Testing for uncovered interest rate parity using distributions implied by FX options.

Martin Cincibuch and David Vávra*

Abstract

Some of the rejections of the uncovered rate parity (UIP) might be due to restrictive distributional assumptions inherent to conventional testing methods. We tested UIP using the information which is carried by currency options and summarized by the risk neutral distributions. Using statistical tests that do not rely on the conventional Fama regression we found empirical support for the hypothesis. Moreover, we designed a new test of the stronger claim that the distributions implied by options well approximate the true distributions of subsequently realized spot rates and found that also this hypothesis was consistent with data.

JEL Classification: F31, G12, G14

Keywords: uncovered interest rate parity – forward unbiasedness – risk neutral distributions – market efficiency – rational expectations

*Monetary Department of the Czech National Bank and CERGE-EI. Correspondening author: Martin Cincibuch, Czech National Bank, Na Prikope 28, 115 03, Prague 1, Czech Republic; e-mail: martin.cincibuch@cnb.cz; phone: +420 22441 2737; fax: +420 22441 2329. The views of the authors are not necessarily those of the institutions. Authors acknowledge support of the Czech Grant Agency under the grant 402/00/0814. We are also grateful to Ron Anderson, Mirek Beneš Chris Gilbert, Jan Hanousek, Peter Hördahl, Michael Koňák for valuable comments and discussions. The usual disclaimer applies.

List of Tables

List of Figures

1 Motivation and summary of results

A popular test of market rationality and risk neutrality has been that of the uncovered interest parity (UIP) hypothesis by regressing of the spot rate changes on the forward premium (interest rate differential). To circumvent the binding restrictions of this estimation strategy, which relies on a wide array of assumptions other than rationality and risk neutrality, we suggest testing the UIP directly by using information revealed by option prices and summarized in implied risk neutral distributions.

The UIP hypothesis relates the current and expected future exchange rates with returns on assets denominated in appropriate currencies by asserting that expected returns from investing in these currencies should be identical with expectations made rationally. The UIP hypothesis (or some of its modifications) has become a cornerstone of most theoretical models of international macroeconomics, regardless of whether they stem from the Keynesian paradigm or are derived from micro-foundations. Hence the large reliance of empirical economic models, which various institutions have used for forecasting and for answering politically relevant questions about economic policy, on this hypothesis (e.g. Bank of England, 2000).

Nevertheless, a large part of the empirical literature rejected the UIP hypothesis and often a significantly negative relationship between interest rate differential and appropriate spot rate changes was reported. Engel (1996) in his survey cited almost 200 articles and concluded that regression based empirical tests almost always rejected the hypothesis. Engel (1996) also cited many potential economic and also econometric explanations for the rejections, but none of them was able to interpret regression results completely.

However, the literature on the peso problem (Krasker, 1980), on the regime changes (Evans and Lewis, 1995) and on the learning hypothesis (Lewis, 1989) suggests that rejections might be at least partly caused by the short length of the available samples. The scope of the small sample bias was shown by Baillie and Bollerslev (2000), who constructed a model with persistent daily volatility and in which they impose the UIP. The model is able to simulate results reported in the literature. Indeed, some of the more recent empirical findings point in this direction. Flood et. al (2001) found that UIP performed much better in the 1990s than in previous decades and that the case for the rejection appears weaker if one allows for longer samples, and Chinn and Meredith (2002), who used post-Bretton Woods time series of interest rates on longer maturity bonds, found coefficients of the correct sign close to unity. Huisman et. al (1998) tested the UIP using a random time effects panel model, which granted more efficiency than the usual bilateral regression and which allowed the control of several potentially biasing factors. In contrast with previous results, their estimates of the slope coefficient are significantly positive and average to 1/2.

A more serious attack on the traditional regression as a testing strategy for the UIP hypothesis was launched by Barnhart, McNown and Wallace (1999), who argued that regression-based tests might be uninformative because of inconsistent coefficient estimates due to simultaneity bias. Although this possibility was understood earlier (Fama, 1984; Liu and Maddala, 1992), Barnhart et. al (1999) argued that this problem is more pervasive than commonly recognized. They show that when the simultaneity is accounted for then the estimates of slope coefficients with the forward premium are virtually one for several major currency pairs. Further, by simulations they show that widely cited tests wrongly reject forward unbiasedness at high rates even if data were generated by models consistent with the UIP. Perhaps, a less severe simultaneity problem might arise when the longer term interest rates are taken into account, and if so then the results of Barnhart et. al (1999) may shed some light on the findings of Chinn and Meredith (2002).

Such results also correspond with the intuitive view of practitioners in central banks and other institutions that the correlation of the exchange and interest rates likely depends on the nature of the underlying shock in a fully specified dynamic model, in which the UIP is one of the structural relationships (e.g. Beneš et. al, 2002b). For example, when interest rates react to a demand inflationary shock, exchange rates move together as a response. On the other hand, when the exchange rate depreciates owing to, say, a rise in the risk premium, the interest rates react to counter the inflationary consequences and hence move against the exchange rate. Within such a model, the sign of the observed exchange-interest rate correlation depends on the nature of the prevailing disturbances. Reduced form econometrics would deliver biased results in any case, however.

Our paper adds to this recent literature which is supportive of the UIP hypothesis. We argue that the normality of residuals is not met in the samples which we have and that skewness and kurtosis in the distributions of the future exchange rate may account for some of the UIP rejections. In fact, it is common knowledge that the financial returns are not normal, that they usually have heavy tails and that they might be skewed. Therefore, it seems odd to test efficiency, which involves the notion that rational market players utilize all available information, and restrict the expectation error to be normal.

In testing the UIP, we propose using additional information provided by option data, which may reveal the structure of the exchange rate expectations. We estimate risk neutral distributions implied by option prices and use them directly for testing. By favouring this approach we commit ourselves to the assumption of investors' risk neutrality. At least for the first approximation, we deem the risk neutrality assumption appropriate for the case of currencies of two major countries with a comparable macroeconomic environment and financial system.

The idea of utilizing the information carried by options in this context is not new. Related literature includes Lyons (1988), who by assuming lognormality used option-implied volatilities to calculate conditional second moments of expected returns. Based on them he modeled the risk premium by using portfolio balance model. Relying on the regression-based tests he concluded, as did the traditional literature based on the time series estimates of variance, that the forward unbiasedness hypothesis fails; he also found implausible parameter estimates for the rational risk premium model. Malz (1997b) generalized Lyon's (1988) approach by avoiding the assumption of lognormality. He estimated the whole implied risk neutral distributions (RNDs) from the over-the-counter (OTC) options, which allowed him to enhance regression tests with the third and fourth distributions' moments. He found that if currency excess returns are regressed on variance, skewness and kurtosis then plausibly signed coefficient estimates are obtained for several currency pairs including dollar-yen, with a highly significant coefficient for the variance[1]. However, due to a low significance of the skewness for the excess return in the univariate regression, he concluded that there is little evidence of peso problem effects. Moreover, enhancing the traditional Fama regression with distributions' skewness did not help; the coefficient for forward premium remained virtually always negative. The potential of options for capturing the peso problem was also studied by Bates (1996) who used option prices to estimate parameters of the jump-diffusion process[2]. By conditioning on them he extended the traditional regression (e.g. Hodrick and Srivastava, 1987) of futures returns on the interest rate differential. Interpreting the estimates,Bates (1996) concluded that jump-diffusion distributions implicit in dollar-mark month options described the ex-post distribution very poorly and that for dollar-yen these distributions contained no information whatsoever for the subsequent evolution of the dollar - yen futures price. He argued that although the implied distributions might serve as a barometer of market sentiment, it did not appear that the peso problem consequences of skewed and fat-tailed distributions, implied at times by currency futures options, could explain rejections of uncovered interest parity.

[1] This finding was confirmed also by Gereben (2002) for the New Zealand dollar.

[2] Malz (1996) applied a similar model to estimate realignment probabilities before the EMS crisis.

Our way of using options for the tests of the UIP hypothesis differs from the approaches mentioned above. Being aware of pitfalls of single equation estimation pointed out by Barnhart et. al (1999) we avoid the regression estimation altogether. Instead, we estimate a class (parametrized by a single parameter) of implied distributions, and then use these distributions for testing directly. When testing for the zero mean of the observed expectation errors we use the implied RNDs for mapping the errors, without any information loss, to the independent standard normal observations. We cannot reject the hypothesis in either of the subsamples containing uncorrelated observations of non-overlapping contracts. The same result is obtained when all the observations are pooled into one sample by estimating its covariance structure. This result is quite robust with respect to the parametric form of the distribution. In turn, this means that the UIP hypothesis is not rejected on conventional levels for maturity of one month. By way of comparison, we ran the same tests using the distributional assumptions of the expectation errors assumed by the conventional linear regression. We cannot reject the hypothesis either, although the results appear weaker.

Furthermore, we test the more general hypothesis that RNDs are good approximations for the true distributions of the future spot rate. Our approach to testing differs from the one applied by Ait-Sahalia, Wang and Yared (2001) to the densities implied by S&P options. In contrast to Ait-Sahalia, Wang and Yared (2001), who compared RNDs implied by option prices with RNDs estimated from the dynamics of the underlying asset under the assumption that this dynamics is a diffusion, we avoid making any assumptions about the nature of the true stochastic process. In fact, under the maintained joint hypothesis of market' rationality and risk neutrality and given the parametrization of the option implied RNDs we transform realized asset prices in a way that allows us to test whether these transformations were drawn from a single, e.g. particular multinomial, or standard normal density. To this end, we apply three tests: a variant of the Pearson test, a likelihood ratio test, and the Kolmogorov-Smirnov test for the whole class of distribution parametrizations.

