Estimation and Arbitrage Opportunities for Exchange Rate Baskets

Danilo Mercurio[1]* Costanza Torricelli[2]†

[1] Humboldt University Berlin
Faculty of Economics
Institute of Staistics and Econometrics
Spandauer Str. 1
10178 Berlin
Germany
isedoc02@wiwi.hu-berlin.de
Tel: +49 (0)30 24725566

[2] University of Modena and Reggio Emilia
Faculty of Economics
Department of Economics
V.le J. Berengario 51
41100 Modena
Italy
torricelli@unimo.it
Tel:+39 059 2056733

Abstract:

This paper analyzes short term portfolio investment opportunities in a capital market where a currency is defined as a currency basket, i.e. a linear combination of foreign currencies. In line with the mean-variance hedging approach, we determine a self financed optimal investment strategy which minimizes an expected quadratic cost function. In order to implement such a strategy an estimate of the basket weights is required. To this end we suggest an adaptive nonparametric procedure, which, if compared with standard procedures, provides very satisfactory results both on simulated and real data. We apply the optimal investment strategy to the case of the Thai Bath basket. The basket weights are computed with the adaptive estimator. We also implement a recursive estimator, a rolling estimator and the Kalman filter which serve as benchmark models. The different estimators are compared with profit based criteria.

Keywords: exchange rates, mean-variance hedging, adaptive estimation.

JEL Classification Numbers C53, F37.

*Financial support by the Deutsche Forschungsgemeinschaft, Graduiertenkolleg für Angewandte Mikroökonomik at Humboldt University, is gratefully acknowledged.

†Financial support by MURST is gratefully acknowledged.

1 Introduction

An exchange rate basket is a form of pegged exchange rate regime and it takes place whenever the domestic currency can be expressed as a linear combination of foreign currencies.
Currency baskets are essentially adopted by developing and transition countries in order to obtain a nominal anchor for monetary policy and to gain some adaptability with respect to fluctuations among the exchange rates of the major international currencies.
Recent crises involving emerging market economies have led some scholars, e.g. Eichengreen et al. (1999), to conclude that pegged exchange rate regimes are inherently crisis prone. Others, e.g. Mussa et al. (2000), do not agree with this statement because these regimes have been successfully implemented by many countries. Nevertheless they confirm that pegged exchange rates can become a great source of vulnerability.
In particular, modern capital mobility enables the investors to exploit the interest rate differentials which may arise between the domestic and the foreign currencies. Furthermore, if it is known that the monetary authorities are committed to sustaining the exchange rate, such speculations seem virtually riskless and can threaten the stability of the exchange rate regime.
The aim of this paper is to analyze short term portfolio investments in a capital market where an exchange rates is defined as a currency basket. The results obtained are compared with Christoffersen and Giorgianni (2000), to which our work is strictly related.
The present work is organized as follows.
In Section 2 we analyze an investment decision in the context of an exchange rate basket. We consider a self financed investment and we develop a strategy which minimizes the expected quadratic cost function. This approach is inspired by mean variance hedging (Musiela and Rutkowski (1997) and Schäl (1994)) and it provides a simple explicit solution for a problem of imperfect hedging. Along with the optimal hedging strategy we also derive the expression of the expected profit and of the expected variance of the profit.
The estimate of the basket weights is required for the construction of the investment strategy, while the expected variance of the profit is needed in order to make an assessment about the riskiness of the investment. The problems related with the estimation of these quantities are discussed in Section 3 where a brief review of the standard approaches is also provided.
From a statistical point of view, the currency basket regime can be represented by a regression model with stochastics regressors and possibly time varying coefficients.
We propose an adaptive estimation algorithm which can cope with this question. The procedure is based on a result by Lipster and Spokoiny (1999). A similar application to the volatility of financial time series can also be found in Mercurio and Spokoiny (2000) and in Härdle et al. (2000).
The adaptive estimator is based on the assumption that the coefficients can be well approximated by a constant over some interval. We call this feature *local time homogeneity.*
The estimation strategy consists in detecting this *interval of time homogeneity* and then estimating the parameters over this interval with standard techniques, such as ordinary least squares.
The adaptive estimator is a nonparametric technique because it does not require that the underlying process belongs to any specific parametric family, such as autoregressive or moving average processes. The only requirement is the local time homogeneity. This hypothesis is fulfilled particularly well by jump processes, which are constant over a certain

Table 1: Speculative strategy.

Time	**Actions**	
t	**borrow**	**lend**
	$\sum_{j=1}^{K} M_j X_{j,t}$	$M_0 Y_t$
$t+h$	**pay back**	**receive**
	$\sum_{j=1}^{K}(1+r_j) M_j X_{j,t+h}$	$(1+r_0) M_0 Y_{t+h}$

interval and make jumps up or down at random times. A simulation study illustrates the performances of the new methodology for jump processes with jumps of different magnitude.
,Finally, in Section 4 we apply the optimal investment strategy to the case of the Thai Bath basket. The basket weights are computed with the adaptive estimator. Furthermore we also implement a recursive estimator, a rolling estimator and the Kalman filter which serve as benchmark models. We calculate the expected and realized profits and the value at risk and we compare the performances of the different estimators with profit based criteria. The last section concludes.

2 Investing in an exchange rate basket

An exchange rate basket regime takes place whenever a currency Y_t can be written as a linear combination of K other currencies. Taking the currency 1 as numeraire, i.e. $X_{1,t} \equiv 1$, one can express the value of the basket by the following equation:

$$Y_t = \sum_{j=1}^{K} \alpha_j X_{j,t}, \tag{1}$$

where $X_{j,t}$ is the amount of currency 1 per unit of currency j, i.e. the cross currency exchange rate.
We remark that for notational convenience we do not isolate the numeraire out of the sum.
The above relationship usually holds only on the average, because the central bank cannot control the exchange rate exactly. Moreover the weights α_j may also change over time because of the reaction of the monetary authorities to changes in macroeconomic fundamentals, and/or speculative pressures.
The aim of this paper is to analyze the possible strategies of an investor who wants to speculate on interest rate differentials which may arise among the countries whose currencies are in the basket. Suppose for simplicity that the country with currency Y_t has the highest interest rates. Then the main idea consists of going long in currency Y_t and short in a portfolio of the other currencies $X_{j,t}$, for $j = 1, \ldots, K$. This concept is summarized by Table 1.
According to a definition of arbitrage which is standard in mathematical finance (see e.g. Elliot and Kopp (1999) or Musiela and Rutkowski (1997)), we would have an arbitrage

opportunity, i.e. a riskless profit, if we can find a portfolio such that the following relationships hold:

$$
\begin{aligned}
0 &= \sum_{j=1}^{K} M_j X_{j,t} - M_0 Y_t \\
0 &\leq (1+r_0) M_0 Y_{t+h} - \sum_{j=1}^{K} (1+r_j) M_j X_{j,t+h} \\
0 &< \mathsf{E}\left((1+r_0) M_0 Y_{t+h} - \sum_{j=1}^{K} (1+r_j) M_j X_{j,t+h} \right)
\end{aligned}
$$

Such a possibility is very unlikely to happen in practice and will not be investigated here. Instead we want to consider an investment which leads to a positive profit on the average. Therefore we face some intrinsic risk and we have to construct our investment strategy in order to minimize it.

The *mean-variance hedging* approach may provide a solution. The mean-variance hedging has been developed for the hedging of non-attainable contingent claims. Furthermore, it can be used as a cheaper alternative to perfect hedging and super hedging and focuses on the minimization of the tracking error at the terminal date (Musiela and Rutkowski (1997) and Schäl (1994)).

In our context, we construct an optimal portfolio made of short positions in the currencies composing the basket: $\sum \xi_j^* X_{j,t+h}$. This portfolio can be regarded as a derivative which has to be replicated in mean by one unit of currency Y_{t+h}.

We say that the portfolio is optimal because it is constructed in such away that the expected deviations between $\sum \xi_j^* X_{j,t+h}$ and Y_{t+h} are zero and the variance is minimized.

2.1 The time varying currency basket regime

We now consider a generalization of the exchange rate basket described by equation (1). The basket weights are an expression of the monetary policy of the central bank and therefore they are regularly updated in order to follow changes in macroeconomic fundamentals, such as trade patterns, or because of speculative pressures.

In order to model the variability of the coefficients we define now the basket regime in the following way:

$$
Y_t = \sum_{j=1}^{K} \alpha_{j,t} X_{j,t} + \varepsilon_t. \qquad (2)
$$

A stationary error term ε_t with $\mathsf{E}\varepsilon_t = 0$, and $\mathsf{E}\varepsilon_t^2 = \sigma^2$, is also added in order to underline the fact that the monetary authorities cannot control the exchange rate exactly and they let it fluctuate around a central parity.

In the remainder we assume that the random variables: Y_t, $X_{j,t}$, $\alpha_{j,t}$ and ε_t have finite second moments.

