}$, $F_3(0,1,1) = \frac{67}{1000}$

and λ is reset to the value

$$\underline{\lambda}^2 = \max\left\{\frac{117}{1000} \ ; \ \frac{67}{1000}\right\} = \frac{117}{1000}$$

Moving away from x^2 the index sets become $I^0 = \{1\}$, $I^* = \{2\}$, $I^1 = \{3\}$, with risk 2 being the directional risk. The direction is given by $F_2(0, x_2, 1) = \lambda$, from which we get

$$\frac{1}{1000} (90 x_2 + 27) = \lambda \implies x_2 = \frac{1000 \lambda - 27}{90}$$

Now the conditions in (0) correspond to the inequalities:

$$0 \le x_2 \le 1 \qquad \Longrightarrow \qquad \frac{27}{1000} \le \lambda \le \frac{117}{1000} ,$$

$$F_1(0, x_2, 1) = \frac{1}{1000} \left(96 \frac{1000 \,\lambda - 27}{90} + 24 \right) \ge \lambda \qquad \Longrightarrow \qquad \lambda \ge \frac{9}{125} ,$$

$$F_3(0, x_2, 1) = \frac{1}{1000} \left(27 \frac{1000 \,\lambda - 27}{90} + 40 \right) \le \lambda \qquad \Longrightarrow \qquad \lambda \ge \frac{319}{7000} .$$

Since

$$\max\left\{\frac{27}{1000} \ ; \ \frac{9}{125} \ ; \ \frac{319}{7000}\right\} = \frac{9}{125} \ ,$$

this means that we have reached a matching corner, $x^3 = (0, 1/2, 1)$, with matching value $\lambda^3 = 9/125$. Moving away from x^3 the index sets become $I^* = \{1, 2\}$, $I^0 = \emptyset$, $I^1 = \{3\}$, and we have the two directional risks 1 (coming back from full retention) and 2. To find the direction we have to solve the system

$$F_1(x_1, x_2, 1) = \lambda$$
, $F_2(x_1, x_2, 1) = \lambda$,

i.e., from (12) and (13)

$$\frac{1}{1000} \begin{pmatrix} 160 & 96\\ 96 & 90 \end{pmatrix} \begin{pmatrix} y_1\\ y_2 \end{pmatrix} = \begin{pmatrix} 1\\ 1 \end{pmatrix}, \qquad \frac{1}{1000} \begin{pmatrix} 160 & 96\\ 96 & 90 \end{pmatrix} \begin{pmatrix} z_1\\ z_2 \end{pmatrix} = \frac{1}{1000} \begin{pmatrix} 24\\ 27 \end{pmatrix}$$

from which

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -\frac{125}{108} \\ \frac{1000}{81} \end{pmatrix} \lambda - \begin{pmatrix} -\frac{1}{12} \\ \frac{7}{18} \end{pmatrix} .$$

Now the conditions in (0) correspond to the inequalities:

$$0 \le x_1 \le 1 \quad \Longrightarrow \quad -\frac{99}{125} \le \lambda \le \frac{9}{125} ,$$
$$0 \le x_2 \le 1 \quad \Longrightarrow \quad \frac{63}{2000} \le \lambda \le \frac{9}{80} ,$$
$$F_3(x_1, x_2, 1) = \frac{1}{1000} \left(24 \left(-\frac{125}{108} \lambda + \frac{1}{12} \right) + 27 \left(\frac{1000}{81} \lambda - \frac{7}{12} \right) + 40 \right) \le \lambda \quad \Longrightarrow \quad \lambda \ge \frac{567}{12500}$$

Since

$$\max\left\{-\frac{99}{125} \; ; \; \frac{63}{2000} \; ; \; \frac{567}{12500}\right\} = \frac{567}{12500}$$

this means that we have reached a matching corner $x^4 = (37/1200, 77/450, 1)$ with matching value $\lambda^4 = 567/12500$. In order to leave x^4 the index sets become $I^* = \{1, 2, 3\}$. Hence all risks are directional and the direction is given by the solution of $\mathbf{V} x = \lambda \mathbf{1}$, which provides the last segment of X^* .

$$x = \begin{pmatrix} \frac{4625}{6804} \\ \frac{2750}{729} \\ \frac{12500}{567} \end{pmatrix} \lambda \,.$$