The tests are powerful enough to discriminate well between different parametrizations of the RND estimated from option prices. We found that lognormality was strongly rejected, as was the Malz (1997a) distribution. However, for some more fat-tailed distribution in the class, we found quite a good fit measured by the p-values of the tests. This result legitimizes the hypothesis that actual realizations of the exchange rate were drawn from the implied risk neutral distributions.

For estimation of the RNDs we utilized two methods, each applied to a different market segment. To estimate RNDs from the Chicago Mercantile Exchange data we employed a non-parametric Jackwerth and Rubinstein (1996) methodology. For estimation of RNDs from OTC options we used the technique of Malz (1997a), of a quadratic extrapolation of the volatility smile in the delta space, which was generalized by Cincibuch (2003) to account for very heavy tails of the distributions.

2 Risk neutrality, rationality and the UIP hypothesis

The uncovered interest rate parity hypothesis asserts that an expected change in the spot rate compensates the interest rate differential between two different currencies when a risk neutral agents form their expectations rationally:

$$\frac{E_t(S_T)}{S_t} = e^{(i_t - i_t^*)(T-t)}, \tag{1}$$

where the right hand side of the equation refers to an interest rate differential between the domestic (i) and foreign (i^*) currency continuously compounded interest rates with appropriate maturities, and where S denote spot exchange rates, observed in the respective time period. The operator $E_t(.)$ denotes rational expectations based on the information at t. As long as the investors are rational, the market expectations coincide with mathematical expectations. Let $\Lambda_{t,T}$ denotes the actual distribution over which the expectation in equation (1) is taken.

There is a close relationship between the expectations related to the UIP hypothesis and the arbitrage based covered interest rate parity (CIP). Under condition of no arbitrage opportunity, the synthetic forward rate determined by the CIP is equal (up to a difference possibly due to transaction costs) to the actual market forward rate $F_{t,T}$. The CIP market clearing condition is then:

$$\frac{F_{t,T}}{S_t} = e^{(i_t - i_t^*)(T-t)}. \tag{2}$$

The CIP holds in practice with greatest accuracy between major currencies, which was confirmed by numerous studies, e.g. Fratianni and Wakeman (1982). Therefore, it follows that the parity (1) is practically equivalent to the condition that forward rate is an unbiased predictor of the future spot rate

$$E_t[S_T] = F_{t,T}. \tag{3}$$

Therefore, conditions (1) and (3) are often used interchangeably and we may conveniently focus on a later formulation (3) in search of testable specifications of the hypothesis.

The hypothesis can be generalized if assumptions of rational and risk neutral pricing are extended from forwards to other contingent contracts. It is a well established theoretical result that the price of a traded security can be expressed as an expected discounted security payoff where the expectation is taken with respect to an appropriate risk-neutral measure (Cox and Ross, 1976; Ross, 1976). If $c\,(S,X)$ is the price of a European call option with underlying asset S, strike X and maturity T then the the risk neutral distribution of the spot rate on the future date T denoted by $\Lambda_{t,T}^{RN}$ is implicitly defined by

$$c\,(S,X) \equiv e^{-i(T-t)} \int_0^\infty \max\,(S_T - X, 0)\, d\Lambda_{t,T}^{RN}\,(S_T)\,, \tag{4}$$

where r denotes the domestic risk free interest rate and S_T is the random spot rate at the option's maturity T. The distribution $\Lambda_{t,T}^{RN}$ can be estimated from option prices and other market data as of the current date t.

In order to reformulate (3) in terms of risk neutral measure let us first observe that under no arbitrage opportunity[3] the European option with zero strike is priced as

$$c\,(S,0) = Se^{-i^*(T-t)}, \tag{5}$$

where i^* denotes foreign interest rate. Further, from equation (4) it follows that

$$c\,(S,0) = e^{-i(T-t)} \int_0^\infty S_T d\Lambda_{t,T}^{RN}\,(S_T) = e^{-i(T-t)} E_t^{RN}[S_T]. \tag{6}$$

From (5) and (6) and from arbitrage based covered interest rate parity (2) it directly follows that forward rate is always the mean of the risk neutral distribution:

$$F_{t,T} = E_t^{RN}[S_T].$$

[3] Under no arbitrage opportunity two portfolios with the same payoff profile should have the same price. The call option $c\,(S_t,0)$ pays at maturity in any case S_T and the foreign currency riskless bond market investment with valued $S_t e^{-r^*(T-t)}$ also pays at maturity S_T. Thus, the relationship (5) holds.

Thus one may restate the hypothesis (1) or (3) as

$$E_t^{RN}[S_T] = E_t[S_T]. \tag{H_0}$$

In other words, the risk neutral form of the UIP stipulates equality of expectations taken with respect to the risk-neutral measure implied by prices of traded securities and the one taken with respect to the actual distribution of the asset.

The natural generalization of the claim H_0 is to hypothesize not only equality of the first moments but equality of the whole distributions:

$$\Lambda_{t,T}^{RN} \approx \Lambda_T. \tag{H_0^*}$$

In a logical sense this hypothesis is obviously stronger than that of the ordinary UIP, but economically the same argument is behind both H_0 and H_0^*. The UIP claims that on average it should not make any difference whether one takes a short or long position in the forward contract. However, from (5) and from covered interest rate parity it follows that $F = e^{i(T-t)}c\,(S, 0)$ and therefore we may understand the hypothesis H_0 as a claim that one option contract with one particular strike is fairly priced. Indeed, it is natural to assume that on average it should not make any difference whether one takes a short or long position in any option contract with any exercise price. This directly leads to hypothesis H_0^*, which is thus conceptually equivalent to hypothesis H_0.

3 Data

We base our tests on the dollar-yen currency pair, because it represents one of the deepest option markets. Moreover, due to the relatively wide interest rate differential and distinct phases of the business cycle in Japan and USA, we might expect severe violations of the UIP for dollar-yen pair. For the parametric methods and econometric regressions we work with the OTC market, where we dispose of time series of dollar-yen spot, forward and option quotes. The series start in 1992 and finish August 2000, i.e. they represent 2124 daily observations, provided by two large market makers. Specifically, the data consist of time series of at-the-money-forward (ATMF) volatilities, 25-delta risk reversals and 25-delta strangles for one-month options together with appropriate forward rates. From the OTC quotes, we backed out implied volatilities for three exercise prices. Due to the fact that the OTC market quotes options in terms of implied volatilities, the data do not suffer from the problem of stalled prices which sometimes occurs in the data from the exchanges.

Another source of option data is the Chicago Mercantile Exchange, in particular we use close-of-business data for dollar-yen currency futures and actually traded dollar-yen currency futures option contracts. The dollar-yen parity is convenient, because here options are traded with enough liquidity at the CME and therefore the estimation of RNDs can be double-checked. The CME data is also employed in the nonparametric method of deriving RNDs. The source of interest rate data is the Bloomberg database.

In the sampling procedure we have to confront possible problems stemming from overlapping data. To avoid mutually dependent observations we construct non-overlapping samples. Specifically, we design a sampling procedure whereby we obtain several subsamples, each of them containing such dates that between the trade dates of options and their maturity dates no other observation occurred in a particular subsample. We undertake the tests only on subsamples containing at least 50 observations. We found 13 such subsamples and we perform each test individually for every subsample. While this is correct in principle, potential efficiency may be lost from considering fewer data in isolation. Ideally, we would like to have one statistic to decide about a particular hypothesis, and not 13. We therefore design a method that corrects for serial correlation of the observations and provides with full sample results as well.

4 Regression results

Our approach has been motivated by some unrealistic assumptions behind UIP tests based on the Fama regression estimates. It is therefore instructive, by a means of reference, to see how the conventional tests perform in our data sample. Let s_T and $f_{t,T}$ denote the natural logarithms of S_T and $F_{t,T}$ respectively. Then the following testing specification

$$s_T - s_t = a + b(f_{t,T} - s_t) + u_T, \tag{7}$$

is easily derived from (3) if rational expectations are assumed and the Jensen inequality issue is neglected. Then standard assumptions are usually placed on the regressors and the error term. The specification was used in many tests of the forward rate unbiasedness (e.g. Fama, 1984) and hence also the UIP hypothesis. The UIP hypothesis implies $a = 0$ and $b = 1$ in (7), but the research on UIP has concentrated mostly on testing $b = 1$, because it is usually argued that the non-zero constant term could be accounted for by an average risk premium and it might include also the Jensen's inequality term. Rewritten in the matrix form the regression (7) reads

$$\Delta s = \beta X + u. \tag{8}$$

The vector $\beta = (a, b)$ is formed by regression coefficients, the data matrix X has two columns, the first consisting of ones and the second containing forward premia (interest rate differentials) and u is the vector of prediction errors.

Data are sampled daily, yet the prediction horizon is about one month. Because of ensuing serial correlation, methods based on the assumption that observations are independently and identically distributed can not be immediately utilized. One option is to avoid mutually dependent observations altogether and test the hypotheses on subsamples consisting only of forward contracts that do not overlap as it was done for example by Fama (1984). Estimation results for these subsamples are shown in the first thirteen rows of Table 1.

However, potential efficiency is likely to be lost from considering fewer data in isolation. Therefore, we make an attempt to use all data in one test by estimating a covariance structure among them. In the Section 11.1 we derive the explicit form of the covariance structure $var(u)$ stemming from the maturity of the contracts which is larger than the sampling period. The same approach we were able to use again when the moving average correlation problem is tackled for other tests of hypotheses.

