Norsk Regnesentral
NORWEGIAN COMPUTING CENTER

Note

Optimal rebalancing of portfolios with transaction costs

Note no ADMIN/01/2007

Authors Helge Holden
Lars Holden

Date 11th September 2007

Norsk Regnesentral
NORWEGIAN COMPUTING CENTER

The authors

Helge Holden is professor in mathematics at NTNU
Lars Holden is managing director at NR and has a dr. philos. in mathematics

Norwegian Computing Center

Norsk Regnesentral (Norwegian Computing Center, NR) is a private, independent, non-profit foundation established in 1952. NR carries out contract research and development projects in the areas of information and communication technology and applied statistical modeling. The clients are a broad range of industrial, commercial and public service organizations in the national as well as the international market. Our scientific and technical capabilities are further developed in co-operation with The Research Council of Norway and key customers. The results of our projects may take the form of reports, software, prototypes, and short courses. A proof of the confidence and appreciation our clients have for us is given by the fact that most of our new contracts are signed with previous customers.

Norsk Regnesentral
Norwegian Computing Center
Postboks 114, Blindern
NO-0314 Oslo, Norway

Besøksadresse
Office address
Gaustadalléen 23
NO-0373 Oslo, Norway

Telefon · telephone
(+47) 22 85 25 00
Telefaks · telefax
(+47) 22 69 76 60

Internett · internet
www.nr.no
E-post · e-mail
nr@nr.no

Title	**Optimal rebalancing of portfolios with transaction costs**

Authors	**Helge Holden** <Helge.Holden@math.ntnu.no> **Lars Holden** <Lars.Holden@nr.no>

Date	11th September 2007
Publication number	ADMIN/01/2007

Abstract

Rebalancing of portfolios with a concave utility function is considered. It is proved that transaction costs imply that there is a no-trade region where it is optimal not to trade. For proportional transaction costs it is optimal to rebalance to the boundary when outside the no-trade region. With flat transaction costs, the rebalance from outside the no-trade region should be to a state at an internal surface in the no-trade region but never a full rebalance. The standard optimal portfolio theory is extended to n-symmetric assets, general utility function, and more general stochastic processes. Examples are discussed. The paper studies rebalancing of portfolios with a general concave utility function. It is proved that transaction costs imply that there is a no-trade region in the state space where it is not optimal to perform trading. If the transaction costs are proportional, then it is optimal to rebalance when outside the no-trade region to a state at the boundary of this region. If the transaction costs have fixed or flat elements, then the rebalance from outside the no-trade region should be to a state at an internal surface in the no-trade region but never a full rebalance. This extends previous results by considering n-symmetric assets, using a general utility function, and a larger class of stochastic processes. We find an analytic approximation to the boundary of the no-trade region for a quadratic utility function. The results are illustrated with an explicit example and a numerical example with simulation. The paper is a natural extension of the optimal portfolio theory (Markowitz, (7)) since it includes transaction costs and handles more general stochastic processes and utility functions.

Keywords	Rebalance portfolio, portfolio optimization, transaction costs, no-trade region, utility function
Target group	
Availability	Open

Project	
Project number	
Research field	
Number of pages	34
Copyright © 2007	Norwegian Computing Center

Contents

1 Introduction **6**

2 The model **9**

3 Utility functions **18**

4 An explicit example **20**

5 A numerical example **24**

5.0.1 Estimate U when $n = 2$ 25

5.0.2 Estimate U when $n > 2$ 28

5.0.3 Simulation of portfolios 29

6 Closing remarks **33**

References **34**

1 Introduction

Financial portfolios usually consist of several different assets and rebalancing the portfolio is necessary in order to control the weights in the different assets. Most papers neglect transactions costs, but it is documented in, e.g., Atkinson *et al.* (1), Donohue and Yip (4) and Leland (5) that this may be a significant problem. In the present paper it is proved, assuming only a concave utility function, that transaction costs imply that there is a no-trade region where it is not optimal to perform trading. If the transaction costs are proportional, then it is optimal to rebalance when outside the no-trade region to a state at the boundary of this region. If the transaction costs have fixed or flat elements, then the rebalance from outside the no-trade region should be to a state at an internal surface in the no-trade region but never a full rebalance. This extends previous results by considering n-symmetric assets, using a general utility function, and a larger class of stochastic processes. Furthermore, we provide an approximate formula and a more precise iterative simulation algorithm for the boundary of the no-trade region and the internal surface it should be rebalanced to in the case of fixed elements in the transaction costs.

The problem of optimally rebalancing a portfolio with transactions costs is studied in several papers, see Chang (3) and Liu (6) and references herein. The case of two assets can be analyzed analytically, see, e.g., Taksar *et al.* (12) and Øksendal and Sulem (10). The multi-asset problem under strong assumptions has been studied by, e.g., Donohue and Yip (4). There are also some papers with dynamic programming algorithms for determining the no-trade region in higher dimensions, see, e.g., Sun *et al.* (11) and Leland (5). Atkinson *et al.* (1) give an approximate formula for the boundary. Most papers conclude that there is a no-trade region where trade should not be performed, and all rebalance from a state outside the no-trade region is to the boundary of the no-trade region. Donohue and Yip (4) formulate this as a general result with general n and only assuming concave utility function, but does not prove the result. Furthermore, they consider different rebalance strategies for portfolios with one risk-free asset and up to seven risky assets. Liu (6) gives a thorough discussion of the problem with constant absolute risk aversion and one risk-free investment and n uncorrelated geometric Brownian motion investments. The paper discusses both proportional and fixed transaction costs and shows the existence of a no-trade region that is a fixed threshold for each investment with risk. All rebalance is between an investment with risk and the risk-free investment.

There are different formulations of the optimal portfolio problem. The classical paper Markowitz (7) employs a utility function defined on a portfolio. Leland (5) and Donohue and Yip (4) focus on portfolios with targets ratios and the utility function is deviance from the target ratios and the total transaction costs. The recent paper Liu (6) builds on the work of Merton (8) that optimizes the consumption that is possible based on the portfolio. All models only seem to imvolve geometric Brownian motion. The properties of the different formulations appear quite similar, but they offer different mathematical challenges. The present paper extends the framework of Markowitz (7) with n symmetric investments by also incorporating transaction costs in addition to a more general utility function and a larger class of stochastic processes. Our formulation includes the formulation of Leland (5) and the formulation of Merton (8) if we assume the consumption and investment are known. Our assumption with n symmetric assets generalize the assumption in Liu (6) where there is one risk-free asset at the cost of more complex analysis and complex boundary to the no-trade region.

The theoretical agreement regarding the existence of a no-trade region is in contrast to the current practice in most portfolios. Donohue and Yip (4) state that the typically reduction of transaction costs by using an optimal rebalance strategy is 50%. It seems to be most common to rebalance to what is considered the optimal balance at fixed time intervals, often monthly or each quarter, see, e.g., Leland (5). Other portfolios define intervals for the weights in each asset and adjust to the boundary of these intervals either at fixed periods or continuously. Frequently, these decision criteria are often combined with a full rebalance at certain situations. In the Norwegian Petroleum Fund (9) the rebalance is mainly performed when deciding which new assets to buy. In addition, the portfolio is rebalanced to the target weights if the weights are outside certain intervals over two consecutive months. The bank states that as a large investor it is an advantage that the time of rebalances is not known in the marked and that the size of each rebalance is not too large. The strategy presented in this paper satisfies these criteria. The document from Norges Bank (9) includes a numerical simulation study of the transaction costs under different rebalancing strategies.

Section 2 describes the model, and the theorem regarding the no-trade region is proved. Different utility functions are discussed in Section 3. Sections 4 and 5 illustrate the theory with examples using different utility functions. The optimal relative weights and the no-trade region are described in both cases. In Section 4 analytic calculations are sufficient. In this section is it also described how to rebalance a portfolio to the boundary of the no-trade region. In Section 5 it is necessary with approximations and simulation. A procedure on how to determine the no-trade region approximately is provided, and it is shown how to improve this approximation by simulation. This example shows that the transaction costs are reduced to 1/4 by the use of an optimal no-trade region. The paper is ended by some closing remarks in Section 6.

2 The model

Consider n assets, and let $V_{i,t}$ denote the stochastic value of asset i, for $i = 1, 2, \ldots, n$ at time t. We assume that the stochastic properties of $V_{i,t}$ are known, and that the processes are Markovian, i.e., if $V_{i,t}$ is known, we do not get more information regarding the value at a later point in time by knowing the value of $V_{i,s}$ for all values $s < t$. We assume these stochastic processes are sufficiently regular in order that a utility function is defined on the stochastic process. For most utility functions this implies that the expectation and a measure for variability is defined over a time period (t, T) with $T > t$. In particular, this will be satisfied if $\log(V_{i,t})$ are correlated Levy processes. Levy processes include Brownian motion.