Finally for $\lambda = 0$ we obtain the end point of X^* , i.e. the vertex **0**. Summing up, the optimal x_i and the optimal expected profit E as functions of λ are (note the non monotonic behaviour of x_1):

	x_1	x_2	x_3	E
$\lambda = \frac{7}{25}$	1	1	1	3
$\frac{7}{25} \geq \lambda \geq \frac{3}{25}$	$\frac{25}{4}\lambda-\frac{3}{4}$	1	1	$\frac{25}{4}\lambda+\frac{5}{4}$
$\frac{3}{25} \geq \lambda \geq \frac{117}{1000}$	0	1	1	2
$\frac{117}{1000} \geq \lambda \geq \frac{9}{125}$	0	$\frac{100}{9}\lambda-\frac{3}{10}$	1	$\frac{100}{9}\lambda+\frac{7}{10}$
$\frac{9}{125} \geq \lambda \geq \frac{567}{12500}$	$- \frac{125}{108} \lambda + \frac{1}{12}$	$\frac{1000}{81}\lambda-\frac{7}{18}$	1	$\frac{3625}{324}\lambda+\frac{25}{36}$
$\frac{567}{12500} \ge \lambda \ge 0$	$\frac{4625}{6804}\lambda$	$\frac{2750}{729}\lambda$	$\frac{12500}{567}\lambda$	$\frac{540875}{20412}\lambda$

In Figures 1 and 2 we show the non monotonic function $x_1(\lambda)$ and the set X^* within the unit cube (note that X^* consists of four segments, i.e., one more than the number of risks).



It is interesting to note that this is an example of a non regular case where the last segment is fully internal, that is all variable jointly reach the value 0 (but not for the first time at least for the risk 1). We also summarize the set of transitions involved.

type of transition	risks involved	corner point x^h	corner value λ^h
UP-IN	1	(1, 1, 1)	7/25
IN-DOWN	1	(0, 1, 1)	3/25
UP-IN	2	(0, 1, 1)	117/1000
DOWN-IN	1	(0, 1/2, 1)	72/1000
UP-IN	3	(37/1200, 77/450, 1)	567/125000
IN-DOWN	1,2,3	(0, 0, 0)	0

This shows that the internality of the last segment is a necessary but not a sufficient condition for regularity. Hence it may be said that the problem of regularity goes far beyond the simple behaviour of the last segment on which the criticism of Markowitz (2006 p.6), regarding the so called last segment conjecture, was mainly based.

9. Conditions for regularity

Let us restate here the concept of regularity for an optimum path:

Definition: The optimal path X^* is regular if at corner points risks move from I^1 to I^* only.

In other words regularity implies that the violation of the complementarity conditions (8) is always referred to $F_i(x) \leq \lambda$. As already stated, no correlation among the risks guarantees regularity of the optimum path. However, regularity can be satisfied even if the risks are correlated provided certain conditions are met. In order to investigate these conditions it is convenient to rewrite m as $m_i = \eta_i \sigma_i$ (with $\sigma_i = \sqrt{\sigma_{ii}}$), where each η_i plays the role of a safety loading coefficient under a standard deviation premium principle. Then it is $m = \Sigma \eta$, with $\Sigma := \text{diag} \{\sigma_1, \dots, \sigma_n\}$. Moreover we have

$$\mathbf{R} = \Sigma^{-1} \, \mathbf{V} \, \Sigma^{-1}$$

where $\mathbf{R} := \{r_{ij} : i, j = 1, ..., n\}$ is the correlation matrix, i.e.

$$r_{ij} = \begin{cases} 1 & \text{if } i = j \\ \\ \frac{\sigma_{ij}}{\sigma_i \, \sigma_j} & \text{if } i \neq j \end{cases}$$

It is possible to provide sufficient conditions for regularity under the following assumptions: $r_{ij} = r \ge 0$ for all $i \ne j$ in the correlation matrix **R** and $\eta_i/\eta_j \le K$ for all i, j, where K is a positive constant. For instance, if all η_i are equal then K = 1. One preliminary result (proved in the Appendix) is given by:

Proposition 1: If rn(K-1) < 1-r, then no index $j \in I^*$ can move to I^1 by decreasing λ .