Specifically, we assumed that the prediction errors per business day are homoskedastic with variance σ^2 and ucorrelated. As $\tau(t)$ denotes maturity in business days for the contract concluded on t, we write

$$u_t \sim N\left(0, \tau(t)\,\sigma^2\right). \tag{9}$$

Under theses assumptions for day to day disturbances we derived matrices P and F such that the components of new series $\tilde{u} = uPF$ are identically independently distributed, i.e. $\tilde{u} \sim N\left(0, \sigma^2 I\right)$. Thus, instead of regression (8) the following transformed equation is estimated[4]

$$\Delta s PF = \beta XPF + \tilde{u}.$$

[4]For this purpose, either the Hansen and Hodrick (1980) or Newey and West (1987) way of estimating asymptotic covariance matrix was often pursued in the literature.

Full sample estimation results of these are shown in the last line of the Table 1. They are consistent with regressions on subsamples, and the narrower confidence intervals show higher efficiency gained by pooling the data.

The results are similar to those routinely found by other authors for most currencies and time periods. As usual, $b = 1$ can be rejected with enough confidence and the estimates are negative for all subsamples, a is insignificantly different from zero and the overall fit of the regression, measured to be R^2 is negligible. The latter fact is disturbing, because it may point to a misspecified equation. Alternatively, a low R^2 may come from a high variance of the error term.

The full sample result is close to what was found by Flood and Rose (2001) for dollar-yen in 1990 in their estimate also based on the daily observations. For the one month maturity they report an estimate of b equal to -1.17 with the Newey-West standard error of 1.11. The standard error of our estimate is about 1.25[5].

Because $\hat{a}$ is insignificant, we impose this as a restriction and re-run this more parsimonious specification:

$$s_T - s_t = b(f_{t,T} - s_t) + u_T. \tag{10}$$

Again, estimate of b turns out to be negative ($\hat{b} = -1.1$) and insignificant, but remarkably different from unity with a 5% confidence interval of $(-2.87, 0.66)$. This form of the Fama regression is interesting, because we may use it to express the expectation error (or, equivalently, the excess return on the domestic asset) under rational expectations as:

$$\begin{aligned} s_T - E_t(s_T) &= s_T - f_{t,T} \\ &= s_T - s_t + s_t - f_{t,T} \\ &= (b-1)(f_{t,T} - s_t) + u_T, \end{aligned} \tag{11}$$

where in the last equality we made use of equation (10). Under rationality, expectation errors, $s_T - E_t(s_T)$,, should be zero on average, which under the conventional assumptions for u_T implies zero mean of the term $(b-1)(f_{t,T} - s_t)$ as well. Surprisingly, as the first line in Table 2 shows, we cannot reject nullity of the excess return either for the full sample or for any of the subsamples. This means that for a risk neutral investor it does not pay to speculate against the interest rate parity. However, examining the right hand side of (11) we find that the forward premium in our sample is significantly different from zero and we showed earlier that b is significantly different from unity, which is a contradiction.

5 Measuring the "UIP failure" by implied distributions: An example of non-parametric approach

The purpose of this section is to undertake an illustrative 'test' of the UIP hypothesis by constructing the RND implied by currency option prices on two different days when the UIP gave seemingly misleading predictions about a change in the dollar-yen exchange rate. We used end-of-day settlement prices from

[5]The difference between the two estimates is most likely mainly due to a slightly different data sample. While we have fewer than 9 years of data between 1992 and 2000, the Flood and Rose (2001) estimates are based on a decade of data of the 1990s.

the CME from 4 January 1999 on the options on yen futures contracts expiring in June 1999[6]. The date was chosen because it lies significantly ahead of the expiration date and the forward rate (i.e. the unbiased predictor according to the UIP) predicted an exchange rate movement over the expiration period in an opposite direction from what actually happened. On 4 January yen spot rate was at 112. The futures rate (June contract) on this day stood at 109.7 Yen/USD. In other words, the UIP predicted an appreciation of the yen by 2% over the 6 months left to maturity. In reality, though, on 7 June 1999 (the last day of options trading) the exchange rate stood at 120.8 Yen/USD, i.e. a depreciation more than 7%.

Next, we checked how the prediction given by the UIP in January 1999 about the exchange rate in June 1999 is consistent with implied distributions. We estimated a risk neutral distribution implied by the options' prices on this day using a non-parametric method developed by Jackwerth and Rubinstein (1996). They construct a smooth density function by minimizing the norm that measures the density function's second derivative, constrained by the condition that option prices are discounted expectations of the options' payoffs at maturity. Like these authors, we chose among the proposed objective functions the maximum smoothness criterion as it does not require any prior about the true distribution[7].

The method is designed for European options, but for the sake of simplicity, using only out-of-the-money options, we chose to ignore the issue of the early exercise premium associated with the CME contracts, which are American futures options. Indeed, Whaley (1986) reports an early exercise premium to be significant only for deep in-the-money options. For a other way of estimating RNDs implied by American futures options see Cincibuch (2003).

The estimation procedure involves a subjective choice of the smoothness parameter; however for the given purpose all reasonable candidate distributions are very similar and they yield virtually the same test statistics. We report one of the obtained distributions in the Figure 1, marking all the relevant information: the spot rate on the date of purchase, the mean of the distribution (almost identical to the futures rate on the same date), and the actual (futures) rate at the option's expiration.

We note that the distribution is slightly skewed towards appreciation of the yen, causing that the modus - most likely outcome - represents some depreciation. Perhaps more importantly, we observe that the distribution is relatively wide. Because the means of the distributions coincide with the forward rate expectations, we may actually test for the hypothesis that market expectations are consistent with the UIP. The probability level at which the realization of the exchange rate in June would be sufficient to reject the null of the mean equal to the forward rate is 21%. This shows that if the RND reflected the consensus expectations then the actual outcome would not be unexpected

6 Estimation of parametric RNDs from OTC data

The non-parametric approach is not suitable for formal testing of the UIP hypothesis as it is relatively intensive in the number of strikes available for each maturity date. Specifically, in order to derive RNDs of market expectations in the above example more than a dozen of exercise prices was used. Such numbers of strikes are only available on organized market exchanges. However, the relatively small number of

[6]The use of end-of-day prices is sometimes criticised for the possible existence of asynchronous, so-called stalled quotes and using actual tick data is suggested. We consider this a minor objection given that settlement prices are set at the end of each trading day by a committee of CME officials, because these prices are used for margining purposes. As such they should have a good information value.

[7]In fact, it is equivalent to fitting a cubic spline between the available option prices. In addition, the smoothness criterion has the advantage of providing a closed form solution of the searched option prices, which greatly facilitates the computations. Also other modalities of the estimation including non-negativity constraints and asymptotic conditions can be found in the original paper.

maturity days available there[8] effectively prevents the non-parametric methods from rigorous testing of the UIP hypothesis. Moreover, not always is the organized exchange market deep enough to consider the available option prices for traded strikes as representing true market prices. This also limits the use of exchange traded options in the UIP testing to occasional examples of RNDs for particular data points, such as those shown above.

Instead, we resorted to a parametric approach to estimation RNDs from the option prices traded on the OTC market. This has the advantage of a relative abundance of maturity days (in fact, virtually every business day is a maturity day for some 1 month contracts). The disadvantage is that the OTC market usually quotes only three benchmark strike prices, hence some parametric inter-, and extrapolation of the implied distribution has to be used. This approach nonetheless still allows for far less restrictive assumptions about the distribution of exchange rates than those implied by conventional methods of UIP testing.

The OTC data based techniques are formulated in terms of the market convention of quoting option prices. Not only do OTC market participants quote currency option prices most often in volatility[9] terms, but the OTC market also developed a way of normalizing of the option moneyness. Instead of specifying an exercise price in dollars, OTC market participants use an option's delta. Thus, instead of dollar exercise price - option price pairs, traders usually quote options in delta - volatility terms[10]. This convention allows them to abstract from immediate changes of the spot rate, which together with interest rates determines the first moment of the RND, and focus on the options' substance, i.e. on the nature of uncertainty inherent in the dynamics of the spot rate.

It is not difficult to show that equation (4) together with the Black-Scholes formula establish an equivalence between RND and the volatility smile (e.g. Cincibuch, 2003). Because of this equivalence, it is possible to view the RND estimation as a curve fitting problem for the volatility smile. From the OTC market, only volatilities for three benchmark deltas are readily available[11]. Therefore, it is natural that methods interpolating these three points by a quadratic function have been suggested in the literature.

While Shimko (1993) proposed fitting volatilities by the quadratic function in the volatility - dollar exercise price space, Malz (1997a) put forward the quadratic function for the smile in the volatility - delta space. The transformation between dollar exercise prices and deltas is highly non-linear and therefore both approaches lead to quite distinct RNDs. In particular, the distributions differ in their fat-tailedness. Moreover, Malz (1997a) noted that while the quadratic smile fitted into the dollar space is problematic since it might break non-arbitrage constraints for the volatility function, the quadratic smile in delta space avoids this problem. On the other hand, Cincibuch (2003) showed that a quadratic smile in the delta space significantly underestimates volatilities implied by exchange traded options that were deeply out or in the money.

The method which we used for the OTC data is a straightforward generalization of Malz's approach (Malz, 1997a) and it is documented in Cincibuch (2003). This method allows for controlling the degree of fat-tailedness of the estimated distribution. The market uses hedging deltas instead of dollar exercise

[8] E.g. on CME only twelve per year.