The portfolio is given by $\mathbf{a}_t = (a_{1,t}, a_{2,t}, \ldots, a_{n,t})$ where $a_{i,t}$ denotes the weight of asset i at time t. The value of the portfolio at any given time t equals

$$W_t = \sum_{i=1}^{n} a_{i,t} V_{i,t}. \tag{2.1}$$

We denote by

$$r_{j,t} = \frac{a_{j,t} V_{j,t}}{W_t} \tag{2.2}$$

the relative weights. In order to simplify the formulation we let $a_{i,t} \in R$, i.e., both positive and negative values are considered, where negative values indicate a short position in the asset. The results will be similar if we consider only non-negative weights. The investor may at any time rebalance the portfolio. Let the time of rebalance be at times $t_1, t_2, \cdots$. We will consider only one rebalance at a time and therefore omit the index in order to simplify the notation. A rebalance at time t implies that the portfolio is changed from

$$W_{t-} = \sum_{i=1}^{n} a_{i,t-} V_{i,t} \tag{2.3}$$

to

$$W_{t+} = \sum_{i=1}^{n} a_{i,t+} V_{i,t} \tag{2.4}$$

where $a_{i,t}$ are functions of time t that are constant between each rebalancing, and $a_{i,t-}$ and $a_{i,t+}$ denote the limits when approaching t from below and above, respectively. In the case there are discontinuities in the value of the assets, $V_{i,t}$, we will always let $V_{i,t}$ denote the limit from the right, i.e., $V_{i,t} = V_{i,t+}$. This also includes (2.3). The reason for this definition is that if a jump in $V_{i,t}$ implies a rebalance, this is performed immediately, and obviously based on the values after the jump, i.e., $V_{i,t} = V_{i,t+}$.

Assume that

$$W_{t+} = W_{t-} - c(D_t) \tag{2.5}$$

where the function $c(D_t) \geq 0$ is the cost of selling or buying assets at time t. The set D_t contains all relevant information or data regarding the assets up to time t, i.e.,

$$D_t = \{(a_{i,s}, V_{i,s}) \mid i = 1, \ldots, n,\ s \leq t+\}.$$

We will assume that the transaction costs have proportional and fixed terms, i.e.,

$$c(D_t) = \sum_{i=1}^{n} \left[c_{i,1}|a_{i,t+} - a_{i,t-}|V_{i,t} + c_{i,2}\chi(a_{i,t+} - a_{i,t-})\right] \tag{2.6}$$

where $c_{i,j} \geq 0$ for $i = 1, 2, \ldots, n$ and $j = 1, 2$. The function $\chi(a_{i,t+} - a_{i,t-}) = 1$ if $a_{i,t+} - a_{i,t-} \neq 0$, i.e., if there is a rebalance in asset i at time t, and it equals 0 otherwise. For each asset i the cost function consists of a fixed fee if $c_{i,2} \neq 0$ and a cost proportional to the weight if $c_{i,1} \neq 0$. This formula covers the properties of interest from a theoretical point of view, and it is trivial to extend it to let the transaction costs, e.g., depend on whether we sell or buy an asset. If we include tax, the cost function is more complicated depending on when the asset was bought, the difference between value when sold and bought and depends on the country. We have $c(D_t) = 0$ if there is no rebalance at time t.

We need to define a utility function $U(W_t)$ as a real valued function of the stochastic process W_t. It is natural that the utility increases in $E\{W_T\}$ for some value of $T > t$ and decreases in a measure for the variability. See Section 3 for a discussion on possible utility functions.

At every time t the investor optimizes $U(W_t)$. Since the investor at each time t focuses on the future, i.e., $s > t$, a different function is optimized for each value of t. When the investor optimizes at time t, we assume the investor applies the same strategy for each later point $t_1 > t$ when optimizing $U(W_{t_1})$. We assume Markov strategies, i.e., the strategy is a function that only depends on the present situation and not the entire previous history; viz.

$$\mathbf{a}_{t+} = S_z(\mathbf{a}_{t-}, \mathbf{V}_t) \tag{2.7}$$

where $\mathbf{a}_{t-}$ and $\mathbf{a}_{t+}$ denotes the weights before and after a possible rebalance at time t, and $\mathbf{V}_t$ denotes the values of the assets at time t. The same strategy is applied for every value of t. The set of admissible strategies are parameterized by $z \in Z$.

The simplest possible strategy is never to rebalance, that is, $a_{i,t}$ is constant in time, and we denote this by $z = 1$. Another strategy is rebalance in order to have the relative weights within a given interval. Denote the interval by $[\hat{\mathbf{r}}^1, \hat{\mathbf{r}}^2]$. Thus

$$\hat{r}_i^1 \leq r_{i,t+} \leq \hat{r}_i^2$$

or

$$\frac{\hat{r}_i^1}{V_{i,t}} W_{t-} \leq a_{i,t+} \leq \frac{\hat{r}_i^2}{V_{i,t}} W_{t-}.$$

We denote this strategy as $z = 2$, or alternatively $z = (2, \hat{\mathbf{r}}^1, \hat{\mathbf{r}}^2)$. For this strategy it is necessary with detailed rules to determine how this rebalance is performed, e.g., only rebalance when one weight reaches the boundary of the admissible interval. Other strategies are of course also possible.

Instead of having discrete rebalancing at times $t_1, t_2, \ldots$, it is possible to consider the case of continuous rebalancing. For instance, one can rebalance continuously in order to have constant relative weights. This may be denoted as $z = 3$ or $z = (3, \hat{\mathbf{r}}_t)$ if we want to include the relative weights in the notation. Thus

$$r_{i,t+} = r_{i,t-}$$

or

$$a_{i,t+} = a_{i,t-} \frac{W_{t+}}{W_{t-}} = \frac{r_{i,t-}}{V_{i,t}} W_{t+}.$$

In the case of continuous rebalancing it is necessary to redefine (2.3)–(2.6) to a continuous setting. In this case the value of the portfolio has a continuous reduction $c_c(D_t)$ due to transaction costs where[1]

$$c_c(D_t) = \sum_{i=1}^{n} c_{i,1} |a'_{i,t}| V_{i,t}. \tag{2.8}$$

The coefficient $a'_{i,t}$ denotes the derivative of $a_{i,t}$. Many stochastic processes, including Brownian motion, will have unbounded transaction costs if there is continuous rebalancing in order to have constant relative weights if not all $c_{i,1} = 0$.

Since a portfolio depends on the present weights $\mathbf{a}_{t+}$ and the strategy z, we will use the notation $W(\mathbf{a}_{t+}, z)$ where $\mathbf{a}_{t+}$ is the weights after a possible rebalance at time t. We will often omit the plus sign if this is clear from the context.

We want as general formulation of the utility as possible and need the following definitions:

Definition 2.0.1 *We collect the basic definitions below:*
We are given n assets with value $V_{i,t}$, $i = 1, \ldots, n$, at time t.
(i) A portfolio is given by weights $\mathbf{a}_t = (a_{1,t}, a_{2,t}, \ldots, a_{n,t})$ of some assets where the value of the assets $V_{i,t}$ are modeled by stochastic processes. The value W_t of the portfolio is given by equation (2.1). *The investor has the opportunity to rebalance the portfolio according to the equations* (2.3)–(2.6) *or* (2.8) *in case of continuous rebalancing.*

(ii) An investment strategy is a function S_z defined by (2.7). *We denote such strategies*

1. The subscript c indicates continuous rebalancing.

admissible. Let the set of admissible investment strategies may be parameterized by $z \in Z$. We write $z = (z_1, z_2)$ where $z_1 \in \{1, 2, \ldots, m\}$, and z_2 denotes (when required) the parameter values needed in the definition of the strategy.

(iii) A utility function is a real-valued function $U(W_t(\mathbf{a}_t, z))$ where W_t defined by the equations (2.1)–(2.6), *or* (2.8) *in case of continuous rebalancing, is assumed to be sufficiently regular such that U is well-defined. The portfolio value W_t depends on $\mathbf{a}_t$, the portfolio weights, the stochastic processes $\mathbf{V}_t$ describing the value of each asset, and the investment strategy S_z parameterized by z.*

(iv) $U(W_t)$ is continuous/continuously differentiable/concave if $U(W_t(\mathbf{a}_t, z))$ is continuous/continuously differentiable/concave in $\mathbf{a}_t$ for every value of z. The utility function is concave if for any $\mathbf{a}_1$ and $\mathbf{a}_2$ satisfying $\sum_i a_{1,i} V_{i,t} = \sum_i a_{2,i} V_{i,t} = A$ where A is a constant, we have

$$U(W_t(\beta \mathbf{a}_1 + (1-\beta)\mathbf{a}_2, z)) \geq \beta U(W_t(\mathbf{a}_1, z)) + (1-\beta) U(W_t(\mathbf{a}_2, z)) \tag{2.9}$$

for any constant $0 < \beta < 1$. We say that $U(W_t)$ is strictly concave if (2.9) *holds with strict inequality.*

(v) We say that $U(W_t(\mathbf{a}_t, z))$ has compact level sets in the weights if for each combination of $A > 0$, $D \in R$, and $z \in Z$, the set

$$\Omega_{A,D,z,t} = \{\mathbf{a}_t \in R^n \mid \sum_i a_{i,t} V_{i,t} = A, \quad U(W_t(\mathbf{a}_t, z)) \geq D\}$$

is compact.

(vi) We say that the transaction costs are super linear in the weights if the transaction costs due to a rebalance D_t from $\mathbf{a}_{t-}$ to $\mathbf{a}_{t+}$ satisfy

$$c(D_t) \geq B||\mathbf{a}_{t-} - \mathbf{a}_{t+}||_{\mathbf{V}_t} \tag{2.10}$$

for a constant $B > 0$. We use the notation $||\mathbf{a}_t||_{\mathbf{V}_t} = \sum_{i=1}^n |a_{i,t}| V_{i,t}$.

(vii) We say that the utility is super linear in the weights if there exists a constant $M > 0$ such that

$$U(W_t(\hat{\mathbf{a}}_t, z)) \geq U(W_t(\mathbf{a}_t, z)) + M||\hat{\mathbf{a}}_t - \mathbf{a}_t||_{\mathbf{V}_t} \tag{2.11}$$

if $\hat{a}_i \geq a_i$ for all i and for all values of z.