The agents only observe the exchange rates Y_t and $X_{j,t}$ and they know that the central bank has committed itself to control the magnitude of the fluctuations of the home currency around a basket of known foreign currencies, whereby the values of the basket weights are not disclosed to the public. Therefore the information set $\mathcal{F}_t$ of the agents

only includes the values of the exchange rates Y_s, $X_{j,s}$ for $s \leq t$ and $j = 1, \dots, K$. Formally $\mathcal{F}_t$ is the σ-field generated by Y_s, $X_{j,s}$ for $s \leq t$ and $j = 1, \dots, K$.
Finally we assume that ε_{t+h}, $\alpha_{j,t+h}$ and $X_{j,t+h}$ for $j = 1, \dots, K$ are conditionally independent given $\mathcal{F}_t$. No distributional assumption is required on the ε_t for the development of the financial strategy, while for the sake of the estimation we assume normality.
We are considering a two stage model, where an investment decision is made at time t and the position is kept until time $t+h$. In the section devoted to the empirical analysis we will consider holding periods 30 and 90 days, respectively. This represents a simplification. Nevertheless, it is required because we estimate our model using inter-bank interest rates. Therefore, our assets consist of bank deposits and cannot be traded until maturity.

2.2 Mean-variance hedging

The mean-variance hedging problem can be formulated as follows: we have to determine at time t a strategy $(\xi_1, \dots, \xi_K)$ such that the expected squared deviation of $\sum \xi_j X_{j,t+h}$ form one unit value of currency Y_{t+h} is minimized. We consider therefore the following quadratic cost function:

$$\mathsf{E}\left(\left(Y_{t+h} - \sum_{j=1}^{K} \xi_j X_{j,t+h}\right)^2 \middle| \mathcal{F}_t\right). \tag{3}$$

and substituting (2) for Y_{t+h} we get:

$$\mathsf{E}\left(\left(\sum_{j=1}^{K} \alpha_{j,t+h} X_{j,t+h} - \sum_{j=1}^{K} \xi_j X_{j,t+h} + \varepsilon_{t+h}\right)^2 \middle| \mathcal{F}_t\right)$$
$$= \mathsf{E}\left(\left(\sum_{j=1}^{K} (\mathsf{E}(\alpha_{j,t+h}|\mathcal{F}_t) - \xi_j) X_{j,t+h} + \zeta_{t+h}\right)^2 \middle| \mathcal{F}_t\right),$$

where:

$$\zeta_{t+h} = \varepsilon_{t+h} + \sum_{j=1}^{K} (\alpha_{j,t+h} - \mathsf{E}(\alpha_{j,t+h}|\mathcal{F}_t)) X_{j,t+h}.$$

The cross term is still zero by the assumption of conditional independence and $\mathsf{E}(\zeta_{t+h}^2|\mathcal{F}_t)$ does not depend on ξ_j for $j = 1, \dots, K$. Differentiating with respect to ξ_j, for $j = 1, \dots, K$, we obtain the following system of K first order conditions:

$$-\mathsf{E}\left(X_{j,t+h}\left(\sum_{j=1}^{K} (\mathsf{E}(\alpha_{j,t+h}|\mathcal{F}_t) - \xi_j^*) X_{j,t+h}\right) |\mathcal{F}_t\right) = 0, \quad \text{for } j = 1, \dots, K;$$

whose solution provides the optimal strategy:

$$\xi_j^* = \mathsf{E}(\alpha_{j,t+h}|\mathcal{F}_t) \quad \text{for } j = 1, \dots, K.$$

This optimal strategy implies that the expected value of revising the hedging portfolio is zero:

$$\mathsf{E}\left(Y_{t+h}-\sum_{j=1}^{K}\xi_j^* X_{j,t+h}\middle|\mathcal{F}_t\right)=0$$

and the expected quadratic costs are:

$$\begin{aligned}&\mathsf{E}\left(\left(Y_{t+h}-\sum_{j=1}^{K}\xi_j^* X_{j,t+h}\right)^2\middle|\mathcal{F}_t\right)\\ &\qquad=\mathsf{E}(\zeta_{t+h}^2|\mathcal{F}_t)=\sigma^2+\mathsf{E}\left(\left(\sum_{j=1}^{K}(\alpha_{j,t+h}-\mathsf{E}(\alpha_{j,t+h}|\mathcal{F}_t))X_{j,t+h}\right)^2\middle|\mathcal{F}_t\right)\end{aligned} \tag{4}$$

We find out that $(1+r_0)^{-1}Y_t$ is the amount of money which is needed at time t in order to hedge the portfolio $\sum\mathsf{E}(\alpha_{j,t+h}|\mathcal{F}_t)X_{j,t+h}$ in the mean-variance sense, where r_0 is the interest rate paid on a h day deposit in currency Y_t.
On the other hand, if we prefer avoiding any risk we can simply buy at time t the discounted value of the portfolio: $\sum(1+r_j)^{-1}\mathsf{E}(\alpha_{j,t+h}|\mathcal{F}_t)X_{j,t}$. Therefore, the implementation of the mean-variance hedging strategy may appear profitable only if the following inequality holds:

$$(1+r_0)^{-1}Y_t<\sum_{j=1}^{K}(1+r_j)^{-1}\mathsf{E}(\alpha_{j,t+h}|\mathcal{F}_t)X_{j,t} \tag{5}$$

which means that the initial cost required by the mean-variance hedging is smaller than the one required by the perfect hedging.

2.3 The speculative strategy

If inequality (5) holds, then the speculative strategy can be implemented in the following way:

- borrow the portfolio $\sum(1+r_j)^{-1}\mathsf{E}(\alpha_{j,t+h}|\mathcal{F}_t)X_{j,t}$,
- lend $(1+r_0)^{-1}Y_t$,
- invest the difference $\sum(1+r_j)^{-1}\mathsf{E}(\alpha_{j,t+h}|\mathcal{F}_t)X_{j,t}-(1+r_0)^{-1}Y_t$ at the risk-free rate r_1.

The profit, its expected value and its variance are respectively the following:

$$\begin{aligned}\mathbf{\Pi}_{t+h}^A &= Y_{t+h}-\sum_{j=1}^{K}\mathsf{E}(\alpha_{j,t+h}|\mathcal{F}_t)X_{j,t+h}\\ &\quad+(1+r_1)\left(\sum_{j=1}^{K}(1+r_j)^{-1}\mathsf{E}(\alpha_{j,t+h}|\mathcal{F}_t)X_{j,t}-(1+r_0)^{-1}Y_t\right)\\ \mathsf{E}(\mathbf{\Pi}_{t+h}^A|\mathcal{F}_t) &= (1+r_1)\left(\sum_{j=1}^{K}(1+r_j)^{-1}\mathsf{E}(\alpha_{j,t+h}|\mathcal{F}_t)X_{j,t}-(1+r_0)^{-1}Y_t\right)>0\\ \mathsf{Var}(\mathbf{\Pi}_{t+h}^A|\mathcal{F}_t) &= \mathsf{E}(\zeta_{t+h}^2|\mathcal{F}_t).\end{aligned}$$

This result is comparable with that obtained by Christoffersen and Giorgianni (2000). In order to make the comparison clear, we recall the derivation of their strategy and we show why we consider their result as an *equivalent result* .
They consider a long position in Y_{t+h} and a short position in a portfolio of the currencies which compose the basket: $\sum m_j X_{j,t+h}$. The expression for the profit is therefore:

$$\mathbf{\Pi}^B_{t+h} = Y_{t+h} - \sum_{j=1}^K m_j X_{j,t+h} = \sum_{j=1}^K \alpha_{j,t+h} X_{j,t+h} + \varepsilon_{t+h} - \sum_{j=1}^K m_j X_{j,t+h}$$

and they impose the condition that the initial cash flow has to be zero:

$$0 = \sum_{j=1}^K (1+r_j)^{-1} m_j X_{j,t} - (1+r_0)^{-1} Y_t.$$

The investment strategy $(m_1, \ldots, m_K)$ is constructed in order to remove the currency $j = 2, \ldots, K$ risk (on the average) and therefore it must satisfy the following linear system of equations:

$$\begin{aligned} 0 &= \mathsf{E}\left(\frac{\partial \mathbf{\Pi}_{t+h}}{\partial X_{j,t+h}} | \mathcal{F}_t \right) = \mathsf{E}(\alpha_{j,t+h} | \mathcal{F}_t) - m_j, \quad j = 2, \ldots, K \\ 0 &= \sum_{j=1}^K (1+r_j)^{-1} m_j X_{j,t} - (1+r_0)^{-1} Y_t. \end{aligned}$$