Note that in the two particular cases of all η_i equal or r = 0, the condition is fulfilled. Our main result (proved in the Appendix) is:

Proposition 2: If
$$3 n r (K-1) < 1-r$$
, the optimum path X^* is regular.

The bound provided in Proposition 2 is not tight. The various inequalities used to derive it are not tight and therefore the result is somehow pessimistic. However, it does reveal a qualitative behaviour of the relationship between r and K for the regularity to hold. We have carried out an experimental test for n = 10, n = 40 and n = 100. We have chosen the following grid of values for r,

$$r \in \mathbf{0.01}, 0.02, 0.03, 0.04, 0.05, 0.075, 0.1, 0.125, 0.15, \\0.175, 0.2, 0.25, 0.3, 0.35, 0.4, 0.5, 0.6, 0.8\}$$

For each value of r we have found by binary search the maximum value of K such that the following regularity test is successful: for each pair (r, K) up to 100 random instances are generated and in each instance $\eta_1 := 1$ and $\eta_2 := K$ while the values of $\eta_i \in [1, K]$ for i = 3, ..., n, and $\sigma_i \in [5, 10]$ for i = 1, ..., n, are randomly chosen. If all 100 instances exhibit a regular behavior then the test for the pair (r, K) is declared successful, otherwise the test for the pair (r, K) is stopped at the non regular instance and the test for the pair (r, K)is declared unsuccessful. The binary search is carried out until the tolerance interval for K has length 0.01.

In Figure 3 we show both the curves corresponding to the bound in Proposition 2, i.e. the equation 3nr(K-1) = 1 - r for the values n = 10, n = 40, n = 100, and also the points found by the binary search.

I

As expected the bound is not tight, in the sense that there are values not satisfying the bound for which the procedure exhibited a regular behavior. However, the qualitative hyperbolic relation between r and Kis reflected both in the bound and in the computed points.



Fig. 3: test results for n = 10, n = 40 and n = 100

If we multiply the bound by the constant 5 we find an interesting and rather curious match between the experimental points and the theoretical bound. Empirically we may therefore substitute the bound in Proposition 2 with 3 n r (K-1) < 5 (1-r) (see the new curve and the experimental points in Figure 4



Fig 4: modified bound for n = 10, n = 40 and n = 100

10. Connections with portfolio theory and CAPM

We would also like to acknowledge a suggestion by Bühlmann (2006), non anonymous referee of our paper, which paved us the way to offer an interesting alternative interpretation of the optimality conditions holding under regularity. Denoting by G_i the random profit coming from an integral retention of the *i*-th risk, the suggestion is that in case of no correlation the set of Pareto efficient retention quotas (seen as a function of the non negative parameter λ) should satisfy for each *i*, $V(x_i G_i)/E(x_i G_i) \leq \lambda$ with $x_i = 1$ if the inequality holds in strict form. In case of correlation an analogous condition is suggested replacing only the numerator of the left side of the (in)equalities with $cov(x_i G_i, \sum x_i G_i)$. The conditions are clearly equivalent (as long as $x_i > 0$ for all *i*, that is except for the point **0** of full reinsurance of all risks) to those found by de Finetti. In particolar, under correlation they hold only in the regular case.

Let us give here just a sketch of their actuarial meaning: once we fix a threshold level λ of acceptable normalized risk (variance measured in units of the numeraire, the expectation), efficient proportional reinsurance strategies are those that, whenever needed, cut down through reinsurance the normalized risk of each policy at the treshold, while do not touch those already under the treshold. In case of no correlation the proper (before normalization) risk measure of each policy is its variance (precisely the variance of its random gain after reinsurance); in the correlation case it is (after reinsurance too) its covariance with a reference portfolio. For any given retention choice, the reference portfolio is simply the one defined by the retentions. That this measure of risk may play a role in case of correlation is no more surprising, at least after the diffusion in the actuarial world, mainly due to the path breaking papers Borch (1974) and Bühlmann and Gerber (1978), of the results of the capital asset pricing model (Sharpe (1964), Lintner (1965), Mossin (1966)). In the essence this key role comes from being $V(\sum_i x_i G_i) = \sum_i \operatorname{cov}(x_i G_i, \sum_i x_i G_i)$, which shows that the contribution of each policy (after reinsurance) to the global amount of risk retained (the global variance after reinsurance) is precisely $\operatorname{cov}(x_i G_i, \sum_i x_i G_i)$. We do not want to enter here in more details about this very interesting point, but we signal that an early discussion of connections between de Finetti's work and portfolio theory could be found in Pressacco (1986) as quoted in Rubinstein (2006 a).