(viii) $U(W_t)$ is homogeneous if $U(cW_t) = cU(W_t)$ for all constants c.

(ix) For a strategy z and constant $A > 0$, we say that $\tilde{\mathbf{a}}_{A,z,t}$ is an optimal weight if it maximizes the utility, that is,

$$U(W_t(\tilde{\mathbf{a}}_{A,z,t}, z)) = \sup_{\substack{\mathbf{a}_t \\ \sum_i a_{i,t} V_{i,t} = A}} U(W_t(\mathbf{a}_t, z)). \tag{2.12}$$

Note that when there are transaction costs, equations (2.3)–(2.6) *imply that $\sum_i a_{i,t+} V_{i,t} < \sum_i a_{i,t-} V_{i,t}$.*

(x) An optimal investment strategy $\tilde{z}$ is a strategy with the following property: It requires a rebalance at time t if and only if $U(W_t(\mathbf{a}_t, z))$ may be increased and then it is rebalanced to a value that optimizes $U(W_t(\mathbf{a}_t, z))$. More precisely,

$$U(W_t(\mathbf{a}_{t+}, \tilde{z})) = \sup_{\hat{z}\in Z} U(W(S_{\hat{z}}(\mathbf{a}_{t-}, \mathbf{V}_t), \hat{z})) \tag{2.13}$$

for all values $\mathbf{a}_{t-}$. Note that $\mathbf{a}_{t+} = S_{\tilde{z}}(\mathbf{a}_{t-}, \mathbf{V}_t)$.

(xi) A no-trade region is a region Ω_t where $(\mathbf{a}_{t-}, z) \in \Omega_t$ if and only if it is not possible to increase $U(W(\mathbf{a}_t, z))$ by a rebalance including a change of strategy z and including the transaction costs at time t. More precisely

$$\begin{aligned}\Omega_t = \{(\mathbf{a}_{t-}, z) \mid \mathbf{a}_{t-} = S_z(\mathbf{a}_{t-}, \mathbf{V}_t) \text{ and } \\ U(W_t(\mathbf{a}_{t-}, z)) = \sup_{\hat{z}\in Z} U(W_t(S_{\hat{z}}(\mathbf{a}_{t-}, \mathbf{V}_t), \hat{z}))\}.\end{aligned} \tag{2.14}$$

The no-trade region may be defined both in terms of the weights $\mathbf{a}_t$ and the relative weights $\mathbf{r}_t$.

When defining optimality we use the Bellmann principle of optimality "If a strategy is optimal for each point in time at that point of time, given that an optimal strategy will be used thereafter, then the strategy is optimal", see Bellman (2). We may then formulate the following theorem.

Theorem 2.0.2 *Consider an admissible set of investment strategies parameterized by $z \in Z$. Let $U(W_t)$ be a utility function defined on a portfolio $W_t = \sum_{i=1}^{n} a_{i,t}V_{i,t}$, that is continuously differentiable, concave and has compact level sets in the weights. Let the transactions costs, $c(D_t)$, be of the form* (2.3)–(2.6) *or* (2.8) *in case of continuous rebalancing.*

(A) There is an optimal weight $\tilde{\mathbf{a}}_{A,z,t}$ for each value of $A = \sum_i a_{i,t}V_{i,t} > 0$ and z. The optimal weight is unique when the utility function is strictly concave.

(B) If $c(D_t) = 0$ for all values of t, then the optimal investment strategy is to rebalance continuously, in order to get optimal weights, i.e., one chooses $z = 3$ and thus $(\tilde{\mathbf{a}}_{A,3,t}, 3)$ given by (2.12).

(C) If the transaction costs, (2.10)*, and the utility function,* (2.11)*, are super linear in the weights, then there is a no-trade region, Ω_t, and $(\tilde{\mathbf{a}}_{A,\tilde{z},t}, \tilde{z}) \in \Omega_t$.*

(D) If there is a no-trade region and the transaction costs have $c_{i,2} = 0$ for $i = 1, \ldots, n$, then there exists an optimal strategy where one only rebalances to the boundary of the no-trade region.

(E) If there is a no-trade region and the transactions costs have $c_{i,2} = 0$ for $i = 1, \ldots, n$, and $V_{i,t}$ for $i = 1, 2, \ldots, n$ have stationary relative increments[2] and the utility function

2. The distribution of $V_{i,t+\Delta t}/V_{i,t}$ is independent of t.

is homogeneous, then the no-trade region is time independent in the relative weights $\mathbf{r}_t$.

(F) If $c_{i,j} > 0$ *for* $i = 1, \ldots, n$ *and* $j = 1, 2$ *and the utility function is super linear in the weights,* (2.11)*, there is a no-trade region. From a state outside the no-trade region it is optimal to rebalance to a state at an internal surface in the no-trade region.*

Proof 2.0.1 *(A) Since* U *has compact level sets in the weights, there exists for each strategy* z *some weights* $\tilde{\mathbf{a}}_{A,z,t}$ *where the optimum is obtained. The function* U *is concave. The optimal weight is unique when the utility function is strictly concave.*

(B) Since U *is continuous on the compact set* $\Omega_{A,D,z,t}$*, there exist parameters* $(\tilde{\mathbf{a}}_{A,\tilde{z},t}, \tilde{z})$ *that optimize* $U(W_t)$*. More precisely,*

$$U(W_t(\tilde{\mathbf{a}}_{A,\tilde{z},t}, \tilde{z})) = \sup_{\substack{(\mathbf{a}_t, z) \\ \sum_i a_{i,t} V_{i,t} = A}} U(W_t(\mathbf{a}_t, z)) \tag{2.15}$$

for each value of t*. When there are no transactions costs, it is optimal always to rebalance to have these weights according to property* (2.12).

(C) Consider a rebalance from $\mathbf{a}_{t-}$ *to* $\mathbf{a}_{t+}$*. Define a third state* $\hat{\mathbf{a}}_t$ *such that* $\hat{a}_{i,t} = a_{i,t+}$ *if* $a_{i,t+} \geq a_{i,t-}$ *and* $\hat{a}_{i,t} = a_{i,t+} + \beta(a_{i,t-} - a_{i,t+})$ *if* $a_{i,t+} < a_{i,t-}$ *where* β *is determined such that*

$$\sum_{i=1}^{n} \hat{a}_{i,t} V_{i,t} = \sum_{i=1}^{n} a_{i,t-} V_{i,t} = \sum_{i=1}^{n} a_{i,t+} V_{i,t} + c(D_t).$$

These two equations may be rewritten such that

$$\begin{aligned} \beta \sum_{\substack{i \\ a_{i,t-} > a_{i,t+}}} (a_{i,t-} - a_{i,t+}) V_{i,t} \\ = \sum_{\substack{i \\ a_{i,t-} > a_{i,t+}}} (a_{i,t-} - a_{i,t+}) V_{i,t} + \sum_{\substack{i \\ a_{i,t-} < a_{i,t+}}} (a_{i,t-} - a_{i,t+}) V_{i,t} = c(D_t). \end{aligned} \tag{2.16}$$

Since $c(D_t) > 0$*, the second equality implies that* β *is positive, while the first one gives that* $\beta \in (0, 1)$*. Thus* $\hat{a}_{i,t} \geq a_{i,t+}$ *and* $\hat{a}_{i,t}$ *is between* $a_{i,t+}$ *and* $a_{i,t-}$ *including the end-points for all* i*. Let* D_t *be a rebalance from* $\mathbf{a}_{t-}$ *to* $\mathbf{a}_{t+}$*. Since* $\beta < 1$ *and the transaction costs are super linear we have the following inequalities*

$$||\hat{\mathbf{a}}_t - \mathbf{a}_{t+}||_{\mathbf{V}_t} \geq c(D_t) \geq B||\mathbf{a}_{t-} - \mathbf{a}_{t+}||_{\mathbf{V}_t}$$

for a constant $B > 0$.

According to the item (A) above, there exist optimal weights $\tilde{\mathbf{a}}_{A,z,t}$ *for each value of* t*. However, when there are transactions costs, it is not necessarily optimal to rebalance to these parameter values. At the optimal weights* $\tilde{\mathbf{a}}_{A,z,t}$*, the derivative of* U *with respect to* $\mathbf{a}_t$ *vanishes on the set* $\sum_i a_{i,t} V_{i,t} = A$*, since we assume that* U *is continuously differentiable. We have that* $\sum_i \hat{a}_{i,t} V_{i,t} = \sum_i a_{i,t-} V_{i,t} = A$*. The derivative will also vanish along each* *line in this plane. Then a Taylor expansion in one variable gives that for all $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$U(W(\hat{\mathbf{a}}_t, z)) \leq U(W(\mathbf{a}_{t-}, z)) + \varepsilon ||\mathbf{a}_{t-} - \hat{\mathbf{a}}_t||_{\mathbf{V}_t}$$

for all $||\mathbf{a}_{t-} - \tilde{\mathbf{a}}_{A,z,t}||_{\mathbf{V}_t} < \delta$, $||\hat{\mathbf{a}}_t - \tilde{\mathbf{a}}_{A,z,t}||_{\mathbf{V}_t} < \delta$ and all values of z.