The solution is:

$$\begin{aligned} m_j &= \mathsf{E}(\alpha_{j,t+h} | \mathcal{F}_t), \qquad j = 2, \ldots, K \\ m_1 &= (1+r_1) \left((1+r_0)^{-1} Y_t - \sum_{j=2}^K (1+r_j)^{-1} \mathsf{E}(\alpha_{j,t+h} | \mathcal{F}_t) X_{j,t} \right). \end{aligned}$$

We say that our investment strategy, based on mean-variance hedging arguments, is equivalent to the one of Christoffersen and Giorgianni (2000) because the profits are equal:

$$\mathbf{\Pi}^A_{t+h} = \mathbf{\Pi}^B_{t+h}.$$

Recall that $X_{1,t} \equiv 1$, than it is easy to see that:

$$\begin{aligned} \mathbf{\Pi}^A_{t+h} &= Y_{t+h} - \sum_{j=1}^K \xi_j^* X_{j,t+h} + (1+r_1) \left(\sum_{j=1}^K \frac{\xi_j^* X_{j,t}}{(1+r_j)} - \frac{Y_t}{(1+r_0)} \right) \\ &= Y_{t+h} - \sum_{j=2}^K \xi_j^* X_{j,t+h} + (1+r_1) \left(\sum_{j=2}^K \frac{\xi_j^* X_{j,t}}{(1+r_j)} - \frac{Y_t}{(1+r_0)} \right) \\ &= Y_{t+h} - \sum_{j=1}^K m_j X_{j,t+h} \\ &= \mathbf{\Pi}^B_{t+h}. \end{aligned}$$

Therefore we conclude that an investor is indifferent between the two strategies.
Nevertheless the derivation of the strategy by means of the mean-variance hedging approach has one main advantage: it highlights under which conditions speculation on the interest rate differentials can be convenient (equation (5)).
Furthermore the mean variance approach also provides a simple mean of insuring a short position in a portfolio of the hard currencies when the initial capital is lower than the one required by perfect hedging.

3 The estimation problem

In this section we discuss the problems connected with the estimation of the basket weights and the conditional variance of the profits. As for the former, we first illustrate the techniques used so far and then, in Section 3.1, the approach proposed in this paper. The issue of estimating the basket weights is central to this paper. To solve this problem one has to take into account the following points:

- The estimates of the basket weights are needed in order to take an investment decision at a certain time t. The outcome of this decision will be only known at a certain future time $t+h$. Therefore, only observations up to time t can be used for the estimation.
- If the investment decision is made independently at every stage: t, $t+1$, ..., then we have to consider an "on-line" or "real-time" estimator which regularly updates the value of the estimate as a new observation becomes available.
- It is important to take into account the possible randomness of the basket weights, because the success of the investment strategy directly depends on the accuracy of the estimation.

The most simple approach to the estimation of the basket weights is probably recursive ordinary least squares (OLS). This algorithm suits the problem very well **if the basket weights are indeed constant** . In this case (2) can be seen as a regression equation, and one has only to perform an OLS estimation at each time t, with all the observations available at that time. The regressors are exchange rates which, since Meese and Rogoff (1983), have been usually modeled as random walks. The OLS estimator is therefore even superconsistent, that is it converges at rate t instead of $\sqrt{t}$, because we are estimating a cointegrated system (Hamilton, 1994).
On the other hand, **if the basket weights are not constant**, then recursive OLS produces in general very poor results. A pragmatic and popular way of taking into account the variability of the coefficients consists in choosing a rolling estimator. A fixed window of the data is defined and the estimation is performed. Once a new observation becomes available, the last observation is dropped from the end window and the new one is added at the beginning. Such an algorithm is very easy to implement. Nevertheless, it has many potential drawbacks. It deletes automatically from the sample many observations which could still fulfill the assumptions of parameter constancy, and it does not even try to prevent the fact that a structural break may be present just in the middle of the window which is currently used for the estimation.
Estimators which can cope with time varying parameters have also been proposed. Cooley and Prescott (1973) have introduced in the econometric literature the Kalman filter

for estimating stochastic regression coefficients. This technique has been suggested by Granger (1986) for the context of time varying cointegration and has been applied by Canarella et al. (1990) to test purchasing power parity and by Christoffersen and Giorgianni (2000) for the present problem. Other approaches for the estimation of time varying parameter models can be found for example in Hamilton (1994) and in Elliot et al. (1995).
We now briefly present the specification of the dynamics of the basket weights proposed by Christoffersen and Giorgianni (2000).
They first define the $K \times 1$ vector $\alpha_t^\top := [\alpha_{1,t} \dots \alpha_{K,t}]$, and then propose the following vector autoregressive process for α_t:

$$\alpha_t = \Gamma \alpha_{t-1} + \nu_t \qquad \nu_t \sim N(0, \Sigma), \tag{6}$$

where Γ and Σ are $K \times K$ matrices. Together with equation (2), the above equation constitutes a *state space model*, where (2) is called the measurement equation and (6) is called the transition equation.
If the model is correctly specified and Γ, Σ and σ^2 are known, then α_t can be estimated recursively with the Kalman filter, which has the property of being the best linear estimator. Furthermore, the Kalman filter recursions provide also an expression for the conditional variance of the profit (4).
But the main drawback is that one neither knows Γ and Σ, nor if the model is correctly specified. For this reason one has to plug into the algorithm some "reasonable" values which are believed to be close to Γ and Σ.
This problem is actually very similar to choosing the two smoothing parameters in the nonparametric procedure, which we propose in Section 3.1.
Christoffersen and Giorgianni (2000) choose Γ equal to the identity matrix, and they estimate Σ and σ^2 on-line: σ^2 is estimated from the residuals of recursive OLS and Σ is estimated by taking the sample covariance matrix of the first differences of the past basket weights estimated with recursive OLS.
In this study we want to consider another approach for the estimation of time varying regression coefficients. The related statistical theory has been developed by Lipster and Spokoiny (1999), and an application to the estimation of the volatility of financial time series can be found in Mercurio and Spokoiny (2000) and in Härdle et al. (2000). The aim of the procedure is to improve the rolling estimator. The basic idea of the rolling estimation is maintained, so that the estimation is performed only on a subset of the complete sample. Nevertheless the length of this subset is not fixed but it is estimated adaptively from the data. We try to keep as many observations as possible, if the coefficient are constant, but at the same time we try to detect changes in the parameters as quick as possible.

Another important issue is the estimation of the conditional variance. In fact, the expected variance of the profit can be used for evaluating the riskiness of the investment. The estimation and in particular the forecast of the variance are quite difficult tasks. One of the reason is that the realizations of the variance are not observed, so that one can hardly measure the goodness of the prediction. Furthermore in this case we have to provide an h-step ahead forecast, where $h = 30$ and 90, which is a very long horizon.
The state space modeling of the basket weights proposed by Christoffersen and Giorgianni (2000) may provide a solution to this problem. Nevertheless the estimate of the variance also depends on unknown quantities which have to be plugged into the estimation algorithm. Furthermore, this approach does not offer a solution if one avoids to

model explicitly the dynamics of the basket weights, for example using a rolling OLS. We recall that the expected variance of the profit, which is also the expected quadratic cost of hedging, is given by:

$$\mathsf{E}\left(\left(Y_{t+h}-\sum_{j=1}^{K}\xi_j^* X_{j,t+h}\right)^2 \middle| \mathcal{F}_t\right) = \mathsf{E}\left(\left(Y_{t+h}-\sum_{j=1}^{K}\mathsf{E}(\alpha_{j,t+h}|\mathcal{F}_t)X_{j,t+h}\right)^2 \middle| \mathcal{F}_t\right).$$

Define now $\hat{\alpha}_{j,t+h|t}$ as an estimator of $\mathsf{E}(\alpha_{j,t+h}|\mathcal{F}_t)$, and assume that the conditional variance of the profit is constant for all t. Then we can estimate it by averaging over the past realizations of the square hedging costs.

$$\hat{\boldsymbol{\sigma}}^2_{t+h|t} = \frac{1}{t-h}\sum_{s=1}^{t-h}\left(Y_{s+h}-\sum_{j=1}^{K}\hat{\alpha}_{j,s+h|s}X_{j,s+h}\right)^2 \tag{7}$$

The assumption that the conditional variance is constant may sound quite restrictive. Nevertheless the above estimator should give good results even if the quadratic cost follow some ergodic process, because in this case any 30 to 90 step ahead forecast is close to the unconditional mean.
The above formulation of the estimator of the conditional variance has the advantage of being model-free and it has the appealing interpretation of directly linking the risk of choosing an estimator to its past performance.