11. Conclusions

It is today plainly recognized that de Finetti introduced the mean-variance approach in financial problems under uncertainty and, treating a problem of proportional reinsurance, offered a correct procedure to find the mean-variance Pareto efficient set in case of no correlation. As far as de Finetti's contribution in the general case of correlation is concerned, this is still an open question. We show in this paper that de Finetti gave a fully correct procedure (plain extension of the one offered for the no correlation case), able to solve the problem with correlated risks only under a convenient regularity hypothesis. Quite likely, convinced that in the practical actuarial applications the non regularity could not be relevant, he did not discuss a procedure of general validity (working also in the non regular case).

To fill this gap and to complete the picture, we offer here a natural adjustment of de Finetti's procedure, which, respecting his logic and applying his simple and original tools, provides a correct solution to find the optimum set also when non regularity holds.

It is well known that (as discussed in some details in Markowitz (2006)) the mean-variance proportional reinsurance problem may also be solved through an application of modern constrained optimization techniques. We show as well that de Finetti's tools play also in this advanced frame a meaningful role and reveal to be the most natural way to build up a computational procedure coherent with the constrained optimization techniques.

We also discuss sufficient conditions for the regularity to hold also in case of correlated risks.

Appendix

Here we provide the proof of Propositions 1 and 2 previously stated.

Proof of Proposition 1. Suppose that for a value λ there are sets $I^0 = \emptyset$, I^* and I^1 such that the optimality condition is verified. Let us partition the matrix **V** into submatrices \mathbf{V}_1^1 , \mathbf{V}_*^1 , \mathbf{V}_*^* , \mathbf{V}_*^* , and similarly the vectors a, w and x into subvectors a_1 , a_* , w_1 , w_* , x_1 and x_* respectively, where the superscript refers to the column indices and the subscript to the row indices. Then the optimality condition can be rewritten (note that we are assuming $I^0 = \emptyset$)

$$\mathbf{V}_{*}^{*} x_{*} + \mathbf{V}_{*}^{1} \mathbf{1} = \lambda a_{*}, \qquad \mathbf{V}_{1}^{*} x_{*} + \mathbf{V}_{1}^{1} \mathbf{1} \le \lambda a_{1}$$

and so

$$x_* = \mathbf{V}_*^{*-1}(-\mathbf{V}_*^1 \,\mathbf{1} + \lambda \,a_*)$$

Then the various conditions become

$$\begin{aligned}
x_{i}^{*} \geq 0, & i \in I^{*} & \text{if} \quad \mathbf{V}_{*}^{*-1}\mathbf{V}_{*}^{1}\mathbf{1} \leq \lambda \mathbf{V}_{*}^{*-1}a_{*} \\
x_{i}^{*} \leq 1, & i \in I^{*} & \text{if} \quad \lambda \mathbf{V}_{*}^{*-1}a_{*} \leq \mathbf{1} + \mathbf{V}_{*}^{*-1}\mathbf{V}_{*}^{1}\mathbf{1} \\
F_{i}(x) \leq \lambda, & i \in I^{1} & \text{if} \quad -\mathbf{V}_{1}^{*}\mathbf{V}_{*}^{*-1}\mathbf{V}_{*}^{1}\mathbf{1} + \mathbf{V}_{1}^{1}\mathbf{1} \leq \\
& \lambda(a_{1} - \mathbf{V}_{1}^{*}\mathbf{V}_{*}^{*-1}a_{*})
\end{aligned} \tag{16}$$

If all entries of $\mathbf{V}_*^{*-1} a_*$ are non-negative, then the condition $x_i^* \leq 1$ can never be violated by decreasing λ . So we first investigate about the question $\mathbf{V}_*^{*-1} a_* \geq 0$. We rewrite $\mathbf{V}_*^* = \Sigma_*^* \mathbf{R}_*^* \Sigma_*^*$, $a^* = \Sigma_*^* \eta_*$ so that

$$\mathbf{V}_{*}^{*-1} a_{*} = \Sigma_{*}^{*-1} \mathbf{R}_{*}^{*-1} \eta_{*}$$

and $\Sigma_*^{*-1} \mathbf{R}_*^{*-1} \eta_* > 0$ if and only if $\mathbf{R}_*^{*-1} \eta_* > 0$.