Using that the utility function increases at least linear in the weights and that δ may be arbitrary small, we have

$$\begin{aligned} U(W(\mathbf{a}_{\mathbf{t}+}, z)) &\leq U(W(\hat{\mathbf{a}}_t, z)) - M||\hat{\mathbf{a}}_t - \mathbf{a}_{t+}||_{\mathbf{V}_t} \\ &\leq U(W(\mathbf{a}_{t-}, z)) + \varepsilon||\mathbf{a}_{t-} - \hat{\mathbf{a}}_t||_{\mathbf{V}_t} - MB||\mathbf{a}_{t+} - \mathbf{a}_{t-}||_{\mathbf{V}_t} \\ &\leq U(W(\mathbf{a}_{t-}, z)). \end{aligned}$$

Hence, it is not possible to increase the utility by rebalancing from $\mathbf{a}_{t-}$ to $\mathbf{a}_{t+}$. The above calculation is valid for any value of z and hence also for the optimal strategy $\tilde{z}$. Since it is not possible to increase the utility by a rebalance from $\mathbf{a}_{t-}$, there must be a no-trade region containing $(\tilde{\mathbf{a}}_{A,\tilde{z},t}, \tilde{z})$.

(D) We will prove that if the transaction costs only have proportional elements and not flat elements (i.e., $c_{i,1} > 0$ and $c_{i,2} = 0$), then all rebalance is to the boundary of the no-trade region. In the following argument, we will assume that we apply the optimal strategy $\tilde{z}$. Let $\mathbf{a}$ be outside the no-trade region and assume the optimal rebalance is to another state $\mathbf{d}$. By definition, the state $\mathbf{d}$ cannot also be outside the no-trade region, since then it will be possible to increase the utility by rebalancing from $\mathbf{d}$. If $\mathbf{d}$ is in the interior of the no-trade region, we may define a third point $\mathbf{b}$ which is on the line between $\mathbf{a}$ and $\mathbf{d}$ and is on the boundary of the no-trade region. Let $c_{\mathbf{a},\mathbf{b}}$, be the transaction costs between $\mathbf{a}$ and $\mathbf{b}$. When the transaction costs are proportional, then $c_{\mathbf{a},\mathbf{d}} = c_{\mathbf{a},\mathbf{b}} + c_{\mathbf{b},\mathbf{d}}$. Then we may consider a rebalance from $\mathbf{a}$ to $\mathbf{d}$ to first be a rebalance from $\mathbf{a}$ to $\mathbf{b}$ and then from $\mathbf{b}$ to $\mathbf{d}$. From the assumptions we have $U(W(\mathbf{a}, \tilde{z})) < U(W(\mathbf{d}, \tilde{z}))$. Since $\mathbf{b}$ is on the boundary of the no-trade region and $\mathbf{d}$ is inside the region, we have $U(W(\mathbf{b}, \tilde{z})) \geq U(W(\mathbf{d}, \tilde{z}))$. This implies that we will not be able to improve the rebalance by rebalancing to the interior of the no-trade region compared to a rebalance to the boundary of the no-trade region.

(E) If all the elements that determine the no-trade region are time independent, then also the no-trade region is time independent. The transaction costs (2.6) *are time independent if it is proportional with the transaction and there are no flat elements (i.e, $c_{i,2} = 0$). It is necessary to assume that the utility function is homogeneous in order to change corresponding to proportional transaction costs. Hence, if $V_{i,t}$ has stationary relative increments, then the no-trade region is time independent in the relative weights $\mathbf{r}_t$.*

(F) When all $c_{i,j} > 0$, the transaction costs are super linear and the existence of the no-trade region follows from item (D). When the utility function is continuous and $c_{i,2} > 0$, it cannot be optimal to perform infinitesimal changes in the weights. Hence from a state outside the no-trade region it is optimal to rebalance to a state inside the no-trade region. When the fixed transaction costs are paid, there still are proportional transaction

costs and item (D) above implies that there is another no-trade region. Hence the optimal rebalance is to a point at the boundary of this interior no-trade region. There may be an interior no-trade region for each combination of assets to sell and buy.

In a practical problem the investor will often increase or decrease the investment. This is an opportunity to rebalance to a lower additional cost than when there are no investment or consumption. Hence, these changes are important for the strategy to optimize the utility including reducing the transaction costs. If these changes are known in advance, the size of the no-trade region will increase when approaching the time when there is a change. The result may be that we only rebalance at these time-points.

In general, rebalancing implies that the investor sells assets that have increased in value over the last period. However, if the knowledge about the change in the value of an asset over a period implies a change in the expected further performance of the value of the asset, e.g., due to time dependent variance, it is critical that this is modeled in a satisfactory way and included in the rebalancing strategy. It is well known that the optimal weights $\tilde{\mathbf{a}}_{A,z,t}$ are sensitive to small changes in the parameters in the stochastic process of the assets. If our expectation regarding the stochastic processes may change, e.g., due to new information, then this uncertainty should be included in the model. This will imply that the utility is more stable in $\mathbf{a}_t,$ the no-trade region will be larger and the probability that we rebalance to weights that we soon after find out are far from optimal due to changes in the expected performance of the stochastic processes, is smaller.

Usually, there are some assets that have high expectations and high uncertainty. If these increase as expected, we may expect to sell these assets regularly in order to maintain the relative weights. However, if these assets have decreased in value, it may be optimal to wait and hope that these assets will increase in value such that a rebalancing will not be necessary. This asymmetry is shown in Tables 5.1 and 5.4. See also discussion at the end of Section 5.0.1. The significance of this effect depends on the difference in expectation and the variability. With the utility function discussed in Section 4 however, the no-trade region is symmetric. This utility function only considers properties at the moment it is evaluated and not expected development later. Therefore, the difference in expectation between the different assets does not influence the no-trade region.

The size of the no-trade region depends on the importance of the transaction costs relative to the importance that the weights in the portfolio are close to the optimal weights, $\tilde{\mathbf{a}}_{A,z,t}$. By making the no-trade region larger, transactions will be less frequent and the associated costs go down. However, this also implies that the weights in the portfolio may be further from the optimal weights. The optimal size of the no-trade region may be difficult to assess, but it is not critical that we know the exact position since the derivative of the improvement vanishes at this point as we will see later in Figure 4.1. Assuming there is only proportional transaction costs, most of the reduction in transaction costs may be ensured if one rebalances only to the estimated boundary of the no-trade region, rather than a full rebalance to the optimal weights. In the case with fixed cost, there is obtained a similar improvement by rebalancing to an internal surface in the no-trade region rather than a full rebalance to the optimal weights.

3 Utility functions

The definitions in this section are motivated by Markowitz (7). An investor wants high expectation and low variability in the portfolio. Hence, the utility should increase in $E\{W_T\}$ for some value of $T > t$ and decrease with the variability. Some authors argue that the utility function also should be decreasing in the expected deviance from a reference portfolio. If we know that we will end the portfolio at time T, it may be natural with the utility function

$$U(W_t) = E\{W_T\} - d(\text{var}(W_T))^{\frac{1}{2}} \tag{3.1}$$

or

$$U(W_t) = E\{W_T\} - d\frac{\text{var}(W_T)}{W_t}. \tag{3.2}$$

Then the end date is coming closer each day, implying that the no-trade region is gradually increasing, since the effect of more optimal weights decreases. If we do not know when the portfolio is ended, it is natural to weight the time with $\exp(-\beta t)$. Then the relative weights of the different time-points in the future is always the same and the no-trade region is stationary if the assets have stationary relative increments. This leads to the following utility function

$$U(W_t) = \int_t^\infty (E\{W_s\} - d(\text{var}(W_s))^{\frac{1}{2}}) \exp(-\beta s) ds \tag{3.3}$$

for constants $d > 0$ and $\beta > 0$. Other alternatives include value at risk $\text{VaR}_\alpha W = P(W < \alpha)$, e.g.,

$$U(W_t) = \int_t^\infty (E\{W_s\} - d\text{VaR}_\alpha(W_s)) \exp(-\beta s) ds, \tag{3.4}$$

or expected shortfall $E\{S_\alpha(W)\} = E\{W | W < \alpha\}$

$$U(W_t) = \int_t^\infty (E\{W_s\} - dE\{S_\alpha(W_s)\}) \exp(-\beta s) ds. \tag{3.5}$$

The threshold α may depend on the time s and the value of the portfolio $W_{s'}$ for $s' < s$, e.g.,

$$\alpha_s = 0.8 \max_{t \le s' \le \max\{t, s-1\}} \{W_{s'}\}.$$

It is possible to increase the investment $K(t) > 0$ or consume part of the investment $K(t) < 0$ by the following equation

$$W_{t+} = W_{t-} + K(t) - c(D_t). \tag{3.6}$$

New investments and consumption are opportunities to rebalance with smaller additional transaction costs than to transfer between different assets. The properties of $K(T)$ for $T > t$ may be known at time t, may be stochastic or may be part of the optimization.

Merton (8) formulates the following utility function

$$U(W_t) = E\{\int_t^T U_1(K_c, s)ds + U_2(W_T, T)\} \tag{3.7}$$

where the two utility functions U_1 is strictly concave in $-K_c$ (i.e., consumption) and U_2 is concave in W_T. The problem is both to rebalance the portfolio and to find the consumption $K_c \leq 0$ that optimizes U. It is assumed to be one risk-free asset with no transaction costs. Liu (6) finds the constant thresholds of the no-trade region for each asset in this model, assuming uncorrelated geometric Brownian motion. All rebalance is a transfer between the risk-free asset and one of the other assets.