3.1 Adaptive window estimation

In this subsection, we discuss an adaptive estimation procedure which can cope with the problem of estimating on-line the regression coefficients of a system with stochastic regressors. The regression coefficients are not assumed to be constant, but local homogeneous, i.e. there exists some interval where they can be well approximated by a constant. This approach suits very well the problem of estimating time varying exchange rate basket weights.
This algorithm is based on a result of Lipster and Spokoiny (1999). Previous applications to the estimation and forecasting of the volatility of financial time series are due to Mercurio and Spokoiny (2000) and Härdle et al. (2000).
Consider the regression equation:

$$Y_t = \mathbf{X}_t^\top \alpha_t + \varepsilon_t, \text{ with } \varepsilon_t \sim N(0,\sigma^2) \quad \forall \quad t, \tag{8}$$

where $\mathbf{X}_t$, α_t K-dimensional random vectors with finite second moments both independent from each other and from ε_t.
We now describe a statistical approach based on the assumption of *local time homogeneity* of the unknown coefficients α_t. The assumption of local time homogeneity means that α_t is nearly constant within an interval $I = [\tau - m, \tau[$. The main objective of the procedure which we propose is to determine the largest of such intervals in a data driven way. Over this interval I we estimate with OLS:

$$\widehat{\alpha}_I = \left(\sum_{t\in I}\mathbf{X}_t\mathbf{X}_t^\top\right)^{-1}\sum_{t\in I}\mathbf{X}_t Y_t. \tag{9}$$

This estimator has the following useful properties. Due to our assumption of local homogeneity, the value of α_t is close to a constant vector for each $t \in I$. This means that the value

$$\Delta_I = \sup_{t \in I} ||\alpha_t - \alpha_\tau||_2$$

is small. Define also the random matrices V_I and W_I as:

$$V_I = \sigma^{-2} \sum_{t \in I} \mathbf{X}_t \mathbf{X}_t^\top \qquad W_I = V_I^{-1}.$$

Hereafter, the elements of the matrix W are denoted by $w_{ij,I}$, $i, j = 1, \ldots, K$.
In the case of a standard regression model with deterministic design, the estimate $\widehat{\alpha}_I$ is the least square estimate and W_I is its covariance matrix. In particular each diagonal element $w_{ii,I}$ of this matrix is the variance of the estimated $\widehat{\alpha}_{i,I}$, $i = 1, \ldots, K$. In our situation the design points are random. By analogy with the regression case, we call $w_{ii,I}$ the conditional variance of $\widehat{\alpha}_{i,I}$, $i = 1, \ldots, K$.
Since the matrix V_I is random, we introduce a random set where certain regularity conditions are satisfied. In particular we want to ensure that V is not degenerated and in the sequel we restrict our considerations to this set. For some positive constants $b > 0$, $B > 1$, $\rho < 1$, $r \geq 1$, $\lambda > \sqrt{2}$ and for $i = 1 \ldots, K$ define the random set, where the following conditions are fulfilled:

$$A_{i,I} = \left\{ \begin{array}{c} b \leq w_{jj,I}^{-1} \leq bB \\ w_{jj,I} ||V||_\infty \leq r \\ |w_{ji,I}/w_{jj,I}| \leq \rho \quad \forall i = 1, \ldots, K \end{array} \right\}$$

where $||V_I||_\infty$ denotes the sup-norm of the matrix V_I:

$$||V_I||_\infty = \sup_{\{\mu \in \mathbb{R}^K : ||\mu|| = 1\}} ||V_I \mu||_2.$$

Theorem 1 *Let $(Y_1, X_1) \ldots (Y_\tau, \mathbf{X}_\tau)$ obey (8), then it holds for the estimate $\widehat{\alpha}_I$:*

$$\begin{aligned} \mathsf{P}\left(|\widehat{\alpha}_{i,I} - \alpha_{i,\tau}| > \Delta_{i,I} + \lambda\sqrt{w_{ii,I}}\,; A_{i,I}\right) \\ \leq 4e\ln(4B)(1 + 2\rho\sqrt{r(K-1)}\lambda)^{K-1}\lambda\exp(-\lambda^2/2), \qquad i = 1, \ldots, K. \end{aligned}$$

The proof of this statement can be found in Härdle et al. (2000).
Given observations $(Y_1, \mathbf{X}_1), \ldots, (Y_\tau, \mathbf{X}_\tau)$ following the locally time homogeneous model (8), we aim at estimating the current value of the parameter α_τ using the estimate $\widehat{\alpha}_I$ with a properly selected time interval I of the form $[\tau - m, \tau[$ in order to minimize the corresponding estimation error. Below we discuss an approach which goes back to the idea of pointwise adaptive estimation, see Lepski (1990), Lepski and Spokoiny (1997) and Spokoiny (1998).
The idea of the method can be explained as follows. Suppose that I is an interval-candidate, that is, we expect time-homogeneity in I and hence in every subinterval of I. This particularly implies that the value Δ_I is negligible and similarly for all Δ_J, $J \subset I$

and that the mean values of the α_t over I and over J nearly coincide. Our adaptive procedure roughly corresponds to a family of tests to check whether $\widehat{\alpha}_I$ does not differ significantly from $\widehat{\alpha}_J$. The latter is done on the basis of Theorem 1 which allows under the assumption of homogeneity within I to bound each $|\widehat{\alpha}_{i,I} - \widehat{\alpha}_{i,J}|$ by $\mu\sqrt{w_{ii,I}} + \lambda\sqrt{w_{ii,J}}$ provided that μ and λ are sufficiently large. If there exists an interval $J \subset I$ such that the hypothesis $\widehat{\alpha}_{i,I} = \widehat{\alpha}_{i,J}$ cannot be accepted then we reject the hypothesis of homogeneity for the interval I. Finally, our adaptive estimator corresponds to the largest interval I such that the hypothesis of homogeneity is not rejected for I itself and all smaller intervals.

Now we present a formal description. Suppose that a family $\mathcal{I}$ of interval candidates I is fixed. Each of them is of the form $I = [\tau - m, \tau[$, $m \in N$, so that the set $\mathcal{I}$ is ordered due to m. With every such interval we associate an estimate $\widehat{\alpha}_{i,I}$ of the parameter $\alpha_{i,\tau}$ due to (9) and the corresponding conditional standard deviation $\sqrt{w_{ii,I}}$.

Next, for every interval I from $\mathcal{I}$, we suppose to be given a set $\mathcal{J}(I)$ of testing subintervals J. For every $J \in \mathcal{J}(I)$, we construct the corresponding estimate $\widehat{\alpha}_J$ from the observations for $t \in J$ according to (8) and compute $\sqrt{w_{ii,J}}$.

Now, with two constants μ and λ, define the adaptive choice of the interval of homogeneity by the following iterative procedure:

Initialization Select the smallest interval in $\mathcal{I}$

Iteration Select the next interval I in $\mathcal{I}$ and calculate the corresponding estimate $\widehat{\alpha}_{i,I}$ and the conditional standard deviation $\sqrt{w_{ii,I}}$

Testing homogeneity Reject I, if there exists one $J \in \mathcal{J}(I)$, and $i = 1, \ldots, K$ such that

$$|\widehat{\alpha}_{i,I} - \widehat{\alpha}_{i,J}| > \mu\sqrt{w_{ii,I}} + \lambda\sqrt{w_{ii,J}}. \tag{10}$$

Loop If I is not rejected, then continue with the iteration step by choosing a larger interval. Otherwise, set $\hat{I} =$ "the latest non rejected I".

The adaptive estimator $\widehat{\alpha}_\tau$ of α_τ is defined by applying this selected interval $\widehat{I}$:

$$\widehat{\alpha}_{i,\tau} = \widehat{\alpha}_{i,\widehat{I}} \text{ for } i = 1, \ldots, K.$$

As for the variance estimation, note that the previously described procedure requires the knowledge of the variance σ^2 of the errors ε_t. In practical applications, σ^2 is typically unknown and has to be estimated from the data. The regression representation (8) and local time homogeneity suggests to apply a residual-based estimator. Given an interval $I = [\tau - m, \tau[$, with $m > K$ we construct the parameter estimate $\widehat{\alpha}_I$. Next the pseudo-residuals $\widehat{\varepsilon}_t$ are defined as $\widehat{\varepsilon}_t = Y_t - \mathbf{X}_t^\top \widehat{\alpha}_I$. Finally the variance estimator is defined by averaging the pseudo-residuals squared:

$$\widehat{\sigma}^2 = \frac{1}{|I|} \sum_{t \in I} \widehat{\varepsilon}_t^2.$$

3.2 Monte Carlo simulations

In order to illustrate the practical implementation of the adaptive estimation procedure we perform a small simulation study. Specifically, we will evaluate the performance of

the proposed algorithm for the case of a change point model, i.e. a model where the regressor coefficients are constant over some interval and then are subjected to a jump, after which they return to being constant. For such a model the *interval of time homogeneity* coincides with the period where the regressor coefficients are constant and the main issue is to detect the change point as quick as possible.
The estimation algorithm involves the choice of the sets $\mathcal{I}$ and $\mathcal{J}(I)$ of the considered intervals and two numerical parameters λ and μ. We now discuss how these parameters can be selected and how they affect the properties of the adaptive estimation. We start form the set of intervals $\mathcal{I}$ and $\mathcal{J}(I)$.

