

# Description of the Dataset FXBAR1

## with Special Focus on EURUSD\*

This note provides a short description of the dataset FXBAR1, kindly provided for research by Danske Bank. The data have been anonymized before being released from the bank.

### **FXBAR1**

The data consists of daily information about barrier options, both single and double, written on a number of FX-rates. The observation period is January 2 2004 to September 27 2005. The information have been obtained from a risk-management department, who in turn obtained them from an FX-desk explicitly for internal risk-management purposes.

The supplied dataset consists of the fields listed in table 1.

The primary and secondary currencies along with the number of unique contracts and the average time series length are reported in table 2. There are 8 primary and 11 secondary currencies in the dataset. The total number of unique contracts is 1,744 and the total number of observations is 80,424.

---

\*File: papers/static/3/data.description.tex, version November 24, 2005.

The dataset is decomposed by contract-type in table 3.

Note the following:

1. The time to expiry can be extracted as

$$T - t = \frac{\log\left(\frac{F}{S}\right)}{r_{dom} - r_{for}}$$

There is a discrepancy between this time and that calculated by various daycount conventions. Probably a synchronicity problem, possibly between yields and forwards, but certainly because yields and forward quotes are obtained at non-integer times to maturity when measured in days.

2. The daycount-convention is not entirely clear, but actual/365 or actual/actual seem to work. It could differ between forwards and barriers.
3. The barrier option prices are for corresponding sets of dual contracts such as, see e.g. Kholodnyi & Price (1998),

$$\text{CDO}_d(\tau, B, S, K) \rightsquigarrow \text{PUO}_f(\tau, 1/B, 1/S, 1/K)$$

That is, for a down-and-out call denominated in the domestic currency, the price of the corresponding up-and-out put denominated in the foreign currency is reported.

4. The option type, i.e. put or call, refers to the case where the secondary currency is the domestic currency. This corresponds to the currency in which spot, forward, strike and barrier are quoted.
5. There are no observations on weekends.
6. Not all contracts are part of the risk report from the day they are traded. That is, the first RUNDATE at which a contract enters may exceed the TRADEDATE.

7. Some options are triggered although both the spot the day before and on the trigger date is inside the alive region. This could indicate continuous monitoring of the barriers.
8. By using the supplied input, in particular the supplied spot volatility, with the Black-Scholes formulas and comparing with the supplied prices, we can make sanity checks of the data. This ensures a certain degree of internal consistency in the data used.

## **FXBAR1.EURUSD**

In the following, attention is focused on the EUR/USD exchange rate. The reasons for choosing to focus on this exchange rate are that it

1. is a major exchange rate,
2. has the largest number of contracts in the dataset on major currencies,
3. has a large number of contracts in each contract category.

When the underlying of an out-option crosses a barrier, the option is removed from the dataset. On the contrary, when the underlying of an in-option crosses a barrier, the (now vanilla) option stays in the dataset until expiry.

In all, 13 in-contracts are triggered in the EUR/USD dataset. All of them are single barrier options.

One contract (a 2 week single barrier call) knocked-in between being traded and entering the risk report. The option finished OTM and has been removed from the dataset.

Since no double-in-options were triggered in the observation period, the field `VOB_BAR_UPP_TCH_DT` is left out of the matlab dataset as are the fields `PRODUCT`, `MCURRENCY` and `BCURRENCY`.

All dates have been converted to serial numbers using the matlab convention, indexing Jan 1 0000 by 1. The OPTTYPE1 field has been converted according to BARR=1, DBARR=2. The OPTTYPE2 field has been converted according to C=1, P=2. The TRADENUM field has been converted to integers running from 1. All fields are now numeric which allows for easier manipulations.

The specific type (up/down/in/out) of contract was not supplied. To classify the contracts, the Black/Scholes prices corresponding to the given spot, spot volatility, yields and contract specifications were calculated for each of the possible contract types. The calculations assume zero rebates and continuously monitored barriers. Then the contracts were classified as the contract type which gave the minimum absolute deviation of the Black/Scholes price from the supplied price. Existence of deviations different from 0 could indicate data problems such as asynchronicity of observations. Furthermore, since the data are from a risk management department, it is likely that different sources have supplied the different fields, thus leading to inconsistencies in the data.