Leland (5) and Donohue and Yip (4) define ideal weights $\tilde{r}_i$ and use the utility function

$$U(W_t) = \sum_{i=1}^{n} b_i \int_t^T (r_{i,s} - \tilde{r}_i)^2 ds - \sum_j c(D_{t_j}) \tag{3.8}$$

for constants $b_i > 0$ and where the last sum is over all times t_j where there is a rebalance. Instead of (2.5), it is assumed that transaction costs are paid by additional contributions. Leland (5) finds an approximation to the corner points in the no-trade region with this utility function assuming geometric Brownian motion, proportional transaction costs and one risk-free asset.

Note that the utility functions (3.1), (3.2), (3.3), (3.4), and (3.5) satisfy the assumptions in the theorem. The theorem is also valid for (3.7) if we fix the consumption, and for (3.8) if we set $c(D_t) = 0$ and include the transaction costs in the utility function U. Properties (B)–(F) in the theorem depend on the transaction costs, and, in particular, whether the transaction costs (i) are zero, or (ii) is not identically zero; or (iii) does not contain fixed elements. If there are fixed elements in the transactions costs, it is optimal to rebalance to a point inside the no-trade region in order to avoid a new rebalance too soon. But the rebalance should never be to the optimal value $\tilde{\mathbf{a}}_{A,z,t}$ when the transaction costs have proportional elements in addition to the fixed elements, since the marginal improvement vanishes when the weights approach $\tilde{\mathbf{a}}_{A,z,t}$.

4 An explicit example

In this section we illustrate the theorem by an example that is made so simple that it is possible to estimate the no-trade region mostly by analytic formulas. Let the utility function be

$$U(W_t, \mathbf{r}_t) = (1 - \sum_{k=1}^{n} d_k (r_{k,t} - \tilde{r}_k)^2) W_t. \tag{4.1}$$

Note that we here have included the relative weights in the list of arguments of U. The optimal values of the relative weights are obviously $\tilde{\mathbf{r}} = (\tilde{r}_1, \tilde{r}_2, \cdots, \tilde{r}_n)$, but in a rebalance we need to consider how much the value of the portfolio W_t is reduced compared to the improvement due to better relative weights.

We do not know the exact form of the no-trade region. It is natural to rebalance when the relative weights of one asset $r_{i,t}$ is high and the relative weights of another asset $r_{j,t}$ are low. If a rebalance involves more than two assets, it may be considered as several independent rebalances only involving two assets, when we assume there is only proportional transaction costs, i.e., $c_{k,2} = 0$ for all values of k. This makes it natural to assume that we may approximate the no-trade region by a region given by

$$H = \{\mathbf{r} \mid -D_{j,i} < r_i - r_j < D_{i,j} \text{ for all } i, j \text{ where } i \neq j, \text{ and } \sum_{i=1}^{n} r_i = 1\} \tag{4.2}$$

when there are only proportional transaction costs. We want to determine the constants $D_{i,j}$.

Consider a rebalance from asset i to asset j with $i \neq j$. Assume furthermore that before the rebalance we have $r_{i,t-} - r_{j,t-} > D_{i,j}$, and that we want to rebalance to

$$r_{i,t+} - r_{j,t+} = D_{i,j}. \tag{4.3}$$

We will assume the new weights are $\mathbf{a_{t+}} = \mathbf{a_{t-}} + \mathbf{b}$ where $b_i < 0$, $b_j > 0$, and $b_k = 0$, for $k \neq i, j$. With these restrictions on $\mathbf{b}$ equations (2.5) and (2.6) give

$$W_{t+} = W_{t-} - c_{i,1}|b_i|V_{i,t} - c_{j,1}|b_j|V_{j,t} - c_{i,2} - c_{j,2}. \tag{4.4}$$

This equation has the solution, using equations (2.3) and (2.4),

$$b_i = \frac{-\eta - c_{i,2}}{(1 - c_{i,1})V_{i,t}} \tag{4.5}$$

and

$$b_j = \frac{\eta - c_{j,2}}{(1 + c_{j,1})V_{j,t}} \tag{4.6}$$

for a constant $\eta > c_{j,2}$. We find η by combining the above expression with (4.3). This gives the solution

$$\eta = \frac{a_{i,t-}V_{i,t} - a_{j,t-}V_{j,t} + \frac{c_{i,2}}{1-c_{i,1}} - \frac{c_{j,2}}{1+c_{j,1}} - D_{i,j}(W_{t-} + \frac{c_{i,2}}{1-c_{i,1}} + \frac{c_{j,2}}{1+c_{j,1}})}{D_{i,j}(\frac{1}{1-c_{i,1}} + \frac{1}{1+c_{j,1}}) - \frac{1}{1-c_{i,1}} + \frac{1}{1+c_{j,1}}}. \tag{4.7}$$

The above formula is useful when we want to determine the new weights in a rebalance.

Equation (4.4) combined with the expression for the relative weights (2.2) gives the following expression for the utility function after a rebalance

$$U_+ = (1 - \sum_{k=1}^{n} d_k(r_{k,t+} - \tilde{r}_k)^2)\frac{W_{t-}(1 - r_{i,t-}c_{i,1} + r_{j,t-}c_{j,1}) - c_{i,2} - c_{j,2}}{1 - r_{i,t+}c_{i,1} + r_{j,t+}c_{j,1}}. \tag{4.8}$$

According to Theorem 2.0.2, it is optimal to rebalance whenever the relative weights are outside the no-trade region, and then it is rebalanced to the boundary of the no-trade region. We find $D_{i,j}$ for the values of $r_{i,t+}$ and $r_{j,t+}$ where

$$\frac{\partial U}{\partial r_{i,t+}} - \frac{\partial U}{\partial r_{j,t+}} = 0. \tag{4.9}$$

If we neglect higher-order terms in $c_{i,1}$ and $c_{j,1}$, we get, using (4.8), that

$$D_{i,j} = \frac{c_{i,1} + c_{j,1}}{d_i + d_j} + \tilde{r}_i - \tilde{r}_j. \tag{4.10}$$

Since we are able to find an expression for $D_{i,j}$ that is independent of $r_{k,t}$ for $k \neq i, j$, by neglecting higher order terms in $c_{i,j}$, this indicates that (4.2) is a good approximation to the no-trade region.

We will then consider the case with also a fixed transaction fee, i.e., $c_{k,2} > 0$. For simplicity we neglect the case where it is optimal to rebalance more than two assets at the same time. Assume the no-trade region may be approximated by a region on the form

$$G = \{\mathbf{r} \mid -E_{j,i} < r_i - r_j < E_{i,j} \text{ for all } i, j \text{ where } i \neq j \text{ and } \textstyle\sum_{i=1}^{n} r_i = 1\}. \tag{4.11}$$

In this case it is optimal to rebalance whenever the relative weights are outside G, and then it should be rebalanced to the border of H since this is the optimal value when the flat fee is paid. We find $E_{i,j}$ from the values of $r_{i,t-}$ and $r_{j,t-}$ where the equation $U_- = U_+$ is satisfied. Here U_- is defined by (4.1) and U_+ is defined by (4.8) with $r_{i,t+} - r_{j,t+}$ defined by (4.3) and (4.9). This gives

$$E_{i,j} = 2\Big(\frac{c_{i,2} + c_{j,2}}{(d_i + d_j)W_{t-}}B + \frac{(c_{i,1} + c_{j,1})^2}{4(d_i + d_j)^2}(1 - B)^2\Big)^{\frac{1}{2}} + \frac{c_{i,1} + c_{j,1}}{d_i + d_j}B + \tilde{r}_i - \tilde{r}_j \tag{4.12}$$

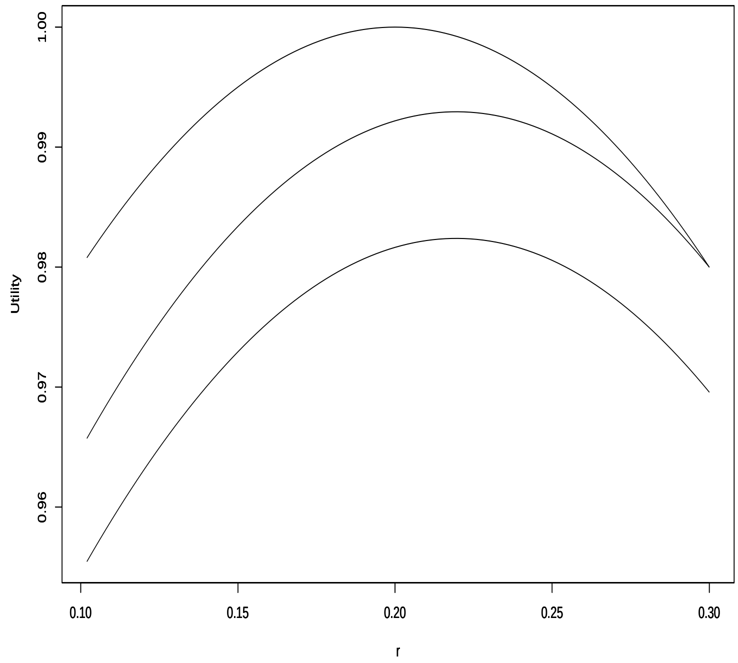

Figure 4.1. The top curve is the utility as a function of r and $W_t = 1$. This has optimum for $\tilde{r} = 0.2$. The two other curves show the utility that may be obtained from a rebalance from $r_{t-} = 0.3$. The top one of these has only proportional transaction fee while the lowest also has flat fee. Notice that both of these curves have optimum for the same value r that is the upper bound of the no-trade region with only proportional costs. The maximum of the lowest curves is equal to $U(0.3)$, illustrating that the boundary of the no-trade region including both proportional and flat transaction costs is 0.3. Since the two curves at the bottom are quite flat close to the optima, we see that the exact position of the no-trade region is not critical.

where

$$B = \frac{4(d_i + d_j) - (c_{i,1} + c_{j,1})^2}{4(d_i + d_j)(1 - \tilde{r}_i c_{i,1} + \tilde{r}_j c_{j,1}) - 2(c_{i,1} + c_{j,1})^2} \approx 1.$$

This expression has the property that $E_{i,j} > D_{i,j}$ and

$$\lim_{c_{i,2}+c_{j,2}\to 0} E_{i,j} = D_{i,j}.$$

Note that $c_{i,2} + c_{j,2}$ scales with W_{t-}. Since the expression for $E_{i,j}$ is independent of r_k for $k \neq i, j$, this indicates that (4.11) is a good approximation for the no-trade region.