[9] Under the Black-Scholes model the spot rate follows a geometric brownian motion and the terminal RND is lognormal. The volatility parameter is the only unobservable variable in the Black-Scholes pricing formula, and therefore this formula can be used as mapping that converts volatilities into prices and vice versa even, in situations where the Black-Scholes assumptions do not hold.

[10] The functional relationship between excercise prices and volatilities is often called the volatility smile. Under the Black-Scholes model the volatility smile degenerates into a horizontal line.

[11] Usually, volatilities for excerice prices corresponding to 0.25, approximately 0.50 and 0.75 delta of call options are quoted.

prices as an alternative measure of option's moneyness and the transformation between these two [12] involves the cumulative density function of the standard normal distribution. Thus, another 'generalized delta' is obtained by replacing the transforming standard normal by another distribution. In particular, by changing the standard deviation of the transforming normal distribution from unity to G we get a class of moneyness spaces. Then G is optimized so that the quadratic function interpolating the OTC quotes fits better peripheral options traded on an organized exchange (CME). It follows from the functional forms involved that higher G makes the resulting RND more fat-tailed.

To enhance the robustness of our conclusions we present the results of extensive testing of the hypotheses for a spectrum of values of the fat-tailedness parameter G. It turns out that at least for the dollar - yen, raising G above 1 significantly improves this fit. It is sufficient, though, to consider G approximately between 2 and 10. It shown in the original paper that the appropriate value of parameter G may enhance the interpolation of the OTC prices by information from exchange traded quotes.

7 Monte Carlo Experiment

Empirical results presented in Section 4 point to a possible misspecification of the Fama regression. We take this as evidence that vindicates alternative approaches to testing for the UIP, such as those based on implied RNDs introduced in this paper. In order to strengthen this point we design a Monte Carlo experiment in which we let hypothetical realizations of the exchange rate in our sample dates be drawn from RNDs implied by option prices at these dates. We generate 20000 vectors of 'future' spot rates and run the regression (8) on these data sets.

These artificial data sets were drawn under assumptions of forward rate unbiasedness, so if the estimator of b is consistent, it should approach unity in large samples. We therefore gradually increase the length of the sample to investigate the small sample properties of the estimator. We display the results in Figure 2. The horizontal axis measures the number of observations used for regression estimates and the vertical axis shows the estimate of the coefficient b. The crosses present the mean of 20000 estimates of b, while the punctuated lines show the 95% confidence intervals of these means. We observe a downward small sample bias, which seems to be disappearing only slowly. Following Krasker (1980), we argue that this small sample bias is likely caused by leptokurtic nature of the sampling distributions.

8 Tests of zero means of expectation errors

The traditional approach to testing market efficiency exploits only exchange rate and interest rate data and imposes a normality restriction on the market's expectations. This assumption seems inappropriate given widespread awareness among market participants of the non-normality of returns. The purpose of this section is to introduce a straightforward test of forward rate unbiasedness, which relaxes the normality assumption and takes into account this market opinion of the uncertainty inherent in the dynamics of the spot rate. This opinion is contained in prices of currency options and can be summarized by implied RNDs.

For the purpose of the test we have to address the fact that for each observation S_T there is a different theoretical distribution $\Lambda_{t,T}^{RN}$. Next, we describe the standardization procedure that allows using familiar statistical methods. Provided that theoretical cumulative density functions $\Lambda_{t,T}^{RN}$ are increasing and if the

[12] It is worthwhile to emphasize that using the options' delta as a measure of moneyness is a pure market convention.

hypothesis $S_T \sim \Lambda_{t,T}^{RN}$ holds then the random variables $\Lambda_{t,T}^{RN}(S_T)$ are uniformly distributed on the unity interval. Further, let Φ_ν denote a cumulative distribution function of the normal distribution with zero mean and variance ν^2. Then, it is obvious that $\Phi_\nu^{-1}\left(\Lambda_{t,T}^{RN}(S_T)\right) \sim N\left(0,\nu^2\right)$. This transformation yields normalized deviations from theoretical means[13].

Moreover, since the observations are sampled more often than option and forward contracts mature, we have to control for the moving average serial correlation. For serial correlation adjustment, which is described in Section 11.1, it is useful to keep the standard deviation of the normalized errors proportional to the number of business days between date t, when a contract was concluded, and its maturity day T. Let this number is denoted by $\tau(t)$. Its dependence on t stems from the fact that interbank market contracts like $1M$ are of fixed maturity only approximately. Weekends and holidays introduce some irregularities. Therefore, the normalized errors are constructed as

$$\epsilon_t^{\tau(t)} \equiv \Phi_{\tau(t)}^{-1}\left(\Lambda_{t,T}^{RN}(S_T)\right) \tag{12}$$

from which the hypothesis follows that

$$\epsilon_t^{\tau(t)} \sim N\left(0, \tau(t)^2\right). \tag{13}$$

It is shown in Proposition 1 of Section 11.1 that under modest assumptions regarding the nature of serial correlation it is possible to give the explicit form of the covariance matrix $var(\epsilon^\tau)$ and to transform the observations $\epsilon_t^{\tau(t)}$ into mutually independent and standard normally distributed variables $\widetilde{\epsilon}_t^\tau \sim N(0,1)$.

We employ a t-test to examine whether $E(\widetilde{\epsilon}^\tau) = 0$ in the case of the full sample (with removed serial correlation), and whether $E(\epsilon^\tau) = 0$ in the case of individual subsamples of uncorrelated observations. In fact, one could use the z-test instead of the t-test, because the theoretical variance is known, but by using only the t-test we try to diminish the dependence on how accurately the distributions' dispersion is estimated. We return to the more general hypothesis of similarity of the theoretical and actual distributions in Section 9.

In the first row of Table 2, we demonstrate that the average excess return in the available samples is indeed insignificantly different from zero, assuming normality of i.i.d. regression error terms. Relaxing the normality and i.i.d. of regression error terms, the last three rows of the table contain confidence intervals for the mean of standardized excess returns for various RNDs estimation methods. Evidently, the confidence intervals center around zero irrespective of the method or sample. We thus conclude that the inability to reject the assumption of forward rate unbiasedness is a robust finding with respect to the RNDs estimation method. In fact, even if the more demanding z-test is used instead of the t-test, there would be no rejections.

9 Tests of distributions equality

In this section, we present the results of three tests of the stronger hypothesis that empirical distributions are well approximated by estimated implied RNDs. Thus, we test rationality and risk neutrality of investors more broadly, since the forward rate unbiasedness concerns only the first moments of the

[13] In fact, the mapping $\Phi_\nu^{-1} \circ \Lambda_{t,T}^M$ generalizes the procedure of standardization of heteroskedastic observations. Indeed, for $\Lambda_{t,T}^M = N\left(\mu_T, \sigma_T^2\right)$ and $\nu = 1$, it would boil down to the usual linear transformation: $\Phi_\nu^{-1} \circ \Lambda_{t,T}^M(S_T) = \frac{S_T - \mu_T}{\sigma_T}$.

distributions. By accepting this hypothesis we would also accept the forward rate as an unbiased predictor of the future exchange rate. Indeed, the converse is not true. If the hypothesis about implied RNDs were rejected, the forward rate unbiasedness would not be yet excluded.

The tests outlined here are used both for full sample data transformed into identically independently distributed random variables and also for subsamples containing non-overlapping contracts with variant theoretical terminal distributions. In the case of subsamples of serially uncorrelated observations the random variables Y_j stand for the actual realizations of the spot rate S_T and the cumulative distribution functions Ψ_j represent estimated risk neutral distributions $\Lambda_{t,T}^{RN}$.

Before the tests can be used for the full sample, the serial correlation has to be accounted for. It is done in the very same way as described in Section 8. In this case, the random variables Y_j stand for the standardized excess returns $\widetilde{\epsilon}_t^{\tau(t)}$ and all the functions Ψ_j collapse to the cumulative density function of the standard normal distribution.

Let $Y_1, Y_2, ..., Y_n$ be independent random variables and let the theoretical cumulative distribution functions $\Psi_{1,}\Psi_2, ..., \Psi_n$ of the random variables are all increasing. Let y_j denote a realization of the variable Y_j, $j = 1, 2, ..., n$. Then, given that $Y_j \sim \Psi_j$, the random variable $\Psi_j(Y_j)$ is uniformly distributed on the unity interval. Also, because Y_j are independent variables, the transformed observations $\Psi_1(y_1), \Psi_2(y_2), ..., \Psi_n(y_n)$ represent independent random draws from the uniform distribution. Further, divide the unity interval into k subintervals $U_i = \left\langle \frac{i-1}{k}, \frac{i}{k} \right)$, $i = 1, 2, ...k$. Then, provided that $Y_j \sim \Psi_j$, events $\Psi_j(Y_j) \in U_i$ are multinomially distributed with equal probabilities $p_0 = \frac{1}{k}$of outcomes. We tested the hypothesis $H_0 : Y_j \sim \Psi_j$ for each j using Pearson and generalized likelihood ratio χ^2 tests for multinomial distributions.

Let N_i be the number of actual realizations that fell into the $i - th$ bracket, i.e.

$$N_i = \left| \{ \Psi_j(y_j) \in U_i \ ; \ j = 1, 2, ..., n \} \right|.$$

Then the Pearson's test statistics is

$$Q_k = \sum_{i=1}^{k} \frac{(N_i - np_0)^2}{np_0}.$$

Under the H_0 the statistics Q_k is asymptotically χ^2 distributed with $k-1$ degrees of freedom. If α is an approximate size of the test and $\chi^2_{1-\alpha}(k-1)$ is $(1-\alpha)\,th$ quantile of $\chi^2(k-1)$ then the test is

$$\text{Reject } H_0 \text{ if and only if } Q_k > \chi^2_{1-\alpha}(k-1).$$

The alternative way for testing the hypothesis is to use a variant of the generalized likelihood ratio test, which is also a uniformly powerful test. The likelihood ratio for the hypothesis of equality of distributions can be computed as

$$\lambda = n^n \prod_{j=1}^{k} \left(\frac{p_0}{N_j} \right)^{N_j}.$$

Under the H_0 the statistic $-2log\lambda$ is asymptotically $\chi^2(k-1)$ distributed.

