10.6 Constructing continuous price and returns series

The problem: many important commodities and financial instruments trade on organized *futures markets*, where buyer and seller agree to exchange the commodity at a future date by means of a *futures contract*. Futures contract prices are quoted just as stock prices are quoted, but in the case of futures contracts, the price quotes refer to an instrument with a specific expiration date. Contracts expire on a fixed schedule: for the most heavily traded commodities, monthly. Many market participants never actually take delivery of the underlying commodity—such as crude oil—but rather maintain a continuous position in crude oil futures, "rolling over" that position when the nearest-to-maturity contract approaches its expiration date. This complicates analysis of data derived from futures prices, such as rates of return from maintaining a futures position. We would like to produce a continuous price series and returns series from the set of futures contract price quotations available, spanning the various contract maturity dates.

Industry analysts have noted that to avoid market disruptions, the large participants in the crude oil futures market roll over their positions from the near contract to the next-near contract over several days before the near contract's expiration date. We can thus define a method of producing a continuous price series from a set of price quotations on multiple months' contracts. Consider the last five trading days before expiration, and consider the midpoint of that interval the *rollover date*. Prior to the rollover date, the price is taken from the near contract: that which will imminently expire. On and after the rollover date, the price is taken from the next-near contract: the contract expiring next month.

For all but the rollover date, the return is the *log price relative*: that is, the logarithm of the ratio of price today to price yesterday, or $\log(P_t/P_{t-1})$. For the rollover date, we assume that the near-contract position is liquidated at the previous trading day's settlement price, and instantaneously rolled into the next-near contract. Thus the return on the rollover date is the log price relative of the next-near contract for the rollover day and the previous trading day.

To illustrate, we display an excerpt of the crude oil futures contract quotations.

```
. list in 1/51, sepby(qmdy)
```

	qmdy	cmdy	settle	contract	qm	cm
1.	02jan1985	01feb1985	25.92	CL1985G	300	301
2.	02jan1985	01mar1985	25.81	CL1985H	300	302
3.	02jan1985	01apr1985	25.69	CL1985J	300	303
4.	02jan1985	01may1985	25.63	CL1985K	300	304
5.	02jan1985	01jun1985	25.60	CL1985M	300	305
6.	02jan1985	01jul1985	25.59	CL1985N	300	306
7.	02jan1985	01aug1985	25.57	CL1985Q	300	307
8.	02jan1985	01sep1985	25.57	CL1985U	300	308
9.	02jan1985	01oct1985	25.57	CL1985V	300	309
10.	02jan1985	01nov1985	25.57	CL1985X	300	310
11.	02jan1985	01dec1985	25.57	CL1985Z	300	311

12.	02jan1985	01jan1986	25.57	CL1986F	300	312
13.	02jan1985	01feb1986	25.57	CL1986G	300	313
14.	02jan1985	01mar1986	25.57	CL1986H	300	314
15.	02jan1985	01apr1986	25.57	CL1986J	300	315
16.	02jan1985	01may1986	25.57	CL1986K	300	316
17.	02jan1985	01jun1986	25.57	CL1986M	300	317
18.	03jan1985	01feb1985	25.84	CL1985G	300	301
19.	03jan1985	01mar1985	25.79	CL1985H	300	302
20.	03jan1985	01apr1985	25.68	CL1985J	300	303
21.	03jan1985	01may1985	25.65	CL1985K	300	304
22.	03jan1985	01jun1985	25.59	CL1985M	300	305
23.	03jan1985	01jul1985	25.58	CL1985N	300	306
24.	03jan1985	01aug1985	25.56	CL1985Q	300	307
25.	03jan1985	01sep1985	25.56	CL1985U	300	308
26.	03jan1985	01oct1985	25.56	CL1985V	300	309
27.	03jan1985	01nov1985	25.56	CL1985X	300	310
28.	03jan1985	01dec1985	25.56	CL1985Z	300	311
29.	03jan1985	01jan1986	25.56	CL1986F	300	312
30.	03jan1985	01feb1986	25.56	CL1986G	300	313
31.	03jan1985	01mar1986	25.56	CL1986H	300	314
32.	03jan1985	01apr1986	25.56	CL1986J	300	315
33.	03jan1985	01may1986	25.56	CL1986K	300	316
34.	03jan1985	01jun1986	25.56	CL1986M	300	317
35.	04jan1985	01feb1985	25.18	CL1985G	300	301
36.	04jan1985	01mar1985	25.19	CL1985H	300	302
37.	04jan1985	01apr1985	25.16	CL1985J	300	303
38.	04jan1985	01may1985	25.13	CL1985K	300	304
39.	04jan1985	01jun1985	25.10	CL1985M	300	305
40.	04jan1985	01jul1985	24.90	CL1985N	300	306
41.	04jan1985	01aug1985	25.06	CL1985Q	300	307
42.	04jan1985	01sep1985	25.06	CL1985U	300	308
43.	04jan1985	01oct1985	25.06	CL1985V	300	309
44.	04jan1985	01nov1985	25.06	CL1985X	300	310
45.	04jan1985	01dec1985	25.06	CL1985Z	300	311
46.	04jan1985	01jan1986	25.06	CL1986F	300	312
47.	04jan1985	01feb1986	25.06	CL1986G	300	313
48.	04jan1985	01mar1986	25.06	CL1986H	300	314
49.	04jan1985	01apr1986	25.06	CL1986J	300	315
50.	04jan1985	01may1986	25.06	CL1986K	300	316
51.	04jan1985	01jun1986	25.06	CL1986M	300	317