So we have (with $m = |I^*|$)

$$\mathbf{R}^*_* = \begin{pmatrix} 1 & r & \dots & r \\ r & 1 & \dots & r \\ \dots & \dots & \dots \\ r & r & \dots & 1 \end{pmatrix}$$

and

$$\mathbf{R}_{*}^{*-1} = \frac{1}{\Delta} \begin{pmatrix} 1 + (m-2)r & -r & \dots & -r \\ -r & 1 + (m-2)r & \dots & -r \\ \dots & \dots & \dots & \dots \\ -r & -r & \dots & 1 + (m-2)r \end{pmatrix}$$

with

$$\Delta = 1 + (m-2)r - (m-1)r^2 = (1-r)(r(m-1)+1)$$

Then $\mathbf{R}_*^{*-1}\eta_* > 0$ if and only if, for all $i \in I^*$,

$$(1 + (m - 1)r)\eta_i - r\sum_{j \in I^*} \eta_j > 0$$

i.e. if r > 0

$$\frac{\eta_i}{\sum_{j \in I^*} \eta_j} > \frac{r}{1 + (m-1)r}$$

Since

$$\frac{\eta_i}{\sum_{j \in I^*} \eta_j} = \frac{1}{\sum_{j \in I^*} \frac{\eta_j}{\eta_i}} \ge \frac{1}{mK}$$
$$\frac{1}{mK} > \frac{r}{1 + (m-1)r}$$

 $1 - r > r m \left(K - 1 \right)$

i.e.

which in turn is fulfilled if (since $n \ge m$ and $K \ge 1$)

$$1 - r > r n \left(K - 1 \right) \tag{17}$$

Hence we may state

the condition is fulfilled if

Proposition 1: If (17) holds, then no index $j \in I^*$ can move to I^1 by decreasing λ .

Proof of Proposition 2. Now we consider the condition $x^* \ge 0$. By using the expressions $\mathbf{V}^*_* = \Sigma^*_* \mathbf{R}^*_* \Sigma^*_*$, $a^* = \Sigma^*_* \eta_*, \mathbf{V}^1_* = \Sigma^*_* \mathbf{R}^1_* \Sigma^1_1$, the condition in (16) can be expressed as

$$\Sigma_{*}^{*-1} \mathbf{R}_{*}^{*-1} \Sigma_{*}^{*-1} \Sigma_{*}^{*} \mathbf{R}_{*}^{1} \Sigma_{1}^{1} \mathbf{1} \leq \lambda \Sigma_{*}^{*-1} \mathbf{R}_{*}^{*-1} \Sigma_{*}^{*-1} \Sigma_{*}^{*} \eta_{*}$$

i.e.,

$$\Sigma_*^{*-1} \mathbf{R}_*^{*-1} \mathbf{R}_*^1 \Sigma_1^1 \mathbf{1} \le \lambda \Sigma_*^{*-1} \mathbf{R}_*^{*-1} \eta_*$$

equivalent to, since Σ^*_* is a positive diagonal matrix,

$$\mathbf{R}_{*}^{*-1} \mathbf{R}_{*}^{1} \Sigma_{1}^{1} \mathbf{1} \leq \lambda \, \mathbf{R}_{*}^{*-1} \, \eta_{*} \tag{18}$$