These two tests stem from transforming the problem of distributions equality to testing parameters of the multinomial distribution. The dimension of this multinomial distribution as well as the values of its parameters are variables of choice; in other words, we may arbitrarily select the number of quantiles and their relative sizes. A legitimate question would then be, what is the optimal multinomial distribution. In fact, there is a trade off; a lower number of quantiles would leave more data for each distribution's parameter and thus it would lead to a more powerful test. On the other hand, a small number of quantiles might poorly capture the distribution's shape. For example, a test based on two equal quantiles would focus only on the equality of the distributions' medians. Therefore, it might cause an acceptance of the hypothesis for two quite distinct distributions with the same median. At the same time due to the relative abundance of data this test would lead to a rejection of the hypothesis for two quite similar, perhaps economically equivalent distributions with only slightly different medians. For the sake of completeness however, we present results for the whole range of quantiles. The roughest division we present is into five quantiles, which corresponds to the four-dimensional multinomial distribution. Note that the estimated distributions are also parametrized by four parameters (three for the quadratic smile, one for the G-space). Regarding the other extreme we recall the rule of thumb for the Pearson test, which stipulates that the theoretical frequencies should be at least six. This would imply for our number of observations the finest division into more than two hundred quantiles. However, we deem that such fine quantiles do not correspond to the nature of the problem. We therefore report a maximum of eighty quantiles, which leads to theoretical frequencies of about fifteen.

The third test we used to assess H_0^* does not hinge on the multinomial distribution. It is a Kolmogorov-Smirnov one sample test. If J_N is the empirical cumulative distribution function then the test statistic is

$$\sup_t \left| J_N\left(\widetilde{\epsilon}_t^{\tau}\right) - \Phi_1\left(\widetilde{\epsilon}_t^{\tau}\right) \right| . \tag{14}$$

The critical values for this statistic are tabulated.

Table 3 exhibits p-values of Pearson's test and the likelihood ratio for the full sample for different values of the fat-tailedness parameter G and number of quantiles k. The last column shows p-values for the Kolmogorov-Smirnov test. Tests are strong enough to distinguish comfortably between different parameters G. For example, the hypothesis that actual observations are consistent with the Black-Scholes lognormal model is comfortably rejected. The empirical distribution of the full 1247 observations into 40 quantiles under the lognormal hypothesis is also plotted on Figure 3 and the reason for the rejection is obvious. In reality, more observations fell to the extreme quantiles than was predicted by lognormal distribution. Further, the record for $G = 1$ shows that neither Malz RND's tails are not heavy enough (Figure 4 illustrates this case). This bodes very well for the efficiency of the option market, because as Cincibuch (2003) found, the Malz's extrapolation of the OTC benchmark volatilities underestimates the prices of CME options with extreme strikes, i.e. the that market thinks that the actual distribution of returns allows for extreme events more often than predicted by Malz's function. Only tails of the distribution with $G = 2$ are already heavy enough to conform to the data. Figure 5 reveals that for this calibration the theoretical distribution leads to a relatively good approximation of the uniform density. In fact, if the sizes of the tests are set to 10%, only this distribution is not rejected. In addition, this parametrization is not completely arbitrary, it was found by Cincibuch (2003) that this parametrization also corresponds relatively well to option prices for extreme strikes traded on the CME.

For the sake of completeness, in Table 4 we report a summary of results for the mutually correlated subsamples of the serially uncorrelated observations. However, it is obvious that the power of the test is quite low for 100 or fewer observations in a single sample.

It emerges from the results of the three tests that the available evidence does not contradict the hypothesis that the implied RND is a good approximation of the true distribution of expectation errors. Moreover, the tests of zero means of expectation errors developed in the preceding subsections are more focused and therefore we argue that they are more powerful as concerns the proper UIP hypothesis. In any case, a variety of tests check for the robustness of our results.

10 Conclusion

The purpose of the paper was to point at the potential significance of non-restricted distributions of expectations in testing for the UIP hypothesis. We corroborate the idea that the conventional tests, which in their majority tend to reject the hypothesis, may have erred by using overly restrictive assumptions about the shape of the distribution of error terms, or by simply misspecifying the reduced form regression relationship. As an alternative, we propose using risk-neutral probability distributions implied by currency option prices. First, we illustrate one such method on a case when the UIP apparently seemed to give misleading predictions about the future development in the dollar-yen exchange rate, yet the kurtosis and partly the skewness of the estimated distribution was such that the observed realization of the exchange rate was not an extreme event. Further, the Monte Carlo simulation shows a significant small sample bias if the traditional regression is run on data drawn from estimated RNDs.

With these results as motivation, we set on rigorous testing of the UIP hypothesis using implied RNDs from OTC option prices. We test whether the observed expectation errors center around zero (as they should under rationality and risk neutrality) using their estimated RNDs. We cannot reject the hypothesis in either of our samples. The same result was obtained in an attempt to pool all the observations into one sample by estimating its covariance structure. By way of comparison, we ran the same tests using the distributional assumptions of the expectation errors assumed by the conventional regression. We cannot reject the hypotheses either, which contrasts with simple regression estimates.

Then we proceeded to the more general hypothesis that actual realizations of the exchange rate are drawn from the implied risk-neutral distribution. A novel way of transforming of the realized spot rates allowed for application of established statistical tests for distributions equality. If the RNDs implied by options are estimated reliably then this course represents a direct testing procedure for joint hypothesis of the option market rationality and risk neutrality. We performed three tests of distributions equality on a class of RNDs estimates based on the OTC data. We found that the tests are strong enough to discriminate among the different parametrizations, which differ mainly in a way how they extrapolate implied distributions' tails. Moreover, we realized that there is a RNDs parametrization which well corresponds with realized spot rates. Essentially, we found an empirical support for market rationality and risk neutrality.

11 Appendix

11.1 Serial correlation

With data sampled daily and a one month prediction horizon, a serial correlation problem arises. In order to exploit all the information contained in the data we model the covariance structure induced by overlapping contracts. Specifically, we assume that the correlation between two normalized errors $\epsilon_t^{\tau(t)}$ and $\epsilon_s^{\tau(s)}$, which were derived in Section 8, is proportional to the size of their mutual overlap. Because variables $\epsilon_t^{\tau(t)}$ are normally distributed according (13), there might be differences of some pure random walk with normal disturbances with some variance σ^2, e.g. $\widetilde{S}_{t+1} = \widetilde{S}_t + \epsilon_t^1$ with $\epsilon_t^1 \sim N\left(0, \sigma^2\right)$ *iid*. Under this notation $\epsilon_t^{\tau(t)} = \widetilde{S}_{t+\tau(t)} - \widetilde{S}_t$ we might write $\epsilon_t^{\tau(t)} = \sum_{i=0}^{\tau(t)-1} \epsilon_{t+i}^1$. The following proposition gives the explicit form for the covariance matrix of the vector $\epsilon^\tau = \left(\epsilon_1^{\tau(1)}, \epsilon_2^{\tau(2)}, ...\epsilon_N^{\tau(N)}\right)$ if day to day spot rate changes ϵ_t^1 are uncorrelated and homoskedastic.

Proposition 1 *If* $cov\left(\epsilon_t^1, \epsilon_s^1\right) = \sigma^2$ *for* $t = s$ *and if* $cov\left(\epsilon_t^1, \epsilon_s^1\right) = 0$ *for* $t \neq s$ *then*

$$var\left(\epsilon^\tau\right) = \sigma^2 \Omega,$$

where components of the matrix Ω *are given by*

$$\Omega_{t,s} = \max\left[0, \min\left[t + \tau\left(t\right), s + \tau\left(s\right)\right] - \max\left[t, s\right]\right].$$

Proof. Let δ_{ts} be components of the identity matrix, $\delta_{ts} = 1$ for $t = s$ and $\delta_{ts} = 0$ for $t \neq s$. Then

$$\begin{aligned}
\left\{var\left(\epsilon^\tau\right)\right\}_{t,s} &= cov\left(\epsilon_t^{\tau(t)}, \epsilon_s^{\tau(s)}\right) = cov\left(\sum_{i=0}^{\tau(t)-1} \epsilon_{t+i}^1, \sum_{j=0}^{\tau(s)-1} \epsilon_{s+j}^1\right) = \\
&= \sum_{i=0}^{\tau(t)-1} \sum_{j=0}^{\tau(s)-1} cov\left(\epsilon_{t+i}^1, \epsilon_{s+j}^1\right) = \sum_{i=0}^{\tau(t)-1} \sum_{j=0}^{\tau(s)-1} cov\left(\epsilon_{t+i}^1, \epsilon_{s+j}^1\right) = \\
&= \sum_{i=0}^{\tau(t)-1} \sum_{j=0}^{\tau(s)-1} \sigma^2 \delta_{t+i,s+j} = \sigma^2 \sum_{m=t}^{t+\tau(t)-1} \sum_{n=s}^{s+\tau(s)-1} \delta_{m,n} = \\
&= \sigma^2 \max\left[0, \min\left[t + \tau\left(t\right), s + \tau\left(s\right)\right] - \max\left[t, s\right]\right].
\end{aligned}$$

■

We assume that under the hypothesis we might write $\epsilon^\tau \sim N\left(0, \sigma^2 \Omega\right)$. Since Ω is a real symmetric matrix there exists a matrix P such that $D = P^{-1} \Omega P$ is diagonal. Moreover P is such that $P^{-1} = P^{'}$ and therefore $D = P^{'} \Omega P$. It follows that $\epsilon^\tau P \sim N\left(0, \sigma^2 D\right)$. Because the matrix Ω is positive definite, all its eigenvalues on the diagonal of D are positive and we can have a diagonal matrix F such that $F_{ij} = \frac{1}{\sqrt{D_{ij}}}$ for $i = j$. Let define $\widetilde{\epsilon}^\tau = \epsilon^\tau P F$. Then one can easily realize that

$$\widetilde{\epsilon}^\tau \sim N\left(0, \sigma^2 I\right). \tag{15}$$

where I denotes an identity matrix.