If we wanted to include rebalance of three assets at a time, we should define the set

$$G_3 = \{\mathbf{r} \mid -E^3_{i,k,j} < r_i - r_j - r_k < E^3_{i,j,k} \text{ for all different } i \text{ and } j > k \text{ and } \textstyle\sum_{i=1}^n r_i = 1\}. \tag{4.13}$$

We could define similar sets for rebalances involving $4, 5, \cdots$ assets at the same time. The no-trade region would then be defined as the intersection of G and the G_i sets. Then each set G_i contributes to a reduction of the no-trade region if there is a corresponding rebalance contribution to an increase in the utility. However, the computational work increases significantly by including rebalances involving more than two assets.

This utility function is illustrated in Figure 4.1 with $n = 2$. The figure shows three utility curves, the utility as a function of $r_{1,t}$ and two utility curves that may be obtained with a rebalance from a $r_{1,t-} = 0.3$ assuming either only proportional transaction costs or both proportional and flat transaction costs. We have chosen constants $\tilde{r}_1 = 0.2$, $c_{k,1} = 0.04$ and $c_{k,2} = 0.0054$ for $k = 1, 2$. This implies the no-trade regions

$$H = \{r_1 \mid 0.18 < r_1 < 0.22\}$$

and

$$G = \{r_1 \mid 0.1 < r_1 < 0.3\}$$

when only proportional transaction fee and both proportional and flat transaction fee.

5 A numerical example

We will use the utility function

$$U(W_s|W_t) = \int_t^\infty \big(E\{W_s|W_t\} - d\frac{\text{var}(W_s|W_t)}{W_t}\big)\exp(-\beta(s-t))ds. \qquad (5.1)$$

The expression $W_s|W_t$ denotes the conditional distribution for W_s for $s > t$ given W_t. This utility function is a function of W_s for $s > t$. When we evaluate it at time t, then W_t is known and used for scaling the variance. The utility is homogeneous and hence gives a stationary no-trade region Ω_t in the relative weights $\mathbf{r}_t$ if the assets have stationary relative increments. Many other utility functions will have similar properties. In this example, we have chosen to be more precise for this utility function instead of making the text more genera,l but also more technical. We assume the portfolio is evaluated once each day, and only then it is decided if one wants to rebalance.

We will first show how to find the optimal weights for this utility function. Then we will find an approximation for the utility function that may be used in order to find an approximate no-trade region. We will show that this approximation gives good estimates for the no-trade region. However, simulations may give better estimates.

We will assume all assets are lognormal distributed and find optimal weights and an approximation for the no-trade region. Let $X_t = (X_{i,t}, X_{2,t}, \cdots, X_{n,t})$ where

$$X_t \sim N(\mu_X t, \Sigma_X t).$$

The correlation matrix is $\Sigma_X = \{\sigma^2_{X,i,j}\}_{i,j}$. Let

$$V_{i,t} = \exp(X_{i,t}),$$

and

$$\nu_i = \mu_{X,i} + \sigma^2_{X,i,i}/2.$$

Then we have the following

$$E\{V_{i,t}\} = \exp(\nu_i t),$$

and

$$\text{Var}(V_{i,t}) = \exp(2\nu_i t)(\exp(\sigma^2_{i,i}t) - 1).$$

Recall that

$$W_t = \sum_i a_{i,t}V_{i,t}.$$

Assume we continuously rebalance such that

$$r_{i,t} = \frac{a_{i,t} V_{i,t}}{W_t}$$

is constant in time. Define

$$\lambda = \sum_i r_{i,t} \nu_{X,i} \tag{5.2}$$

and

$$\gamma = 2 \sum_i r_{i,t} \nu_{X,i} + \sum_{i,j} r_{i,t} r_{j,t} \sigma^2_{X,i,j}. \tag{5.3}$$

Then we have

$$E\{W_t\} = \exp(\lambda t), \tag{5.4}$$

and

$$E\{W_t^2\} = \exp(\gamma t). \tag{5.5}$$

The utility function (5.1) gives

$$U = \frac{1}{\beta - \lambda} - d\Big(\frac{1}{\beta - \gamma} - \frac{1}{\beta - 2\lambda}\Big). \tag{5.6}$$

The optimal relative weights $\tilde{r}_{i,t}$ are the n variables that optimize (5.6), i.e., where $\partial U / \partial r_{i,t} = 0$ under the constraint that $\sum_i r_{i,t} = 1$. When there are no transaction costs it is optimal to rebalance continously to the optimal relative weights $\tilde{r}_{i,t}$.

When there are transaction costs, there is a no-trade region according to Theorem 2.0.2. We assume there is only proportional transaction costs and that the no-trade region has the form (4.2). The optimal values of $D_{i,j}$ may only be found by CPU intensive simulation. However, we will find an approximation that does not need simulation. This approximation is described for $n = 2$ and later it is shown how to apply it for $n > 2$. In Section 4 it is described how to rebalance to the boundary of the no-trade region.

5.0.1 Estimate U when $n = 2$

For $n = 2$ it is possible to prove that the no-trade region is

$$H_2 = \{r_1 \mid D_- < r_1 < D_+\}, \tag{5.7}$$

only assuming that the utility function is concave and that the transaction costs has the form (2.6). In a simulation the ratio r_1 will vary between the two limits D_- and D_+. If the ratio moves outside the interval, it is rebalanced to the boundary. We will first describe how r varies in this interval. Fix the time t and let $V_{1,\Delta t}$ and $W_{\Delta t}$ represent the change from t to $t + \Delta t$ in V_1 and W, respectively. We have

$$V_{1,t+\Delta t} = V_{1,t} V_{1,\Delta t} = V_{1,t} \exp(X_{1,\Delta t})$$

and

$$W_{t+\Delta t} = W_t W_{\Delta t} = W_t(r_1 \exp(X_{1,\Delta t}) + (1 - r_1)\exp(X_{2,\Delta t}))$$
$$\approx W_t(\exp(r_1 X_{1,\Delta t} + (1 - r_1)X_{2,\Delta t})).$$

This gives

$$r_{1,t+\Delta t} = \frac{V_{1,t+\Delta t}}{W_{t+\Delta t}} = r_{1,t}\frac{V_{1,\Delta t}}{W_{\Delta t}} \approx r_{1,t}(\exp((1 - r_{1,t})(X_{1,\Delta t} - X_{2,\Delta t})).$$

This implies with $\Delta r = r_{1,t} - r_{1,t-\Delta t}$ that

$$E\{\Delta r\} \approx r_{1,t-\Delta t}(E\{\exp((1 - r_{1,t-\Delta t})(X_{1,\Delta t} - X_{2,\Delta t}))\} - 1), \tag{5.8}$$

and

$$E\{(\Delta r)^2\} \approx r^2_{1,t-\Delta t}E\{(\exp((1 - r_{1,t-\Delta t})(X_{1,\Delta t} - X_{2,\Delta t})) - 1)^2\}. \tag{5.9}$$

We will approximate the distribution of Δr with a normal distribution $\phi(\Delta r) \sim N(\mu, \sigma^2)$. Let $p(r_1)$ be an approximation to the distribution of r_1 in the interval $[D_-, D_+]$. We want a $p(r_1)$ that satisfies

$$p(r_1)\phi(\mu) = p(r_1 + \mu)\phi(-\mu), \quad r_1, r_1 + \mu \in (D_-, D_+). \tag{5.10}$$

The motivation for this is that density at r_1 multiplied by the probability to increase r_1 to $r_1 + \mu$ should be equal to the density at $r_1 + \mu$ multiplied by the probability to decrease from $r_1 + \mu$ to r_1 since $p(r_1)$ is independent of time. Equation (5.10) has as solution

$$p_1(r_1) = d_1 \exp(4\mu r_1/\sigma^2)$$

on the open interval (D_-, D_+). In addition, there is a positive probability that r_1 is equal the endpoints D_- and D_+. Therefore, we assume $p(r_1)$ has the form

$$p(r_1) = d_1 \exp(4\mu r_1/\sigma^2) + d_2\delta(D_-) + d_3\delta(D_+), \tag{5.11}$$

where

$$d_3 = d_3 P(\Delta r > 0) + P(r_{1,t} > D_+ \mid r_{1,t-\Delta t} < D_+)$$
$$= d_3 P(\Delta r > 0) + \int_{D_-}^{D_+}\int_{D_+ - r_1}^{\infty} p_1(r_1)\phi(\Delta r)d(\Delta r)dr_1.$$

We approximate $p_1(r_1)$ in the double integral with $p_1(D_+)$ and get

$$d_3 = \frac{p_1(D_+)\int_0^\infty \phi(\Delta r)\Delta r d(\Delta r)}{1 - P(\Delta r > 0)}.$$

Similarly, we find

$$d_2 = \frac{-p_1(D_-)\int_\infty^0 \phi(\Delta r)\Delta r d(\Delta r)}{1 - P(\Delta r < 0)}.$$

The three constants d_1, d_2 and d_3 are scaled such that $\int p(r_1)dr_1 = 1$.