3.2.1 Choice of the sets $\mathcal{I}$ and $\mathcal{J}(I)$

The simplest proposal is to use a regular grid $G = \{t_k\}$ with $t_k = mt_k$ for some natural number m and with $\tau = t_{k^*}$ belonging to the grid. We next consider the intervals $I_k = [t_k, t_{k^*}[= [t_k, \tau[$ for all $t_k < t_{k^*} = \tau$. Every interval I_k contains exactly $k^* - k$ smaller intervals $J' = [t_{k'}, t_{k^*}[$. So that for every interval $I_k = [t_k, t_{k^*}[$ and $k' : k < k' < k^*$ we define the set $\mathcal{J}(I_k)$ of testing subintervals J' by taking all smaller intervals with right end point t_{k^*}: $J' = [t_{k'}, t_{k^*}[$ and all smaller intervals with left end point t_k:$J' = [t_k, t_{k'}[$:

$$\mathcal{J}(I_k) = \{J = [t_{k'}, t_{k^*}[\text{ or } J = [t_k, t_{k'}[: k < k' < k^*\}.$$

The testing interval sets $\mathcal{I}$ and $\mathcal{J}(I)$ are therefore identified by the parameter m: the grid step. If m is small the grid is dense and the test of homogeneity is performed very often. On one hand, this increases the sensitivity of the procedure to structural changes, but on the other hand, it also increases the possibility of rejecting a large interval of time homogeneity. As far as change point models are concerned, the value of m also implies a minimal delay in the perception of a jump.
The estimation is performed both on simulated and real data with a grid step: $m = 30$.

3.2.2 The choice of λ and μ

The behavior of the procedure critically depends on the parameters λ and μ. The values of λ and μ influence the estimation procedure like the bandwidth in nonparametric regression (Green and Silverman, 1994) or density estimation (Park and Marron, 1990). They determine the smoothness of the estimate and the sensitivity of the adaptive estimation procedure. This can be seen directly from equation (10).
Smaller values of λ and μ are likely to lead to a rejection of large intervals quite often, so that the estimate tends to have a smaller bias, but a larger variance. The sensitivity to structural changes of the coefficients is very high, but this may also lead to erroneous rejections of a true interval of time homogeneity due to "normal" stochastic variation or to the presence of outliers.
Larger values of λ and μ present the opposite problem. They imply a less frequent rejection of large intervals and they reduce the sensitivity of the procedure to change points. The optimal choice of the values of λ and μ remains an open question. The theory only requires λ, $\mu > \sqrt{2}$. For the simulated data, we have chosen $\lambda = 2$ and $\mu = 4$, which yield a very good performance for the change point model both for large and small jumps of the regressor coefficients. On real data we select the parameters λ and μ by minimizing the sum of the forecasting error.

3.3 Simulation results

The performance of the adaptive estimator is evaluated with data from the following process:

$$Y_t = \alpha_{1,t} + \alpha_{2,t}X_{2,t} + \alpha_{3,t}X_{3,t} + \varepsilon_t.$$

The length of the sample is 300. The regressors X_2 and X_3 are two independent random walks. The regressor coefficients are constant in the first half of the sample, then they make a jump after which they continue being constant until the end of the sample. We simulate three models with jumps of different magnitude. The values of the simulated models are presented in Table 2.

The error term ε is a Gaussian white noise, with zero mean and variance $\sigma^2 = 10^{-4}$.

Table 2: Simulated models.

first half of the sample: 1 < t < 150			
	$\alpha_{1,t} = 1$	$\alpha_{2,t} = 0.006$	$\alpha_{3,t} = 0.04$
second half of the sample: 151 < t < 300			
large jump model	$\alpha_{1,t} = 0.85$	$\alpha_{2,t} = 0.001$	$\alpha_{3,t} = 0.04$
medium jump model	$\alpha_{1,t} = 0.99$	$\alpha_{2,t} = 0.004$	$\alpha_{3,t} = 0.028$
small jump model	$\alpha_{1,t} = 0.9995$	$\alpha_{2,t} = 0.0055$	$\alpha_{3,t} = 0.0255$

For each of the three models above 100 realizations of the white noise ε are generated and the adaptive estimation is performed. We recall the values of the grid step and of the smoothing parameters: $m = 30$, $\lambda = 2$ and $\mu = 4$.
Figure 1 shows the true value of the coefficients ($\alpha_{1,t}$: top plots, $\alpha_{2,t}$: medium plots, $\alpha_{3,t}$: bottom plots) along with the median, the maximum and the minimum of the estimates of all realizations for each model at each time point.
The simulation results are very satisfactory. The change point is quickly detected, almost within the minimal delay of 30 periods for all three models, so that the adaptive estimation procedure show a good performance even for the small jump model.

4 An application to the Thai Bath basket

In this section we apply the financial and statistical theory developed so far to the case of the Thai Bath basket.
The purpose of this section is twofold: on one hand, we will investigate whether arbitrage profits were possible among the currencies composing the Thai Bath basket, on the other we will evaluate how the results of the speculation are affected by the choice of the estimator.
A similar analysis on the same data set has been performed by Christoffersen and Giorgianni (2000). As far as the comparison of the estimator is concerned, they argue that the Kalman filter should be preferred to the recursive and rolling estimators. This conclusion however is not supported by our results.

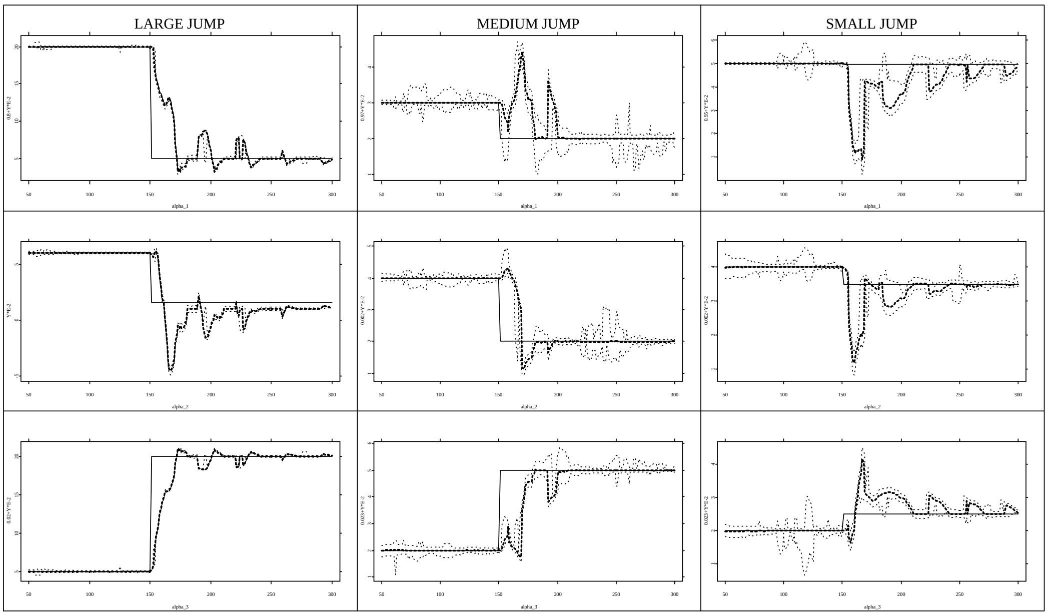

Figure 1: On-line estimates of the regression coefficients with jumps of different magnitude. Median (thick dotted line), maximum and minimum (thin dotted line) among all estimates.

Our set contains the daily exchange rates of the Thai Bath, Japanese Yen and German Mark against the US \$ (Figure 2), together with the nominal interbank 1-and 3-months interest rates on US \$, Mark, Yen and Bath deposits (Figure 3). The period under observation is January 2 1992 to February 12 1997. The source of the data is Bloomberg, L.P. and they were kindly provided by Lorenzo Giorgianni from IMF.
From 1985 until its suspension on July 2, 1997 (following a speculative attack) the Bath was pegged to a basket of currencies consisting of Thailand's main trading partners. In order to gain greater discretion in setting monetary policy, the Bank of Thailand neither disclosed the currencies in the basket nor the weights.
Similarly to Christoffersen and Giorgianni (2000) we assume to know the currencies composing the basket: US \$, Japanese Yen and German Mark. Therefore we can express the US \$/Thai Bath exchange rate in the following way:

$$Y_{\$/th,t} = \alpha_{\$} + \alpha_{gm} X_{\$/gm,t} + \alpha_{jp} X_{\$/jp,t} + \varepsilon_t.$$

The above equation seems to be confirmed by the statistical evidence because the R^2 with respect to the full sample estimation is around 0.8 and the estimated coefficients with fully modified OLS are highly significant.
Unit root tests confirm the hypothesis of nonstationarity for the univariate exchange-rate time series, while Mean-F and Sup-F tests (Hansen, 1992) reject the hypothesis of a stable cointegration relationship among them.