References

Ait-Sahalia, Y., Wang, Y. and Yared, F., 2001, Do Option Markets Correctly Price the Probabilities of Movement of the Underlying Asset ? Journal of Econometrics 102, 67 - 110.

Baillie, R. and Bollerslev, T., 2000, The Forward Premium Anomaly is Not Bad as You Think. Journal of International Money and Finance 19 (4), 471-88.

Bank of England, 2000, Economic Models at the Bank of England. Bank of England. ISBN 1857301021.

Barnhart, S., McNown, R. and Wallace, M., 1999, Non-Informative Tests of the Unbiased Forward Exchange Rate, Journal of Financial and Quantitative Analysis.

Bates, D.S., 1996, Dollar Jump Fears, 1984-1992: Distributional Abnormalities Implicit in Currency Futures Options. Journal of International Money and Finance 15 (1), 65-91.

Beneš, J., Hlédik, T., Stavrev, E. and Vávra, D., 2002, The Quarterly Projection Model and Its Properties, in Laxton D., Polak, S., Rose, D. (eds.), The Czech National Bank's Forecasting and Policy Analysis System, Czech National Bank and IMF, forthcoming.

Cincibuch, M., 2003, Distributions Implied by American Currency Futures Options: A Ghost's Smile ? Journal of Futures Markets, forthcoming.

Chinn, M. and Meredith, G., 2002, Testing Uncovered Interest Parity at Short and Long Horizons during the Post-Bretton Woods Era. SCCIE Working Paper #02-14.

Cox, J. and Ross, S., 1976, The Valuation for Alternative Stochastic Processes. Journal of Financial Economics 3, 145-166.

Engel, C., 1996, The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence. Journal of Empirical Finance 3, 123-192.

Evans, M.D. and Lewis, K., 1995, Do Long Term Swings in the Dollar Affect Estimates of the Risk Premia? Review of Financial Studies 8(3), 709-42.

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

Flood, R. P. and Rose, A.K., 2001, Uncovered Interest Parity in Crisis: The Interest Rate Defence in the 1990s, CEPR Discussion Paper, no. 2943

Fratianni, M. and Wakeman, L.M., 1982, The Law of One Price in the Eurocurrency Market. Journal of International Money and Finance 1, 307-323.

Gereben, A., 2002, Extracting Market Expectations from Option Prices: an Application to Over-The-Counter New Zealand Dollar Options. Reserve Bank of New Zealand Discussion Paper 2002/04.

Hansen, L. P. and Hodrick, R. J., 1980, Forward Exchange Rates as Optimal Predictors of Future Spot Rates. Journal of Political Economy 88, no. 5, 829-853.

Hodrick, R. and Srivastava, S., 1987, Foreign Currency Futures. Journal of International Economics 22, 1-24.

Huisman, R., Koedijk, K., Kool, C. and Nissen, F., 1998, Extreme Support for Uncovered Interest Parity. Journal of International Money and Finance 17, 211-228.

Jackwerth, C. J. and Rubinstein, M., 1996, Recovering Probability Distributions from Option Prices. Journal of Finance 51 (5), 1611-1631.

Krasker, W. S., 1980, The 'Peso Problem' in Testing the Efficiency of Forward Exchange Markets. Journal of Monetary Economics 6 (2), 269-276.

Lewis, K.K., 1989, Changing beliefs and systematic rational forecast errors with evidence from foreign exchange. American Economic Review 79, 621-636.

Liu, P. C. and Maddala, G. S., 1992, Rationality of Survey Data and Tests for Market Efficiency in the Foreign Exchange Markets. Journal of International Money and Finance 11, 366-381.

Lyons, R. K., 1988, Tests of the Foreign Exchange Risk Premium Using the Expected Second Moments Implied by Option Pricing. Journal of International Money and Finance 7, 91-108.

Malz, A.M., 1996, Using Option Prices to Estimate Realignment Probabilities in the EMS: The Case of Sterling-Mark. Journal of International Money and Finance 15, 717-748.

Malz, A.M., 1997, Estimating the Probability Distribution of the Future Exchange Rate from Option Prices. The Journal of Derivatives, Winter 1997, 18-36.

Malz, A.M., 1997, Option-Implied Probability Distributions and Currency Excess Returns. Federal Reserve Bank of New York Staff Reports 32, November 1997.

Newey, W. K. and West, K.D., 1987, A Simple, Positive Semi-definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica 55-3, 703-708.

Ross, S., 1976, Options and Efficiency. Quarterly Journal of Economics 90, 75-89.

Shimko, D., 1993, Bounds of Probability. Risk 6 (4), 33-37.

Whaley, R. E., 1986, Valuation of American Futures Options: Theory and Empirical Tests. Journal of Finance 41, 127-150.

Sample		Estimates and 5% confidence intervals for coefficients of		R^2	F-stat	p-val
		intercept	forward premium			
# 1	(103 obs.)	0.01 (-0.00,0.02)	-2.88 (-6.54,0.78)	0.024	2.43	0.12
# 2	(102 obs.)	0.01 (-0.00,0.02)	-2.95 (-6.74,0.84)	0.023	2.39	0.13
# 3	(102 obs.)	0.01 (-0.00,0.02)	-2.80 (-6.40,0.80)	0.023	2.37	0.13
# 4	(101 obs.)	0.01 (-0.00,0.02)	-2.78 (-6.41,0.85)	0.023	2.32	0.13
# 5	(101 obs.)	0.01 (-0.00,0.02)	-3.17 (-6.77,0.44)	0.030	3.04	0.08
# 6	(99 obs.)	0.01 (-0.00,0.03)	-3.27 (-7.19,0.65)	0.028	2.75	0.10
# 7	(98 obs.)	0.01 (-0.00,0.03)	-3.09 (-6.89,0.70)	0.027	2.61	0.11
# 8	(96 obs.)	0.01 (-0.00,0.02)	-2.89 (-6.71,0.93)	0.023	2.25	0.14
# 9	(92 obs.)	0.01 (-0.00,0.02)	-2.94 (-6.73,0.86)	0.026	2.36	0.13
# 10	(86 obs.)	0.01 (-0.00,0.03)	-2.91 (-7.22,1.39)	0.021	1.81	0.18
# 11	(82 obs.)	0.01 (-0.00,0.03)	-2.89 (-7.20,1.43)	0.022	1.77	0.19
# 12	(70 obs.)	0.00 (-0.01,0.02)	-1.28 (-5.41,2.85)	0.006	0.38	0.54
# 13	(50 obs.)	0.01 (-0.01,0.03)	-0.58 (-6.38,5.21)	0.001	0.04	0.84
Full	(1247 obs.)	0.01 (-0.00,0.01)	-2.20 (-4.66,0.26)	0.003	3.79	0.05

Table 1: Estimates of the traditional regression

Estimation method	Confidence interval for the full sample	Intersection of confidence intervals for subsamples 1 - 13
Residuals from constrained regression for $a = 0$ and $b = 1$	(-0.000, 0.001)	(-0.007, 0.004)
Lognormal RND	(-0.040, 0.089)	(-0.175, 0.149)
Malz's RND	(-0.041, 0.082)	(-0.174, 0.139)
Generalized Δ-space RNDs (Intersections of conf. intervals for $G = 2, 3, ..., 11$)	(-0.044, 0.072)	(-0.172, 0.127)

Table 2: Confidence intervals (5%) for the mean of excess returns

Method	80 quantiles Pearson	80 quantiles LR	40 quantiles Pearson	40 quantiles LR	20 quantiles Pearson	20 quantiles LR	10 quantiles Pearson	10 quantiles LR	5 quantiles Pearson	5 quantiles LR	Kolmogorov Smirnov
Lognormal	**0.02**	**0.06**	**0.01**	**0.03**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.03**
$G = 1$	**0.04**	**0.05**	0.13	0.15	**0.05**	**0.06**	**0.05**	**0.05**	**0.02**	**0.02**	**0.06**
$G = 2$	0.22	0.23	0.60	0.55	0.26	0.23	0.47	0.48	0.10	0.10	0.21
$G = 3$	0.13	0.13	0.71	0.72	0.14	0.14	0.20	0.20	**0.09**	0.10	0.15
$G = 4$	0.25	0.25	0.21	0.25	0.15	0.17	0.12	0.11	**0.03**	**0.03**	**0.07**
$G = 5$	0.11	**0.08**	0.27	0.26	0.22	0.23	**0.05**	**0.05**	**0.01**	**0.01**	**0.06**
$G = 6$	**0.01**	**0.01**	**0.08**	**0.09**	**0.05**	**0.05**	**0.06**	**0.06**	**0.01**	**0.01**	**0.05**
$G = 7$	**0.00**	**0.00**	0.15	0.15	**0.09**	**0.09**	**0.06**	**0.06**	**0.01**	**0.01**	**0.05**
$G = 8$	**0.01**	**0.01**	0.15	0.16	**0.04**	**0.04**	**0.06**	**0.06**	**0.01**	**0.01**	**0.05**
$G = 9$	**0.01**	**0.01**	0.11	0.12	**0.06**	**0.06**	**0.07**	**0.07**	**0.01**	**0.01**	**0.05**
$G = 10$	**0.02**	**0.02**	**0.09**	0.10	**0.06**	**0.07**	**0.09**	**0.09**	**0.01**	**0.01**	**0.06**
$G = 11$	**0.02**	**0.03**	0.10	0.11	**0.08**	**0.08**	**0.07**	**0.07**	**0.01**	**0.01**	**0.06**

Table 3: Pearson, generalized likelihood ratio and Kolmogorov-Smirnov tests of equality of distributions. P-values for the full sample with 1247 observations.