The data has further been cleaned in the following ways

- All records with nonsensical entries in any field have been removed.
- All records where the relative absolute error between the observed price and the price calculated with the given spot volatility exceeds 1% have been removed.
- All records where different option type classifications were inferred on different dates have been removed.

The break-down of the data by contract type is reported in table 4.

Of the total 214 contracts, 192 have time series starting at the day of trade.

The options that knock-in have the following TRADENUM identifiers: 30, 31, 34, 35, 131, 132, 133, 182, 196, 197, 198, 199. The 12 options have a total of 545 post knock-in observations.

The matlab dataset consists of the 23 fields listed in table 5.

## References

Kholodnyi, V. & Price, J. (1998), *Foundations of Foreign Exchange Option Symmetry*, IES Press.

Field	Description
RUNDATE	Observation time, daily
TRADENUM	Identification of each contract
PRODUCT	Noninformative (All equals "BAR")
OPTTYPE1	Single/Double barrier (BARR/DBARR)
TRADEDATE	Initiation time
EXPIRY	Expiry
DELDATE	Noninformative
OPTTYPE2	Put/call (P/C)
MCURRENCY	Primary currency
BCURRENCY	Secondary currency
VOB.BAR.LOW.KURS	Barrier for singles, lower barrier for doubles
VOB.BAR.UPPER.KURS	Upper barrier for doubles
VOB.BAR.LOW.TCH.DT	Passage time for barrier (singles) / lower barrier (doubles)
VOB.BAR.UPPER.TCH.DT	Passage time for upper barrier (doubles)
STRIKEPRICE	Strike
SPOTPRICE	Spot
SPOTVOL	B/S volatility corresponding to prices
FORWPRICE	Forward
MYIELD	Yield in primary currency
BYIELD	Yield in secondary currency
Main_prm	Barrier option price when primary currency is domestic
Bi_prm	Barrier option price when secondary currency is domestic

Table 1: Description of fields in the dataset FXBAR1

PRI / SEC	AUD	CAD	CHF	DKK	GBP	JPY	NOK	PLN	SEK	SGD	USD	Total
AUD	n.a.	0/0	0/0	0/0	0/0	0/0	20/51.80	0/0	0/0	0/0	4/17.00	24
CHF	0/0	0/0	n.a.	240/47.74	0/0	4/20.00	2/30.00	0/0	17/61.41	0/0	0/0	263
EUR	1/44.00	2/4.50	20/32.65	1/16.00	69/61.61	50/16.62	98/16.68	12/27.00	290/35.69	0/0	223/31.83	766
GBP	0/0	1/15.00	1/9.00	65/41.77	n.a.	1/6.00	1/10.00	0/0	8/13.50	0/0	42/49.45	119
JPY	0/0	0/0	0/0	8/30.75	0/0	n.a.	3/10.00	0/0	0/0	0/0	0/0	11
NOK	0/0	0/0	0/0	17/31.65	0/0	0/0	n.a.	0/0	0/0	0/0	0/0	17
SEK	0/0	0/0	0/0	15/27.73	0/0	0/0	2/63.00	0/0	n.a.	0/0	0/0	17
USD	0/0	9/31.56	6/50.33	361/77.40	0/0	47/29.38	23/68.00	2/18.00	77/46.48	2/42.50	n.a.	527
Total	1	12	27	707	69	102	149	14	392	2	269	1,744

Table 2: Number of contracts / average time series length; by primary and secondary currency.