4.1 Practical implementation of the adaptive estimator

The adaptive estimation procedure requires to choice of three parameters: m, λ and μ. The value of m does not influence the results very much and it can be reasonably set to 30. This value represents the minimal amount of data which are used for the estimation, and in the case of a structural break, the minimal delay before having the chance of detecting the change point.
The selection of λ and μ is more critical. These two values determine the sensitivity of the algorithm. Small values would imply a fast reaction to changes in the regressor coefficients, but they would also lead to the selection of intervals of homogeneity which are possibly too small. Large values would imply a slower reaction and consequently the selection of intervals which can be too large.
To overcome this problem we suggest a procedure which is based on heuristic arguments. It has nevertheless some appeal because it focuses on the minimization of the sum of squared forecast errors.
The main idea is that small changes in the values of λ and μ should not affect the estimation results. Therefore we restrict our attention on a set $\mathcal{S}$ of possible pairs (λ, μ). In the present context we chose all the even number between 2 and 8:

$$\mathcal{S} = \{(\lambda, \mu) | \ \lambda, \mu \in \{2, 4, 6, 8\}\}$$

Then we compare the 16 pairs with the following criterion at each time t:

$$(\lambda^*, \mu^*) = \arg\min_{(\lambda,\mu)\in\mathcal{S}} \sum_{s=t-200}^{t-1} \left(Y_s - \sum_{j=1}^{K} \hat{\alpha}_{j,s|s-h} X_{j,s} \right)^2 .$$

Finally, we estimate the value of $\hat{\alpha}_{t+h|t}$ with the selected pair (λ^*, μ^*).
The appeal of the above selection criterion consists in the fact that it leads to the choice of the pair (λ, μ) which has provided the least quadratic hedging costs over the past trading periods.
We remark that the problem of selecting free parameters is not specific to the adaptive estimator. The Kalman filter proposed by Christoffersen and Giorgianni (2000) faces the similar problem of choosing the matrices Γ, Σ and σ^2.

4.2 Three benchmark models

In order to compare the performances of the adaptive estimator with standard approaches we consider three benchmark models: recursive OLS, rolling OLS and the Kalman filter. If the regressors are indeed constant then the recursive OLS estimator:

$$\widehat{\alpha}_t^{rec} = \left(\sum_{s=1}^{t} \mathbf{X}_s \mathbf{X}_s^\top \right)^{-1} \sum_{s=1}^{t} \mathbf{X}_s Y_s$$

equals the Kalman filter of the noisily observed constant process α (Harvey, 1992), and it has the property of being the minimum variance estimator.
The rolling OLS estimator on the other hand represents probably the most simple and popular approach to the estimation of time varying parameters.

$$\widehat{\alpha}_t^{rol} = \left(\sum_{s=t-k}^{t} \mathbf{X}_s \mathbf{X}_s^\top \right)^{-1} \sum_{s=t-k}^{t} \mathbf{X}_s Y_s .$$

It consists in performing the estimation only over a window of the last k observations. In the empirical analysis we choose $k = 250$.
The equations of recursive and rolling OLS are very helpful for understanding the possible advantages which may arise from the implementation of the adaptive estimator. The adaptive estimator is a rolling OLS estimator with a variable window, where the dimension of the window is estimated form the data at each stage:

$$\widehat{\alpha}_t^{ada} = \left(\sum_{s=t-\hat{k}}^{t} \mathbf{X}_s \mathbf{X}_s^\top \right)^{-1} \sum_{s=t-\hat{k}}^{t} \mathbf{X}_s Y_s.$$

Finally the Kalman filter estimate is obtained by the following recursion. Let I be the identity matrix and set the unknown quantities of equation (6) according to the suggestions of Christoffersen and Giorgianni (2000) reported in Section 3. We recall that $\Gamma = I$, Σ and σ^2 are estimated recursively from the data, while the starting value $\widehat{\alpha}_{0|0}$ is estimated by OLS.
The Kalman filter recursions are (Chui and Chen, 1998):

$$\begin{cases} P_{0|0} &= \mathsf{Cov}(\widehat{\alpha}_{0|0}) \\ P_{t|t-1} &= P_{t-1|t-1} + \Sigma\sigma^2 \\ G_t &= P_{t|t-1}\mathbf{X}_t(\mathbf{X}_t^\top P_{t|t-1}\mathbf{X}_t + \sigma^2)^{-1} \\ P_{t|t} &= (I - G_t\mathbf{X}_t^\top)P_{t|t-1} \\ \widehat{\alpha}_{t|t-1} &= \widehat{\alpha}_{t-1|t-1} \\ \widehat{\alpha}_{t|t} &= \widehat{\alpha}_{t|t-1} + G_t(y_t - \mathbf{X}_t^\top \widehat{\alpha}_{t|t-1}), \end{cases}$$

We remark that if the data generating process in equation (6) is well specified and the parameters Γ, Σ and σ^2 are known then the Kalman filter should provide the best performance among all the four estimators. On the other hand, because the data generating process is unknown, estimators which do not require its explicit specification, such as the adaptive estimator or even the rolling estimator can be more robust than the Kalman filter.
Note that for all the above estimators the forecasted value of $\widehat{\alpha}_{t+h|t}$ coincide with the last estimate $\widehat{\alpha}_t$.
We remark furthermore that all the models above require pre-sample values in order to be initialized. For this purpose we use the first 350 observations which are then discarded.

4.3 The results

The estimation results can be seen in Figure 4 and in Figure 5, which show the output of the adaptive procedure calibrated for a 1-month and a 3-month forecast horizon respectively, together with the recursive and rolling OLS estimates.
It is interesting to see that the adaptive estimate tends to coincide with the recursive estimate during the first half of the sample, more or less, while during the second half of the sample it tends to follow the rolling estimate. This may be seen as a hint of a change point, possibly a large realignment of the basket weights. However, these results may be also due to the large outlier which is visible in the middle of the Bath/\$ exchange rate sample (Figure 2).
The main criterion for evaluating the performance of the different estimation procedures

is made by plugging the estimated values of the basket weights into the formula of the optimal quadratic risk minimizing strategy developed in Section 2. There it was also shown that this strategy is equivalent to the one proposed by Christoffersen and Giorgianni (2000). Indeed numerical differences are of the order: 10^{-13}.
The results are displayed in Table 3. We compute the average expected profits, the average expected standard deviations of the profit and the average realized profit.
We also compute the cumulative profits, i.e. the sum over all the realizations of the profits. This quantity is quite interesting since it expresses how much one could have lost or gained by preferring a specific estimator.
In order to evaluate the riskiness of the investment with respect to the different estimators of the basket weights, we also compute the value at risk (VaR) at a level $p = 0.05$. The VaR is a measure of risk which has become popular in recent years and it quantifies the biggest loss that can be incurred with probability p, within a certain horizon h. For example for a standard normal random variable Z it holds that $\text{VaR}(p = 0.05,\ Z) = 1.65$, because $\mathsf{P}(Z < -1.65) = 0.05$, while for a random variable $X \sim N(\mu, \sigma^2)$ it holds that $\text{VaR}(p = 0.05,\ X) = 1.65\sigma - \mu$. As far as our investment is concerned, under the assumption of normality, the VaR of $\mathbf{\Pi}_{t+h}$ is:

$$\text{VaR}(p = 0.05,\ \mathbf{\Pi}_{t+h}) = 1.65\sqrt{\mathsf{E}((\mathbf{\Pi}_{t+h} - \mathsf{E}(\mathbf{\Pi}_{t+h}|\mathcal{F}_t))^2|\mathcal{F}_t)} - \mathsf{E}(\mathbf{\Pi}_{t+h}|\mathcal{F}_t).$$

In Table 3 we report the average values of the VaR and in order to see how correct the measurement of the risk is we compute the relative frequency with which the realizations of the profit overshoot the value at risk:

$$\text{VaR performance} := \frac{\#(\mathbf{\Pi}_{t+h} < -\text{VaR}(p = 0.05,\ \mathbf{\Pi}_{t+h}))}{\#(\mathbf{\Pi}_{t+h})}.$$