Then the transaction costs due to the upper limit are

$$\begin{aligned}\int_{D_-}^{D_+}\int_{D_+-r_1}^{\infty} p(r_1)\phi(r_{\Delta t})(r_1 + r_{\Delta t} - D_+)(c_{1,1} + c_{2,1})d(\Delta r)dr_1 \\ = (c_{1,1} + c_{2,1})\int_0^{\infty}\phi(\Delta r)\int_{D_+-\Delta r}^{D_+} p(r_1)(r_1 + \Delta r - D_+)dr_1 d(\Delta r) \\ \approx \frac{1}{2}(c_{1,1} + c_{2,1})p(D_+)\int_0^{\infty}\phi(\Delta r)(\Delta r)^2 d(\Delta r).\end{aligned}$$

Using a similar calculation for the transaction costs due to the lower limit, gives the following expression for the transaction costs in a time step

$$C_{\Delta t} = \frac{1}{2}(c_{1,1} + c_{2,1})(p(D_-) + p(D_+))E\{r_{\Delta t}^2\}.$$

The expectation and variance of W_t are calculated as follows, assuming it follows the properties of lognormal distributions,

$$E\{W_t^q\} = (E\{W_{\Delta t}^q\})^{t/\Delta t} = (E\{W_{C,\Delta t}^q(1 - C_{\Delta t})^q\})^{N_d t} = (1 - C_{\Delta t})^{qN_d t}(E\{W_{C,\Delta t}^q\})^{N_d t} \tag{5.12}$$

where $q = 1, 2$. In this calculation, we have assumed the same relative increase in W_t in each time step. The variable $W_{C,\Delta t}$ is the increase in W in one time step Δt when we neglect the reduction due to transaction costs in this time step. Let $N_d = 1/\Delta t$ be number of time steps in a year. Let furthermore

$$\hat{\lambda}_2 = N_d \log(\sum_i p_i(R_i E\{V_{1,\Delta t}\} + (1 - R_i)E\{V_{2,\Delta t}\})) + N_d \log(1 - C_{\Delta t})$$

and

$$\hat{\gamma}_2 = N_d \log(\sum_i p_i E\{(R_i V_{1,\Delta t} + (1 - R_i)V_{2,\Delta t})^2)\} + 2N_d \log(1 - C_{\Delta t})$$

where R_i and p_i for $i = 1, 2, \cdots, m$ is a discretization of $p(r)$ for $r \in (D_-, D_+)$. R_i denotes values in the interval and p_i the corresponding probability.

The expression $\exp(\hat{\lambda}_2 t)$ may be used as an estimate for $E\{W_t\}$ and the expression $\exp(\hat{\gamma}_2 t)$ as an estimate for $E\{W_t^2\}$. Then we may estimate the utility function U by

$$\hat{U}_2 = \frac{1}{\beta - \hat{\lambda}_2} - d(\frac{1}{\beta - \hat{\gamma}_2} - \frac{1}{\beta - 2\hat{\lambda}_2}). \tag{5.13}$$

Instead of finding the two parameters D_-, D_+, it is more stable to find the parameters for the length of the no-trade region $D_+ - D_-$ and for the position relative to the optimal weight $\tilde{r}$, e.g., $P = \frac{\tilde{r} - D_-}{D_+ - D_-}$.

Assume asset 1 has higher expected increase than asset 2. Then $\mu > 0$ and the function p defined in (5.11) is increasing. Since the object function is quite symmetric around the optimal weights, it is more important that the right end point of the no-trade region is closer to the optimal weights than the left end point. Hence, the relative position will satisfy $P > 0.5$. This implies that when asset 1 has lower relative weight than the optimal, we are more reluctant to impose a rebalance than for asset 2, since asset 1 is more likely to increase without a rebalance.

Optimization of the above approximation does not give the optimal no-trade region for the utility function U. However, experiments show that the above approximation gives almost as large values of the utility function as when applying the optimal no-trade region. The approximation may also be used as a first guess on the no-trade region. Then it is possible to adjust these values based on a simulation if wanted.

5.0.2 Estimate U when $n > 2$

We approximate the no-trade region by a region on the form (4.2). For $n > 2$ we use the same technique for each pair r_i and r_j. The quantity $r_i - r_j$ varies between two boundaries $-D_{j,i} < r_i - r_j < D_{i,j}$. If we find an approximation to the pair of boundaries $-D_{j,i}, D_{i,j}$ separately from the other $D_{k,m}$, we may find $-D_{j,i}, D_{i,j}$ similarly as we found D_-, D_+. The optimal values may only be found by a simulation of all the boundaries simultaneously.

Let us fix i and j and set $\Delta r = r_{i,t} - r_{j,t} - r_{i,t-\Delta t} + r_{j,t-\Delta t}$. Similar to the argument for $n = 2$, we have for $n > 2$

$$\Delta r = r_{i,t-\Delta t}\big(\frac{V_{i,\Delta t}}{W_{\Delta t}} - 1\big) - r_{j,t-\Delta t}\big(\frac{V_{j,\Delta t}}{W_{\Delta t}} - 1\big).$$

This gives

$$E\{(\Delta r)^q\} = E\{(r_{i,t-\Delta t}(\frac{V_{i,\Delta t}}{W_{\Delta t}} - 1) - r_{j,t-\Delta t}(\frac{V_{j,\Delta t}}{W_{\Delta t}} - 1))^q\}. \tag{5.14}$$

We may use the approximation

$$\frac{V_{i,\Delta t}}{W_{\Delta t}} \approx \exp(X_{i,\Delta t} - \sum_{k=1}^{n} r_{k,t} X_{k,\Delta t}) \tag{5.15}$$

in order to find expressions for $E\{(\Delta r)^q\}$.

For $n = 2$, we defined discrete values R_i in the interval (D_-, D_+). Similarly, we may define

$$R_{i,k} = \frac{1}{2}(-D_{j,i} - \tilde{r}_i + \tilde{r}_j) + \frac{1}{2m}(k - \frac{1}{2})(D_{i,j} + D_{j,i})$$

and $R_{j,k} = -R_{i,k} + \tilde{r}_i - \tilde{r}_j$ for $k = 1, 2, \cdots, m$. Then $R_{i,k} - R_{j,k}$ varies in the interval $(-D_{j,i}, D_{i,j})$. In order to simplify some expressions below we define $R_{q,k} = \tilde{r}_q$ for $q \neq i, j$ and for $k = 1, 2, \cdots, m$. This is a first order approximation to the average value of r_q. Let p_k be the probability for the discrete value $R_{i,k}$ defined using (5.14) and an expression similar (5.11) for $n > 2$. Note that we find p_k for each pair i, j. We then have the following approximations to λ and γ using (5.12)

$$\hat{\lambda} = N_d \log(\sum_{k=1}^{m} p_k \sum_{v=1}^{n} R_{v,k} E\{V_{m,\Delta t}\}) + N_d \log(1 - C_{\Delta t})$$

and

$$\hat{\gamma} = N_d \log(\sum_{k=1}^{m} p_k \sum_{v=1}^{n} \sum_{u=1}^{n} R_{v,k} R_{u,k} E\{V_{v,\Delta t} V_{u,\Delta t}\}) + 2N_d \log(1 - C_{\Delta t}).$$

We have

$$E\{V_{v,\Delta t}\} = \exp(\mu_{\Delta X,v} + \sigma^2_{\Delta X,v}/2)$$

and

$$E\{V_{v,\Delta} V_{u,\Delta t}\} = \exp(\mu_{\Delta X,v} + \mu_{\Delta X,u} + (\sigma^2_{\Delta X,v} + \sigma^2_{\Delta X,u})/2 + \sigma^2_{\Delta X,v,u}),$$

where $\mu_{\Delta X,v}$, $\sigma^2_{\Delta X,v}$ and $\sigma^2_{\Delta X,v,u}$ correspond to the expectation and variance for variable v and the correlation between v and u per time step.

The approximation $\hat{U}$ to U is calculated from the formula

$$\hat{U} = \frac{1}{\beta - \hat{\lambda}} - d\Big(\frac{1}{\beta - \hat{\gamma}} - \frac{1}{\beta - 2\hat{\lambda}}\Big). \quad (5.16)$$

This approximation may by used in order to find an approximation to the optimal boundaries $(-D_{j,i}, D_{i,j})$. In this approach we find the boundaries for each pair i, j separately. Also in this case, it is stable to parameterize the length of the interval $D_{j,i} + D_{i,j}$ and the position relative to the difference between the optimal weights $\tilde{r}_i - \tilde{r}_j$, e.g., $\frac{\tilde{r}_i - \tilde{r}_j + D_{j,i}}{D_{i,j} + D_{j,i}}$. Application of the above approximation will not give the optimal no-trade region for the object function U. Experiments show that optimizing the approximation gives almost the optimal value. The approximation may also be used in order to find a first guess on the no-trade region. Then it is possible to adjust these values based on simulation if wanted.

5.0.3 Simulation of portfolios

In this section we find the no-trade region by simulation using the utility function (5.1). The value of the assets $V_{i,t}$ is modeled by logarithmic Brownian motion. In the simulation it is rebalanced to the boundary of the no-trade region when the portfolio is outside the region using equation (4.7).