The condition $F_i(x) \leq \lambda$, after replacing in (16) $\mathbf{V}^1_* = \Sigma^*_* \mathbf{R}^1_* \Sigma^1_1$, $\mathbf{V}^*_1 = \Sigma^1_1 \mathbf{R}^*_1 \Sigma^*_*$, $\mathbf{V}^1_1 = \Sigma^1_1 \mathbf{R}^1_1 \Sigma^1_1$, $\mathbf{V}^{*-1}_* = \Sigma^{*-1}_* \mathbf{R}^{*-1}_* \Sigma^{*-1}_*$, becomes

$$-\Sigma_{1}^{1} \mathbf{R}_{1}^{*} \mathbf{R}_{*}^{*-1} \mathbf{R}_{*}^{1} \Sigma_{1}^{1} \mathbf{1} + \Sigma_{1}^{1} \mathbf{R}_{1}^{1} \Sigma_{1}^{1} \mathbf{1} \le \lambda (\Sigma_{1}^{1} \eta_{1} - \Sigma_{1}^{1} \mathbf{R}_{1}^{*} \mathbf{R}_{*}^{*-1} \eta_{*})$$

i.e., since Σ_1^1 is positive diagonal

$$-\mathbf{R}_{1}^{*}\mathbf{R}_{*}^{*-1}\mathbf{R}_{*}^{1}\Sigma_{1}^{1}\mathbf{1} + \mathbf{R}_{1}^{1}\Sigma_{1}^{1}\mathbf{1} \le \lambda(\eta_{1} - \mathbf{R}_{1}^{*}\mathbf{R}_{*}^{*-1}\eta_{*})$$
(19)

Let $b := \mathbf{R}_*^{*-1} \mathbf{R}_*^1 \Sigma_1^1 \mathbf{1}, c := \mathbf{R}_*^{*-1} \eta_*, d := \mathbf{R}_1^1 \Sigma_1^1 \mathbf{1}$, so that (18) and (19) can be written as

$$b \le \lambda c, \qquad -\mathbf{R}_1^* b + d \le \lambda (\eta_1 - \mathbf{R}_1^* c)$$
⁽²⁰⁾

We can compute

$$b = \frac{1}{\Delta} \begin{pmatrix} 1 + (m-2)r & -r & \dots & -r \\ -r & 1 + (m-2)r & \dots & -r \\ \dots & \dots & \dots & \dots & \dots \\ & -r & 1 + (m-2)r \end{pmatrix} \cdot \\ \begin{pmatrix} r & r & \dots & r \\ r & r & \dots & r \\ \ddots & \ddots & \cdots & n \end{pmatrix} \Sigma_{1}^{1} \mathbf{1} \implies \\ b_{i} = \frac{r(1-r)}{\Delta} \sum_{j \in I^{1}} \sigma_{j} = \frac{r}{r(m-1)+1} \sum_{j \in I^{1}} \sigma_{j}, \quad i \in I^{*} \\ c_{i} = \frac{\eta_{i} (1+r(m-1)) - r \sum_{j \in I^{*}} \eta_{j}}{\Delta}, \quad i \in I^{*} \\ d = \begin{pmatrix} 1 & r & \dots & r \\ r & 1 & \dots & r \\ \dots & \dots & 1 \end{pmatrix} \Sigma_{1}^{1} \mathbf{1} \Longrightarrow d_{i} = (1-r) \sigma_{i} + r \sum_{j \in I^{1}} \sigma_{j}, \quad i \in I^{1} \\ (d - \mathbf{R}_{1}^{*} b)_{i} = (1-r) \sigma_{i} + r \sum_{j \in I^{1}} \sigma_{j} - mr \frac{r}{r(m-1)+1} \sum_{j \in I^{1}} \sigma_{j} = \\ (1-r) \sigma_{i} + (1 - \frac{mr}{r(m-1)+1}) r \sum_{j \in I^{1}} \sigma_{j} = \\ (1-r) (\sigma_{i} + \frac{r(1-r)}{r(m-1)+1} \sum_{j \in I^{1}} \sigma_{j}) \\ (\eta_{1} - \mathbf{R}_{1}^{*} c)_{i} = \eta_{i} - r \frac{\sum_{j \in I^{*}} \eta_{i} (1+r(m-1)) - mr \sum_{j \in I^{*}} \eta_{j}}{\Delta} = 0$$

$$(1 - r)\sigma_{i} + (1 - r(m - 1) + 1)r\sum_{j \in I^{1}} \sigma_{j} = (1 - r)(\sigma_{i} + \frac{r(1 - r)}{r(m - 1) + 1}\sum_{j \in I^{1}} \sigma_{j})$$
$$-\mathbf{R}_{1}^{*}c)_{i} = \eta_{i} - r\frac{\sum_{j \in I^{*}} \eta_{i}(1 + r(m - 1)) - mr\sum_{j \in I^{*}} \eta_{j}}{\Delta}$$
$$\eta_{i} - \frac{r}{r(m - 1) + 1}\sum_{j \in I^{*}} \eta_{i}$$