If the modeling assumptions and the estimates are "good", then the above quantity should be close to 0.05.
It is interesting to compute the VaR because its value is directly linked to the amount of capital reserves which are required by regulators (see European Commission (1999)) in order to insure a risky portfolio, and therefore it express the cost of detaining a certain position.
If the estimate of the VaR is too conservative, then the investor has to keep more reserves than he would need. On the other hand underestimating the VaR, and consequently the risk, is not only dangerous, but it can also become expensive, because if the performance of the portfolio overshoots the VaR too often, then the regulators will impose higher capital requirements.
We remark that the adaptive estimator sometimes produces negative values of the estimated basket weights. This result is either the sign of a poor estimate and/or a hint for a realignment or instability in the relationship between the exchange rates. Therefore we decide not to invest if the estimated basket weights becomes negative.
The *expected profits* are on the average larger than the realized profits for all the estimators, so that the expectations are in general upward biased. The largest bias is due by far to the recursive estimator, while the Kalman filter shows the smallest one although the adaptive and rolling methods are quite close.
The *expected standard deviation* is similar again across the Kalman filter, rolling and adaptive estimator, while the recursive OLS shows larger results for both investment

horizons. This hints at its poor forecasting performance which is probably due to the randomness of the basket weights.
It must be noted that these last results are quite different from the one obtained by Christoffersen and Giorgianni (2000), where the expected standard deviation estimated under the state space modeling assumptions is much larger than the one for recursive and rolling OLS. This may be due to the fact that their estimate of the expected standard deviation strongly depends on the unknown values of Γ, Σ and σ^2 of equation (6) which have to be plugged into the estimation algorithm. On the other hand the estimator of the expected standard deviation which we propose in Section 4 is model free.
Their conclusion, which cannot be accepted here, was that the rolling OLS causes an underestimation of the standard deviation and therefore reduces the awareness of the risk.
The *Value at Risk* results have to be interpreted with caution, in particular because they rely on the assumption of normality which is hardly fulfilled. Nonetheless one can say that the recursive OLS strongly underestimates the risk (this is a direct consequence of the overestimation of the expected profits). The average values of the VaR are large and negative for both investment horizons, and therefore losses appear to be almost impossible.
The other three estimators perform quite well. The adaptive estimator is the most conservative, while the other two slightly underestimate the risk.
The *average realized profits* are quite similar among all methods. They are positive and, as far as the three month investment horizon is concerned, they are significantly larger than zero.
This provides a clear evidence for the fact that arbitrage profits were possible with in the framework of the Thai Bath basket for the period under study.
The adaptive estimator obtains the largest profits for both investment horizons.
Finally the importance of choosing a good statistical tool is confirmed by the *cumulative profits* which show that even a small improvement in the average profit can be quite rewarding eventually. The best performance is obtained by the adaptive estimator, while the second best performance is obtained by the recursive and rolling OLS together for the one month horizon and by the recursive OLS alone for the three month horizon.

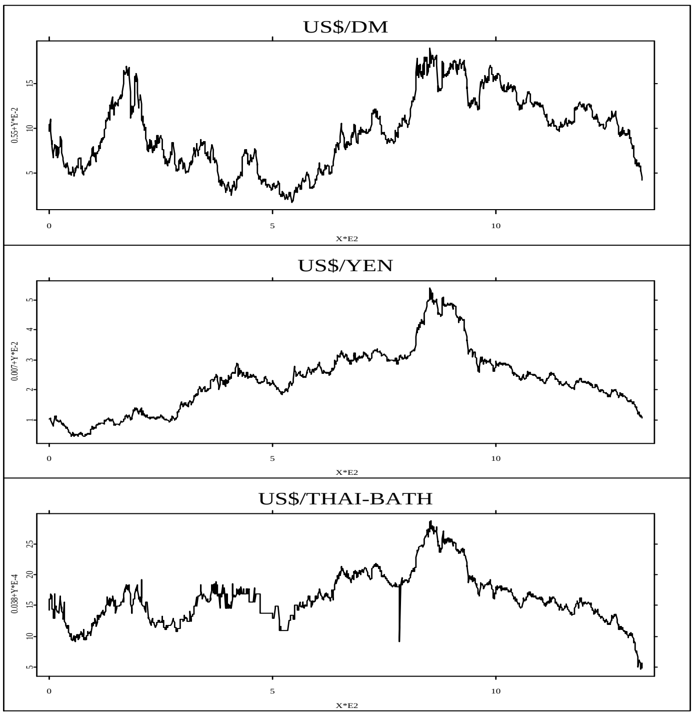

Figure 2: Exchange rate time series.

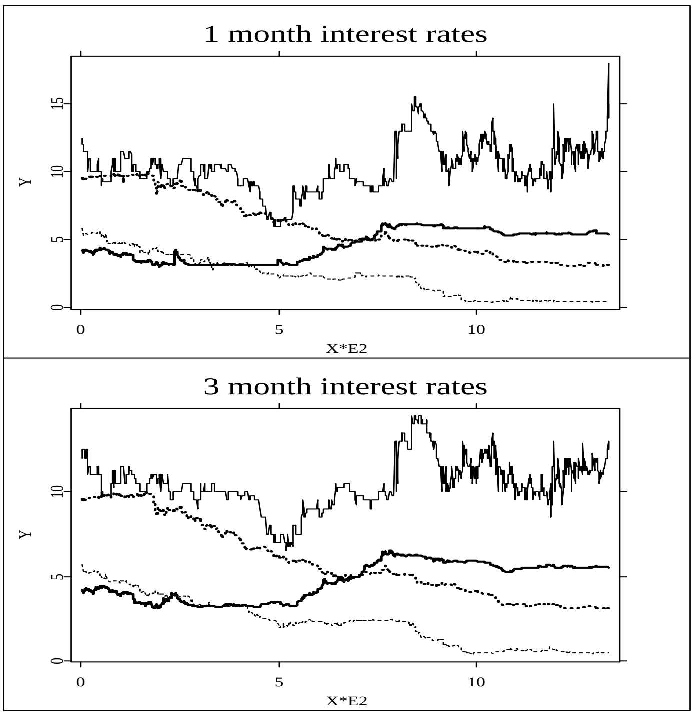

Figure 3: Interest rates time series: German (thick dotted line), Japanese (thin dotted line), American (thick straight line), Thai (thin straight line).

Table 3: Summary statistics of the profits.

one month horizon				
	Recursive	**Rolling**	**KF**	**Adaptive**
Average Expected Profits	0.772	0.565	0.505	0.553
Average Expected Std Deviation	0.542	0.356	0.338	0.390
Average VaR	-0.200	-0.003	-0.023	0.169
VaR Performance (target: 0.05)	0.271	0.059	0.106	0.016
Average Realized Profit	0.403	0.401	0.389	0.420
(Standard errors)	(0.305)	(0.305)	(0.330)	(0.333)
Cumulative Profit	344.1	344.6	331.6	360.8

three month horizon				
	Recursive	**Rolling**	**KF**	**Adaptive**
Average Expected Profits	1.627	1.467	1.375	1.455
Average Expected Std. Deviation	0.633	0.549	0.479	0.455
Average VaR	-0.928	-0.675	-0.663	-0.209
VaR Performance (target: 0.05)	0.271	0.115	0.091	0.004
Average Realized Profit	1.166	1.141	1.147	1.182
(Standard errors)	(0.464)	(0.513)	(0.475)	(0.438)
Cumulative Profit	945.4	925.7	929.1	958.5

The investments are normalized such that at each trading day we take a short position of 100$ in the optimal portfolio of the hard currencies. The result refers to the period April 9 1993 to February 12 1997 for the one month horizon investment and June 7 1993 to February 12 1997 for the three month horizon investment.

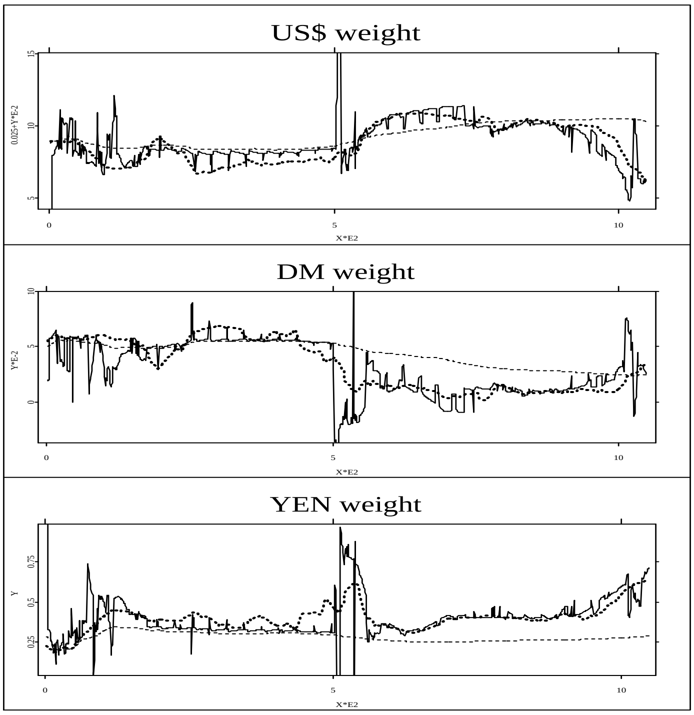

Figure 4: Estimated exchange rate basket weights: 1-month horizon adaptive (straight line), recursive (thine dotted line), rolling (thick dotted line).