Table 5.1 shows the no-trade region for $n = 2$ for different values of the proportional transaction costs and with and without correlation between the different assets. Note that when adding correlation, it is necessary to change d in order to have optimum for $\tilde{r} = 0.2$. Table 5.2 compares the utility of four different strategies. It is shown that optimal rebalance gives highest utility.

d	$c_{i,1}$	ρ	$\tilde{r}$	D_-	D_+
2.72	0.01	0	0.2	0.165	0.212
2.72	0.001	0	0.2	0.186	0.203
2.15	0.01	0.3	0.2	0.167	0.212

Table 5.1. The critical values for $n = 2$ with proportional transaction. The two assets are geometric Brownian motion with annual expectation and standard deviation equal to $\mu_V = (1.08, 1.02)$ and $\sigma_V = \text{diag}(0.2, 0.04)$, and discounted by $\beta = -\log(0.8)$.

strategy	U	EW_1	U_1	C_n	C_c
no rebalance	4.8089	1.0325	1.0254	0	0
monthly rebalance	4.9230	1.0307	1.0240	12	0.0019
optimal rebalance (approximate)	4.9457	1.0318	1.0251	12	0.00039
optimal rebalance (simulated)	4.9460	1.0316	1.0254	13	0.00047

Table 5.2. Comparison of four different rebalancing strategies with $n = 2$. The parameters are as in the top row of Table 5.1 including the threshold for the optimal rebalance. The quantity U is an estimate for the utility function, $E\{W_1\}$ is the expected value and $U_1 = E\{W_1\} - d\text{var}(W_1)$ is another evaluation of the portfolio after 1 year, C_n and C_c denote the annual number of transactions and annual transaction costs, respectively.

Referring to Table 5.2, we see that the simulated optimal rebalance has as expected highest utility function, but using the approximation to the no-trade region gives only slightly lower utility. The optimal no-trade region, assuming the form (4.2), is small adjustments of the no-trade region given by approximate formulas and computed by simulation. The case of no rebalance increases the ratio of the high volatile asset giving higher expected value of portfolio $E\{W_1\}$ but at a cost of higher variance. The difference in variance increases faster than the difference in expected value. This is seen by comparing no rebalance and optimal rebalance after 1 year when we use the same d value as in the utility function. Optimal rebalance reduces the transaction costs to $1/4$ compared to monthly rebalance.

The second example with $n = 5$ correlated assets and proportional transaction costs is shown in Tables 5.3–5.5. Table 5.3 shows the parameters for the five assets with values that are assumed realistic for the Norwegian stock marked, international stock marked, Norwegian real estate, Norwegian bonds, and international bonds. The table shows the optimal relative weights for the five assets with the utility function (5.1) and three different values of d. Table 5.4 shows the no-trade region for $d = 2$. Note how the length L varies between the different combinations of assets and the position in some cases is far from the symmetric 0.5 value. Table 5.5 is similar to Table 5.2 but for the example with five assets. The table

i	μ_i	σ_i	$\rho_{i,1}$	$\rho_{i,2}$	$\rho_{i,3}$	$\rho_{i,4}$	$\rho_{i,5}$	$\tilde{r}_{i,d=0.5}$	$\tilde{r}_{i,d=1}$	$\tilde{r}_{i,d=2}$
1	1.1	0.22	1	0.7	0.1	0.3	0.1	0.223	0.152	0.083
2	1.09	0.20	0.7	1	0.05	0.1	0.2	0.175	0.129	0.092
3	1.05	0.12	0.1	0.05	1	0	0	0.255	0.201	0.157
4	1.035	0.04	0.3	0.1	0	1	0.3	0.071	0.193	0.306
5	1.035	0.04	0.1	0.2	0	0.3	1	0.0277	0.325	0.362

Table 5.3. The critical values for $n = 5$ with proportional transaction costs $c_{i,1} = 0.01$ and three different d values. The parameters are expectation, standard deviation, correlation, and the optimal values for each asset for the different values of d.

i/j	1	2	3	4	5
1	—	0.094	0.094	0.044	0.041
2	0.69	—	0.069	0.053	0.058
3	0.52	0.59	—	0.12	0.10
4	0.75	0.73	0.67	—	0.14
5	0.50	0.73	0.51	0.67	—

Table 5.4. The optimal no-trade region for the example shown in Table 5.3 with $d = 2$. Above the diagonal is the length L of the interval for $r_i - r_j$ and below the diagonal is the position P of the interval. The no-trade region is then $\tilde{r}_i - \tilde{r}_j - LP < r_i - r_j < \tilde{r}_i - \tilde{r}_j + L(1 - P)$.

strategy	U	$E\{W_1\}$	U_1	C_n	C_c
no rebalance	5.2745	1.04761	1.04228	0	0
monthly rebalance	5.3692	1.04549	1.04092	12	0.0022
optimal rebalance (approximate)	5.3988	1.04693	1.04239	36	0.00049
optimal rebalance (simulated)	5.4000	1.04696	1.04240	36	0.00050

Table 5.5. The table is exactly as Table 5.2 except that $n = 5$ and we have used data as in Table 5.3. It compares the strategies, no rebalance, monthly rebalance, and optimal rebalance with no-trade region found by approximation and by simulation. The optimal no-trade region is as in Table 5.4.

compares four different strategies and shows that the optimal rebalance strategy gives highest utility. The approximation to the no-trade region gives almost as good results as the method based on simulation. The no-trade region is slightly different in the two cases and the result is not sensitive to the exact position. Note that we also here get a reduction of transaction costs to 1/4 compared to monthly rebalance. But in this case the expected annual number of rebalances is 36 which is much larger than in monthly rebalance.

The calculation in both examples is based on 10000 simulations until 10 years and then estimated tail for $t > 10$. For $n = 5$ one such simulation takes about 3 hours using the statistical package R on a standard desk top computer. Finding the optimal no-trade region by estimating the 20 parameters from a good starting point, requires at least 100 simulations which gives about 2 weeks of simulation time. The approach based on approximation took about 20 seconds which is an improvement compared with simulation of the order 10^5.

6 Closing remarks

This paper discusses optimal rebalance of portfolios with transaction costs. We have shown that for n symmetric assets and a general utility function, there is a no-trade region. If the transaction costs are proportional, with no flat or fixed elements, it is optimal to rebalance to the boundary of the no-trade region whenever the portfolio is outside the no-trade region. If the transactions costs have flat elements, it is optimal to rebalance to an internal surface in the no-trade region whenever the portfolio is outside the no-trade region. It is never optimal to have a full rebalance or a calendar-based rebalance.

The theory is illustrated on two examples; one using analytic calculations and approximations and one using simulations. The last example is simulated for $n = 2$ and $n = 5$. Three different rebalance strategies, namely, no rebalance, monthly rebalance, and optimal rebalance are tested using simulations. Both for $n = 2$ and $n = 5$ the transaction costs are reduced by a factor 4 compared to monthly rebalance. These figures are slightly better than other papers on optimal rebalance for a particular utility function. The reduction in transaction costs is probably mainly due to the fact that we rebalance to the boundary of the no-trade region instead of a full rebalance. The reduction in transaction costs is probably not very critical to the exact position of the boundary of the no-trade region. But in order to optimize the utility, it is critical to have an optimal no-trade region. The example shows that the size of the no-trade region depends heavily on the properties of the stochastic processes, not only the size of the transaction costs.

Acknowledgement. The authors thank Steinar Holden for valuable comments.

References

[1] C. Atkinson, S. Pliska, and P. Wilmott. Portfolio management with transaction costs. *Proc. R. Soc. London* A **453**, 551–563 (1997).

[2] R. Bellman. *Dynamic Programming.* Princeton University Press, 1957.

[3] M.-H. Chang. Hereditary portfolio optimization with taxes and fixed plus proportional transaction costs — Part 1. *J. of Appl. Math. Stoch. Anal.*, Article ID 82753, 33p (2007).

[4] C. Donohue and K. Yip. Optimal portfolio rebalancing with transaction costs: improving on calendar or volatility-based strategies. *J. Portfolio Management* **29**, 49–63, Summer 2003.

[5] H. E. Leland. Optimal portfolio management with transaction costs and capital gains taxes. *Technical Report*, Haas School of Business, UC Berkeley, (1999).

[6] Hong Liu. Optimal consumption and investment with transaction costs and multiple risky assets. *J. Finance* **59**, 289–338 (2004).

[7] H. Markowitz. Portfolio selection, *J. Finance* **7**, 77–91 (1952).

[8] R. Merton. Optimum consumption and portfolio rules in a continuous-time model, *J. Economic Theory* **3**, 373–413 (1971).

[9] Norges Bank. Rebalancing regime for the Government Petroleum Fund. `www.norges-bank.no/english/publications/submissions/submissions-2001-08-22` (2001).

[10] B. Øksendal and A. Sulem. Optimal consumption and portfolio with both fixed and proportional transaction costs. *SIAM J. Control Optim.* **40**, 1765–1790 (2002).

[11] W. Sun, A. Fan, Li-Wei Chen, T. Schouwenaars, and M. Albota. Optimal rebalancing for institutional portfolios, *J. of Portfolio Management*, 33–43, Winter 2006.

[12] M. Taksar, M. Klass, and D. Assaf. A diffusion model for optimal portfolio selection in the presence or brokerage fees. *Math. of Oper. Res.* **13**, 277–294 (1988).