PRI / SEC	AUD	CAD	CHF	DKK	GBP	JPY	NOK	PLN	SEK	SGD	USD	Total
AUD	n.a.	0/0/0/0	0/0/0/0	0/0/0/0	0/0/0/0	0/0/0/0	20/0/0/0	0/0/0/0	0/0/0/0	0/0/0/0	4/0/0/0	24/0/0/0
CHF	0/0/0/0	0/0/0/0	n.a.	49/95/27/69	0/0/0/0	2/2/0/0	2/0/0/0	0/0/0/0	0/17/0/0	0/0/0/0	0/0/0/0	53/114/27/69
EUR	1/0/0/0	2/0/0/0	5/12/3/0	1/0/0/0	56/12/1/0	19/30/1/0	52/28/1/17	12/0/0/0	110/98/3/79	0/0/0/0	108/71/25/19	365/252/34/115
GBP	0/0/0/0	0/0/1/0	0/0/0/1	60/5/0/0	n.a.	0/1/0/0	1/0/0/0	0/0/0/0	0/8/0/0	0/0/0/0	9/33/0/0	70/47/1/1
JPY	0/0/0/0	0/0/0/0	0/0/0/0	0/8/0/0	0/0/0/0	n.a.	3/0/0/0	0/0/0/0	0/0/0/0	0/0/0/0	0/0/0/0	3/8/0/0/0
NOK	0/0/0/0	0/0/0/0	0/0/0/0	17/0/0/0	0/0/0/0	0/0/0/0	n.a.	0/0/0/0	0/0/0/0	0/0/0/0	0/0/0/0	17/0/0/0
SEK	0/0/0/0	0/0/0/0	0/0/0/0	4/11/0/0	0/0/0/0	0/0/0/0	0/2/0/0	0/0/0/0	n.a.	0/0/0/0	0/0/0/0	4/13/0/0
USD	0/0/0/0	2/7/0/0	2/4/0/0	244/109/0/8	0/0/0/0	8/33/2/4	19/4	0/2/0/0	22/55/0/0	0/2/0/0	n.a.	295/216/2/12
Total	1/0/0/0	4/7/1/0	7/16/3/1	374/229/27/77	56/12/1/0	29/66/3/4	97/34/1/17	12/2/0/0	132/178/3/79	0/2/0/0	121/104/25/19	833/650/64/197

∞

Table 3: Number of contracts (single call/single put/double call/double put); by primary and secondary currency.

Type Number	Type	#Contracts	#Observations
Single calls		103	3384
1	CUO	44	745
2	CUI	18	655
3	CDO	36	1901
4	CDI	5	83
Single puts		69	1882
5	PUO	5	179
6	PUI	0	0
7	PDO	20	181
8	PDI	44	1522
Double calls		23	556
9	DCO	23	556
10	DCI	0	0
Double puts		19	585
11	DPO	19	585
12	DPI	0	0
Total		214	6407

Table 4: Number of contracts and observations by contract type for cleaned EUR/USD data

Field	Description
RUNDATE	Observation time, daily. Matlab serial date
TRADENUM	Identification of each contract. Integer in [1;223]
OPTTYPE1	Single/Double barrier (1/2).
TRADEDATE	Initiation time. Matlab serial date
EXPIRY	Expiry. Matlab serial date
DELDATE	Noninformative. Matlab serial date
OPTTYPE2	Call/put (1/2)
VOB_BAR_LOW_KURS	Barrier for singles, lower barrier for doubles
VOB_BAR_UPPER_KURS	Upper barrier for doubles
VOB_BAR_LOW_TCH_DT	Passage time for barrier (singles) / lower barrier (doubles)
STRIKEPRICE	Strike
SPOTPRICE	Spot
SPOTVOL	B/S volatility corresponding to prices
FORWPRICE	Forward
MYIELD	Yield in primary currency
BYIELD	Yield in secondary currency
Main_prm	Barrier option price when primary currency is domestic
Bi_prm	Barrier option price when secondary currency is domestic
T2E	Time to expiry in days. Matlab serial date
OPTTYPE3	Option type: [(1,CUO);(2,CUI);(3,CDO);(4,CDI); (5,PUO);(6,PUI);(7,PDO);(8,PDI); (9,DCO);(10,DCI);(11,DPO);(12,DPI)]
RELERR	Relative absolute difference between Bi_prm and Black-Scholes price with other fields as input
ABSERR	Absolute difference between Bi_prm and Black-Scholes price with other fields as input
BSPRICE	Black-Scholes <sup>10</sup> price with other fields as input

Table 5: Description of fields in the matrix 'datacleaned' in the matlab dataset euruscleaned.mat