	10 quantiles		5 quantiles	
Method	Pearson	LR	Pearson	LR
Lognormal	2 of 13	3 of 13	0 of 13	0 of 13
$G = 1$	2 of 13	2 of 13	0 of 13	0 of 13
$G = 2$	2 of 13	2 of 13	0 of 13	0 of 13
$G = 3$	2 of 13	2 of 13	0 of 13	0 of 13
$G = 4$	2 of 13	2 of 13	0 of 13	0 of 13
$G = 5$	2 of 13	2 of 13	0 of 13	0 of 13
$G = 6$	2 of 13	2 of 13	0 of 13	0 of 13
$G = 7$	2 of 13	2 of 13	0 of 13	0 of 13
$G = 8$	2 of 13	2 of 13	0 of 13	0 of 13
$G = 9$	2 of 13	2 of 13	0 of 13	0 of 13
$G = 10$	2 of 13	2 of 13	0 of 13	0 of 13
$G = 11$	2 of 13	2 of 13	0 of 13	0 of 13

Table 4: Pearson's and generalized likelihood ratio tests of equality of distributions for non-overlapping subsamples

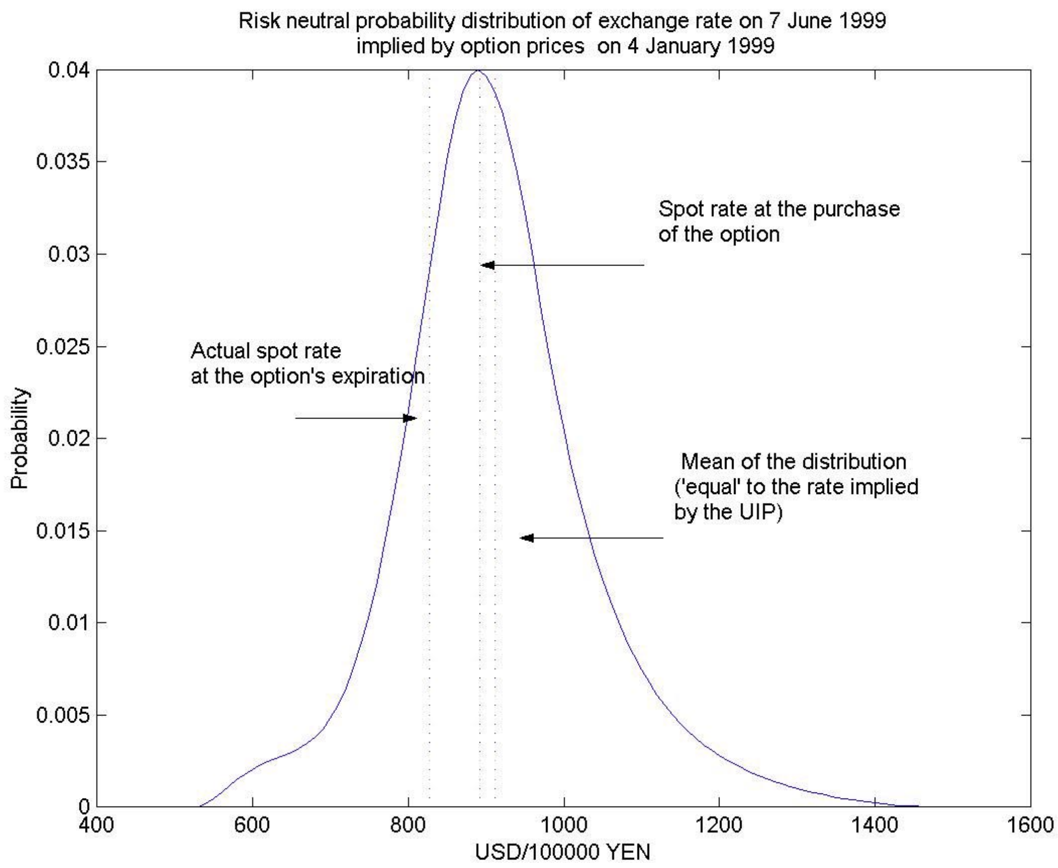

Figure 1: Example: The UIP and implied distribution

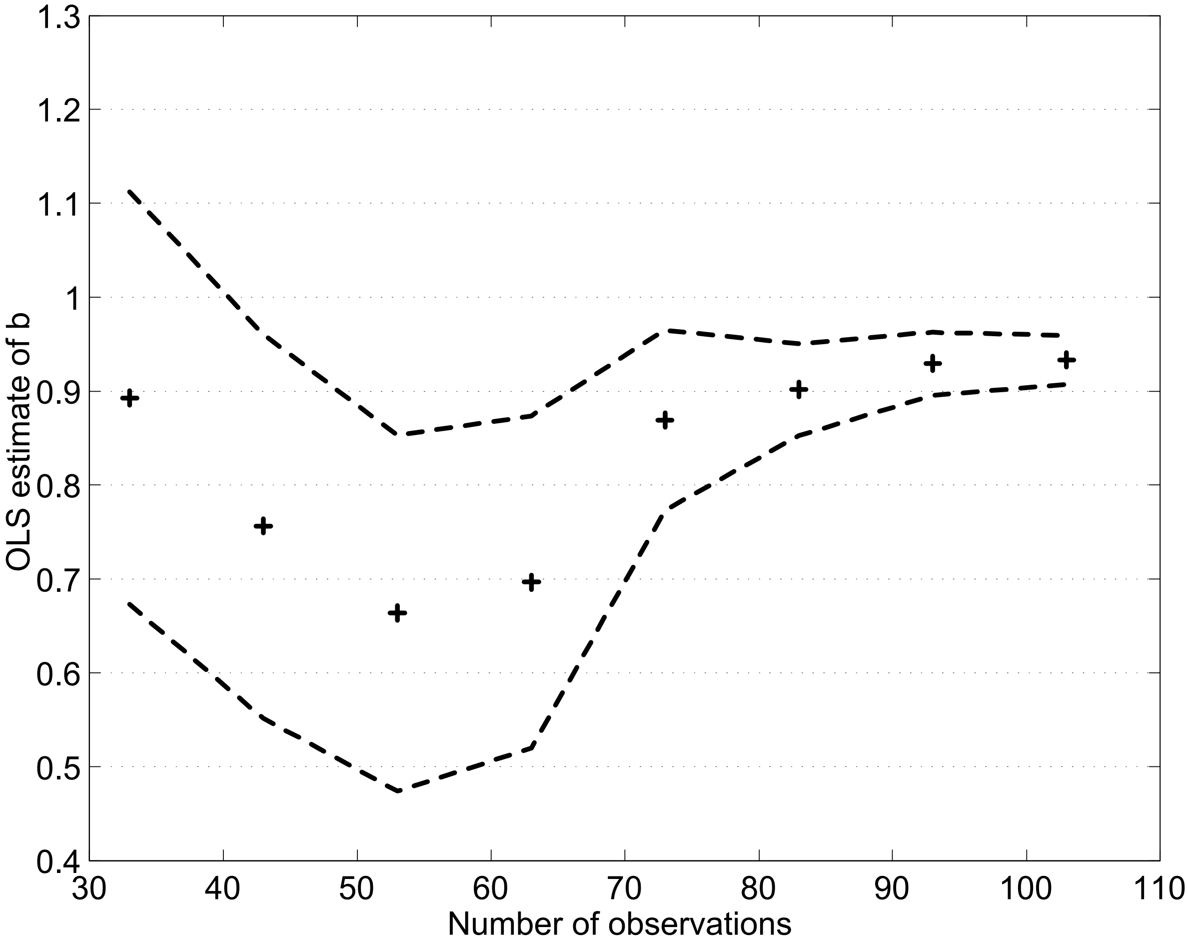

Figure 2: Estimates of b in Monte Carlo

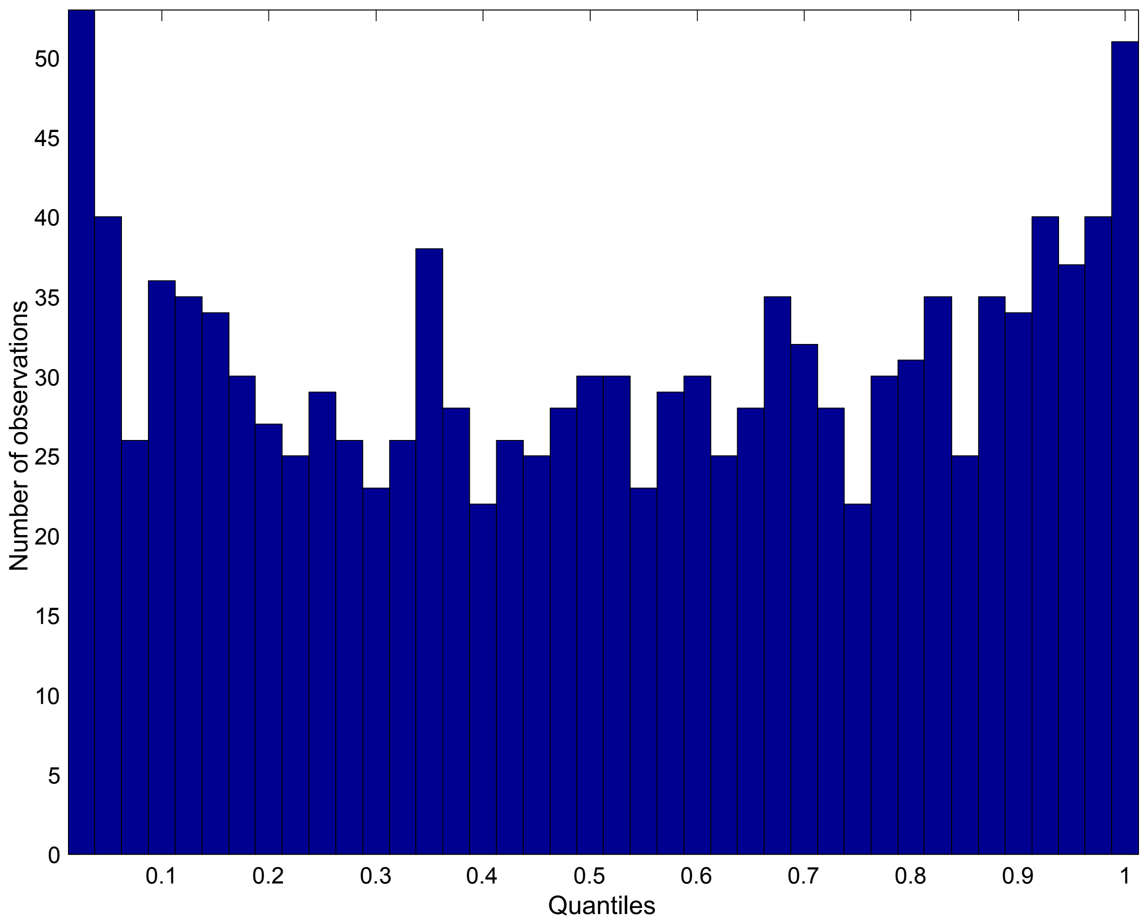

Figure 3: Full sample, method: logn, divison in 40 quantiles

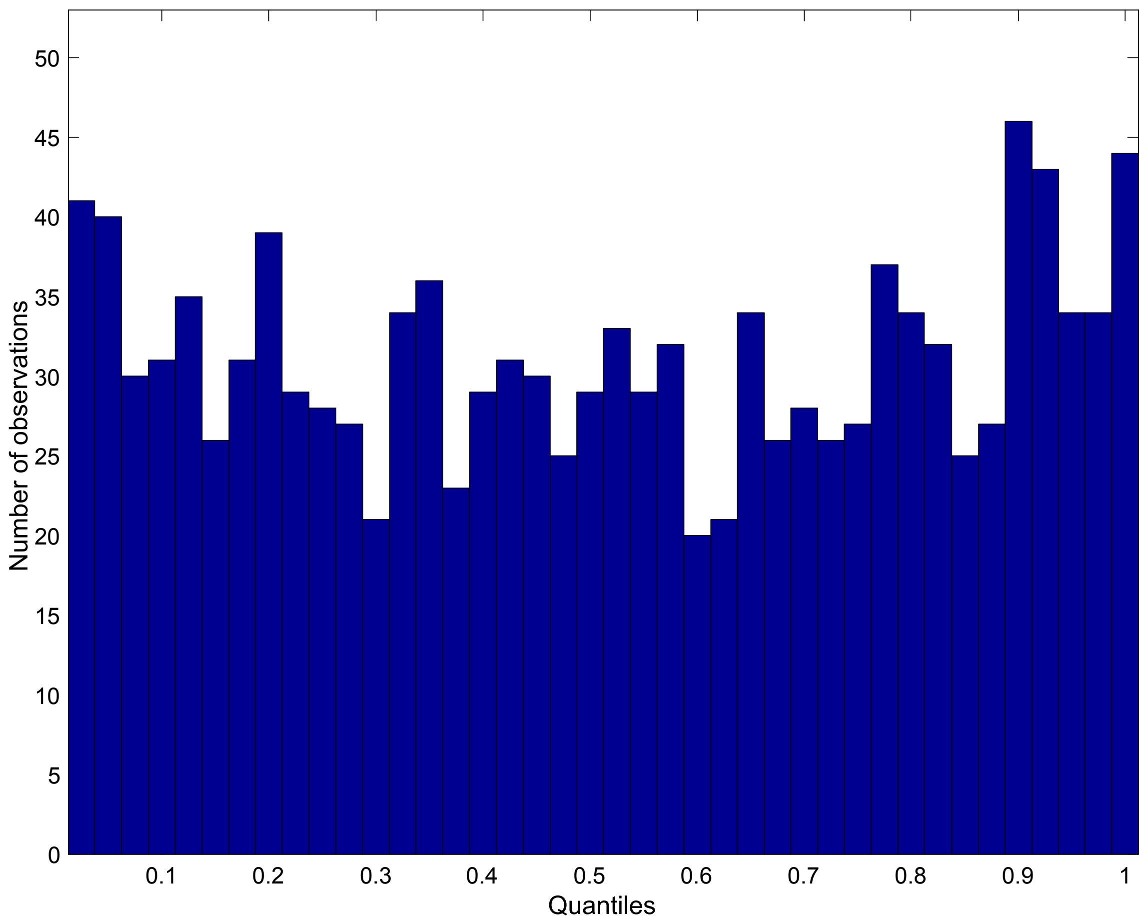

Figure 4: Full sample, method: 1, divison in 40 quantiles

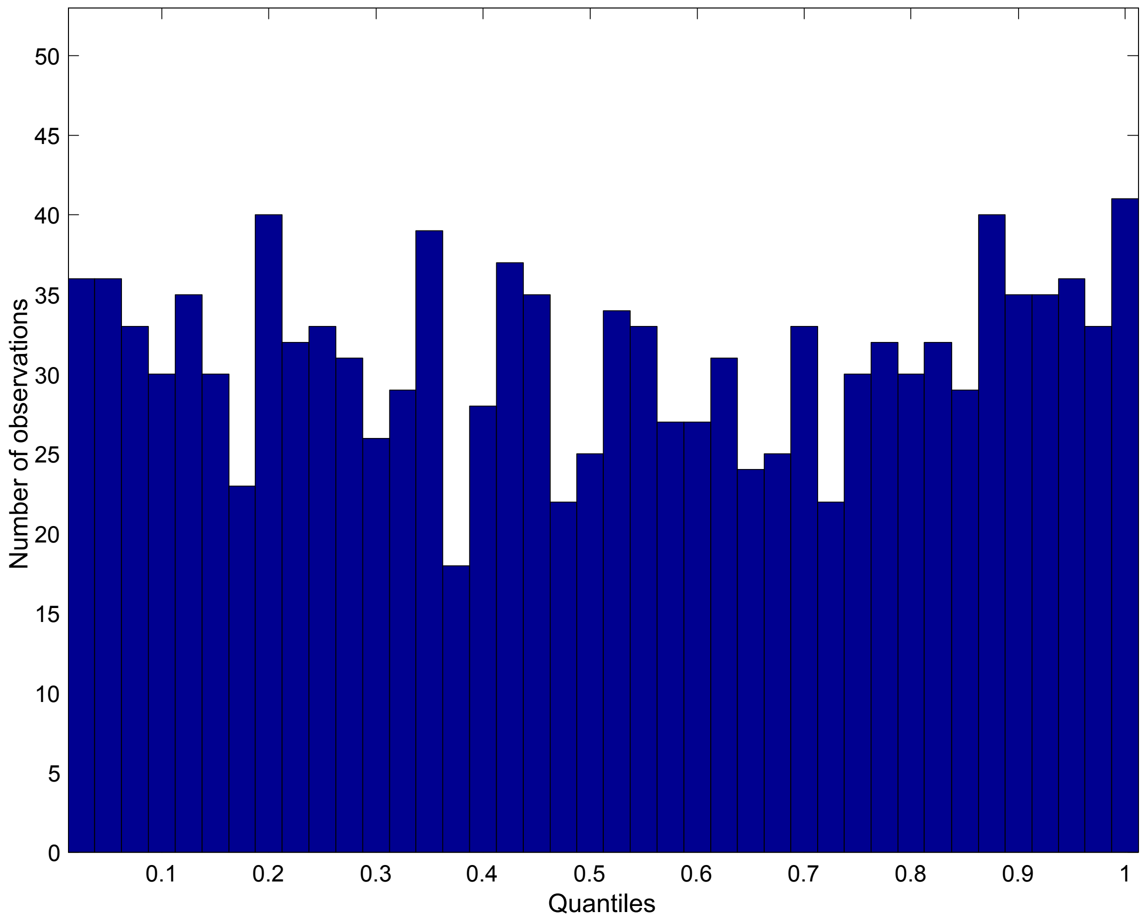

Figure 5: Full sample, method: 2, divison into 40 quantiles