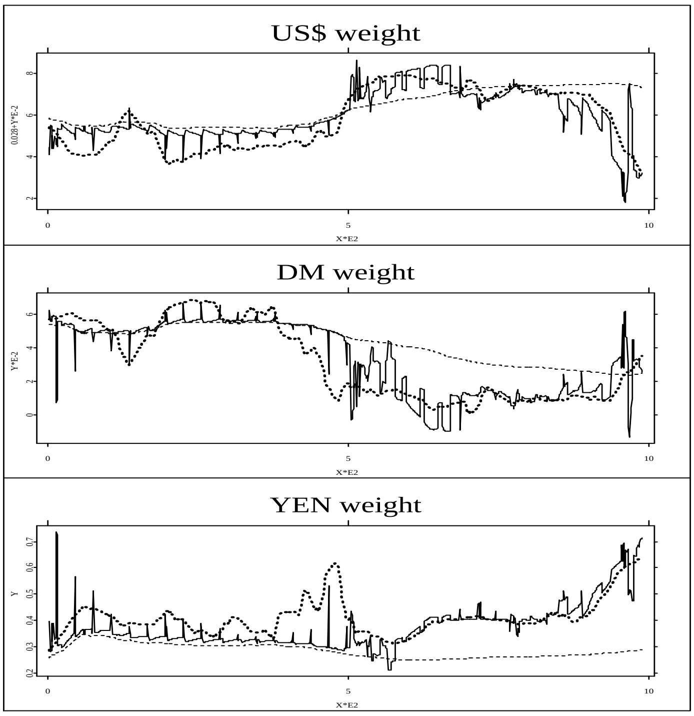

Figure 5: Estimated exchange rate basket weights: 3-month horizon adaptive (straight line), recursive (thine dotted line), rolling (thick dotted line).

5 Conclusions

In this paper we have analyzed the problem of investing and making arbitrage profits in a capital market where the currencies are linked to a basket and it proposes both a theoretical and a practical solution. Furthermore, we have proposed a new adaptive method for the on-line estimation of time varying coefficients for a regression model with random design, which shows a good performance on simulated data and appears to be highly competitive with standard methods on real data.

The main results of the paper are the following.

First, we have analyzed the issue of investing in an exchange rate basket regime and focused on that form of imperfect hedging which is mean-variance hedging. We have derived an expression for the optimal strategy, for the expected profits and for the conditional variance, both for the case of constant and random basket weights.

Secondly, we have tackled the statistical problems connected with the implementation of the arbitrage strategy and the numerical evaluation of the expected profits and conditional variances. In particular, one has to provide an estimate of the basket weights. This problem is statistically very interesting because it consists of a real time estimation of the coefficients of a cointegrated system, which are possibly time varying. We have proposed a new adaptive estimation strategy, which basically assumes that the coefficients are at least almost constant over some unknown intervals. The procedure determines this interval of time homogeneity, where the estimation can be carried out with ordinary least square. A small simulation study shows that the adaptive estimator performs quite well for change point models.

Finally, we have tested the above results in an empirical study of the Thai Bath exchange rate basket. Specifically, we have implemented the optimal investment strategy and evaluated the profits and the VaR as a measure of risk. We have estimated the values of the basket weights by means of the adaptive procedure. However, we also consider three benchmark models: the recursive OLS, the rolling OLS and the Kalman filter. The first one neglects the possible randomness of the parameters, the second one accounts for the time variability of the basket weights because it performs the estimation only on a moving window of the data of fixed length and the third assumes explicitly that the parameters follow a random walk.

All the estimators provide positive arbitrage profits, however the results clearly show the importance of taking care of the possible time variability of the parameters, in particular for a correct evaluation of the risk. The performance of the adaptive method appears to be the best one.

The work performed in this paper still leaves some open questions, which may deserve further research. The choice of an imperfect hedging criterion contains some arbitrariness. We have considered a quadratic cost function, because of its nice analytical properties. The main drawback of this choice lies in the fact that the optimal strategy minimizes the expected deviations from the profits both of positive and negative sign. The minimization of an asymmetric cost function such as $\mathsf{E}(\max(Y_{t+h} - \sum \xi_i X_{i,t+h}, 0)|\mathcal{F}_t)$ may be theoretically more appropriate.

Furthermore, a profitable investment strategy can be carried out only as long as the exchange rate basket regime is sustained by the Central Bank. The break of the Thai Bath basket on July 2, 1997 and the Asian crisis show that this cannot happen indefinitely. It is clear that an approach that is able to infer and predict about the stability of the currency basket would be very useful, but only a larger model which includes also macroeconomic fundamentals or an explicit specification of the data generating process of the basket weights could face these questions. These issues are left for future research.

References

Canarella, G., Pollard, S. K. and Lai, K. S. (1990). Cointegration between exchange rates and relative prices: Another view, *European Economic Review* **34**: 1303–1322.

Christoffersen, P. and Giorgianni, L. (2000). Interest rate arbitrage in currency baskets: Forecasting weights and measuring risk, *Journal of Business and Economic Statistics* **18**: 321–335.

Chui, C. and Chen, G. (1998). *Kalman Filtering*, Information Sciences, third edn, Springer-Verlag, Berlin.

Cooley, T. F. and Prescott, E. C. (1973). An adaptive regression model, *International Economic Review* **14**: 364–371.

Duffie, D. and Pan, J. (1997). An overview of value at risk, *Journal of Derivatives* **4**: 7–49.

Eichengreen, B., Masson, P., Savastano, M. and Sharma, S. (1999). *Transition Strategies and Nominal Anchors on the Road toGreater Exchange Rate Flexibility*, number 213 in *Essays in International Finance*, Princeton University Press.

Elliot, R. J., Aggoun, L. and Moore, J. B. (1995). *Hidden Markov Models*, Springer-Verlag, Berlin.

Elliot, R. J. and Kopp, P. E. (1999). *Mathematics of Financial Markets*, Springer Finance, Springer, New York.

European Commission (1999). A review of regulatory capital requirements for eu credit institutions and investment firms. `http://europa.eu.int/comm/internal_market/en/finances/banks/capaden.pdf%`.

Granger, C. W. J. (1986). Developments in the study of cointegrated economic variables, *Oxford Bulletin of Economics and Statistics* **48**: 213–228.

Green, P. J. and Silverman, B. W. (1994). *Nonparametric Regression and Generalised Linear Models*, Vol. 58 of *Monographs on Statistics and Applied Probability*, Chapman and Hall, London.

Hamilton, J. D. (1994). *Time Series Analysis*, Princeton University Press, Princeton.

Hansen, B. E. (1992). Tests for parameter instability in regressions with i(1) processes, *Journal of Business and Economic Statistics* **10**: 242–254.

Härdle, W., Spokoiny, V. G. and Teyssière, G. (2000). Adaptive estimation for a time inhomogeneouse stochastic volatility model, *Technical Report 6*, Sonderforschungsbereich 373: Quantification and Simulation of Economic Processes. `http://sfb.wiwi.hu-berlin.de`.

Harvey, A. C. (1992). *Time Series Models*, second edn, Harvester Wheatsheaf, New York.

Lepski, O. (1990). One problem of adaptive estimation in gaussian white noise, *Theory Probab. Appl.* **35**: 459–470.

Lepski, O. and Spokoiny, V. (1997). Optimal pointwise adaptive methods in nonparametric estimation, *Annals of Statistics* **25**: 2512–2546.

Lipster, R. and Spokoiny, V. (1999). Deviation probability bound for martingales with applications to statistical estimation, *Stat. & Prob. Letter* **46**: 347–357.

Meese, R. and Rogoff, K. (1983). Empirical exchange rate models of the seventies: Do they fit out of the sample ?, *Journal of International Economics* **14**: 3–24.

Mercurio, D. and Spokoiny, V. (2000). Statistical inference for time-inhomogeneous volatility models, *Technical Report 583*, Weierstrass Institut für Angewandte Analysis und Stochastik, Berlin. http://www.wias-berlin.de/publications/preprints/583.html.

Musiela, M. and Rutkowski, M. (1997). *Martingale Methods in Financial Modelling*, number 36 in *Application of Matemathics. Stochastic Modelling and Applied Probability*, Springer, New York.

Mussa, M., Masson, P., Swoboda, A., Jadresic, E., Mauro, P. and Berg, A. (2000). *Exchange Rate Regimes in an Icreasingly Integrated World Economy*, International Monetary Fund, New York. Occasional Paper 193.

Park, B. U. and Marron, J. S. (1990). Comaprison of data-driven bandwidth selectors, *Journal of the American Statistical Association* **85**: 66–72.

Schäl, M. (1994). On quadratic cost criteria for option hedging, *Mathematics of Operation Research* **19**: 121–133.

Spokoiny, V. (1998). Estimation of a function with discontinuities via local polynomial fit with an adaptive window choice, *Annals of Statistics* **26**: 1356–1378.