Then the first inequality in (20) becomes

$$\frac{r\left(1-r\right)}{\Delta}\sum_{j\in I^{1}}\sigma_{j}\leq\lambda\,\frac{\eta_{i}\left(1+r\left(m-1\right)\right)-r\,\sum_{j\in I^{*}}\eta_{j}}{\Delta}$$

i.e.

$$r(1-r)\sum_{j\in I^1}\sigma_j \le \lambda\left(\eta_i\left(1-r\right)+r\left(\eta_i m-\sum_{j\in I^*}\eta_j\right)\right) \qquad i\in I^*$$

and the second one becomes

$$(1-r)\,\sigma_i + \frac{(1-r)}{r\,(m-1)+1}\,r\,\sum_{j\in I^1}\sigma_j \le \lambda\,(\eta_i - \frac{r}{r\,(m-1)+1}\sum_{j\in I^*}\eta_i)$$

i.e.

$$\Delta \sigma_i + (1-r) r \sum_{j \in I^1} \sigma_j \le \lambda \left(\eta_i \left(1 - r \right) + r \left(\eta_i m - \sum_{j \in I^*} \eta_j \right) \right) \qquad i \in I^1$$

Denote $\gamma_i := (\eta_i (1-r) + r (\eta_i m - \sum_{j \in I^*} \eta_j)), \rho := (1-r) r \sum_{j \in I^1} \sigma_j$ so we have

$$\begin{split} \lambda &\geq \frac{\rho}{\gamma_i}, \quad \text{ for } \quad \gamma_i > 0, \quad i \in I^* \\ \lambda &\geq \frac{\Delta \sigma_i + \rho}{\gamma_i}, \quad \text{ for } \quad \gamma_i > 0, \quad i \in I^1 \end{split}$$

For regularity to hold we must have

$$\max\left\{\frac{\Delta\,\sigma_i+\rho}{\gamma_i}:\gamma_i>0,\quad i\in I^1\right\}>\max\left\{\frac{\rho}{\gamma_i}:\gamma_i>0,\quad i\in I^*\right\}$$

or equivalently

$$\exists i \in I^1, \gamma_i > 0: \quad 1 + \frac{\Delta \sigma_i}{\rho} > \frac{\gamma_i}{\gamma_j} \quad \forall j \in I^*, \ \gamma_j > 0$$
⁽²¹⁾

Let us examine first the question about the existence of $\gamma_i > 0, \ i \in I^1$.

$$\eta_i \left(1 + r \left(m - 1 \right) \right) - r \sum_{j \in I^*} \eta_j > 0 \tag{22}$$

We have

$$\begin{split} \eta_i \left(1 + r \left(m - 1 \right) \right) &- r \sum_{j \in I^*} \eta_j = \eta_i \left(1 + r \left(m - 1 \right) - r \sum_{j \in I^*} \frac{\eta_j}{\eta_i} \right) > \\ \eta_i \left(1 + r \left(m - 1 \right) - r \, m \, K \right) = \eta_i \left(1 - r \, - r \, m \left(K - 1 \right) \right) > \\ \eta_i \left(1 - r \, - r \, n \left(K - 1 \right) \right) \end{split}$$

So if (17) holds, then (22) is satisfied. Let us examine (21)

$$\frac{\gamma_i}{\gamma_j} - 1 = \frac{\eta_i (1 - r) + r (\eta_i m - \sum_{k \in I^*} \eta_k)}{\eta_j (1 - r) + r (\eta_j m - \sum_{k \in I^*} \eta_k)} - 1 = \frac{\eta_i}{\eta_j} \frac{(1 - r) + r (m - \sum_{k \in I^*} \frac{\eta_i}{\eta_j})}{(1 - r) + r (m - \sum_{k \in I^*} \frac{\eta_k}{\eta_j})} - 1 \le K \frac{(1 - r) + r m (1 - 1/K)}{(1 - r) + r m (1 - K)} - 1 = \frac{(K - 1) (1 - r + 2r m)}{(1 - r) - r m (K - 1)}$$

Also, since there always exists an index $i \in I^1$ such that $\sigma_i \ge \sum_{j \in I^1} \sigma_j / (n-m)$ (in a set of numbers there is a number not less than the average), we have that there exists $i \in I^1$ such that

$$\frac{\Delta \, \sigma_i}{\rho} = \frac{\left(r \left(m-1\right)+1\right) \sigma_i}{r \sum_{j \in I^1} \sigma_j} \ge \frac{\left(1-r+r \, m\right)}{r \left(n-m\right)}$$

If

 $\frac{(1-r+rm)}{r(n-m)} > \frac{(K-1)(1-r+2rm)}{(1-r)-rm(K-1)}$ (23)

then (21) holds. Inequality (23) is equivalent to

$$(K-1) r < \frac{(1-r+rm-rmK)(1-r+rm)}{(1-r+2rm)(n-m)}$$

Since

$$\frac{1 - r + r\,m}{1 - r + 2\,r\,m} > \frac{1}{2}$$

and $1 - r + rm - rmK \ge 1 - r + rn - rnK$ we have

$$\frac{(1-r+r\,m\,-r\,m\,K)\,(1-r+r\,m)}{(1-r+2\,r\,m)\,(n-m)} > \frac{1}{2}\,\frac{(1-r+r\,n\,-r\,n\,K)}{n}$$

Then if

$$(K-1)r < \frac{1}{2}\frac{(1-r+rn-rnK)}{n}$$
(24)

(23) and in turn (21) hold. A straightforward computation shows that (24) is equivalent to

$$3nr(K-1) < 1-r$$
 (25)

Note that (25) implies (17). So we may state:

Proposition 2: If (25) holds, the optimum path X^* is regular.

12. References

BORCH, K (1974), The mathematical theory of insurance, Lexington Books.

BÜHLMANN, H. (2006), Private communication.

BÜHLMANN, H. AND H. GERBER (1978), "Risk bearing and the reinsurance market", *The ASTIN Bullettin*, **10**, 12-24.

DANTZIG, G.B. (1963), *Linear programming and extensions*, Princeton University Press, Princeton, NJ, USA.

DE FINETTI, B. (1940), "Il problema dei pieni", *Giornale Istituto Italiano Attuari*, **9**, 1-88.; English translation by L. Barone available as "The problem of "Full-risk insurances", Ch. 1 'The problem in a single accounting period', *Journal of Investment Management*, **4**, 19-43, 2006.

DE FINETTI, B. (1969), Un matematico e l'economia, F. Angeli, Milano.

KARUSH, W. (1939), Minima of functions of several variables with inequalities as side constraints, M.Sc. dissertation, Department of Mathematics, University of Chicago, Chicago, IL, USA.

KUHN, H.W. AND A.W. TUCKER (1951), "Nonlinear programming", in *Proceedings of the Second Berke*ley Symposium on Mathematical Statistics and Probability, J. Neyman ed., University of California Press, Berkeley, CA, USA.

LINTNER, J. (1965), "The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets", *Review of Economic and Statistics*, **47**, 13-37.

MARKOWITZ, H. (1952), "Portfolio selection", Journal of Finance, 6, 77-91.

MARKOWITZ, H. (1956), "The optimization of quadratic functions subject to linear constraints", Naval Research Logistics Quarterly, **3**, 111-133.

MARKOWITZ, H.(2006), "de Finetti scoops Markowitz", Journal of Investment Management, 4, 5-18.

MOSSIN, J. (1966), "Equilibrium in a capital asset market", Econometrica, 34, 768-783.

PRESSACCO, F. (1986), "Separation theorems in proportional reinsurance", in *Insurance and Risk Theory*, M. Goovaerts et al. eds., Reidel Publishing, pp. 209-215.

RUBINSTEIN M. (2006 a), "Bruno de Finetti and mean-variance portfolio selection", Journal of Investment Management, 4, 3-4.

RUBINSTEIN M. (2006 b), A history of the theory of investments, Wiley.

SHAPIRO, J.F. (1979), Mathematical programming: structures and algorithms, Wiley, New York, NY, USA.

SHARPE, W. (1964), "Capital asset prices: a theory of capital market equilibrium under conditions of risk", *Journal of Finance*, **19**, 425-442.