In this listing of three days' quotations, `qm` is the month number of the quote date and `cm` is the month number in which the contract expires. Recall that Stata's dates start from 1 January 1960, so that month 300 is January 1985 and month 301 is February 1985. `qmdy` is the quote date and `cmdy` is the maturity month of the contract. That is, contract `CL1985G` is the "Feb 85" contract which last traded on 18 January 1985, as contracts expire in roughly the third week of the previous calendar month. That contract was quoted at $25.92 (per barrel of crude oil) on 2 January 1985 and $25.84 on 3 January 1985.

We first must define the near contract (that closest to expiration) for each month:

```
. // identify last day of trading for near contract in each month
```

```
. bysort qmdy (cmdy): generate near = contract if _n == 1
(155907 missing values generated)
```

We now can identify the rollover date and the prior trading date:[11]

```
. // qmdy is the first date when the near contract is no longer quoted
. // that minus 3 trading days is the target rollover date
. bysort cmdy (qmdy): generate rolldate = qmdy[_n-3] if near[_n] != near[_n-1]
(161354 missing values generated)
. bysort cmdy (qmdy): generate roll1date = qmdy[_n-4] if near[_n] != near[_n-1]
>
(161354 missing values generated)
. bysort cmdy (qmdy): generate nnear = contract if near[_n] != near[_n-1]
(161353 missing values generated)
. // fixup for first obs
. replace nnear = . in 1
(1 real change made, 1 to missing)
```

We use the `egen mean()` function to set the `rolldate`, `roll1date` and `nnear` values into each trading day of the quote month:

```
. bysort qm: egen rollover = mean(rolldate)
. bysort qm: egen rollover1 = mean(roll1date)
. bysort qm: egen nextnear = mean(nnear)
```

With these variables defined, we are ready to calculate the continuous price and returns series:

```
. // calculate price series as settle(near) for pre-rollover dates
. bysort qm: generate futprice = settle if contract == near & qmdy < rollover
(158572 missing values generated)
. // calculate price series as settle(nnear) for rollover date et seq.
. bysort qm: replace futprice = settle if contract==nextnear & qmdy >= rollover
(2665 real changes made)
. format futprice %9.2f
. // calculate return series for the rollover date
. bysort qm: generate settleprev = settle if contract == nextnear & qmdy == rol
> lover1
(161356 missing values generated)
. bysort qm: egen sprev = mean(settleprev)
. bysort qm: generate double futret = log(settle) - log(sprev) if qmdy == rollo
> ver
(153860 missing values generated)
. // drop obs no longer needed, flagged by missing settle
. drop if futprice == .
(155907 observations deleted)
. // calc returns for all non-settle dates
. sort qmdy
```

11. In the code below, note that references to `near[_n]` could be replaced by `near`. We use the explicit subscript to clarify the meaning of the code.

```
. replace futret = log(settle) - log(settle[_n-1]) if missing(futret)
(5449 real changes made)
```

We now verify that the proper series have been constructed:

```
. sort qmdy cmdy
. list qmdy contract futprice futret if !missing(futprice) in 1/62,  noobs sepb
> y(qm)
```

qmdy	contract	futprice	futret
02jan1985	CL1985G	25.92	.
03jan1985	CL1985G	25.84	-.00309119
04jan1985	CL1985G	25.18	-.02587364
07jan1985	CL1985G	25.56	.01497857
08jan1985	CL1985G	25.48	-.0031348
09jan1985	CL1985G	25.43	-.00196422
10jan1985	CL1985G	25.76	.01289332
11jan1985	CL1985G	25.77	.00038813
14jan1985	CL1985G	26.12	.01349029
15jan1985	CL1985G	25.91	-.00807235
16jan1985	CL1985H	25.57	-.01243699
17jan1985	CL1985H	25.69	.00468205
18jan1985	CL1985H	25.75	.0023328
21jan1985	CL1985H	25.97	.00850737
22jan1985	CL1985H	25.55	-.01630471
23jan1985	CL1985H	25.40	-.00588813
24jan1985	CL1985H	25.28	-.00473556
25jan1985	CL1985H	25.25	-.00118744
28jan1985	CL1985H	25.23	-.00079241
29jan1985	CL1985H	25.38	.00592768
30jan1985	CL1985H	25.67	.01136157
31jan1985	CL1985H	26.41	.02841972
01feb1985	CL1985H	26.74	.01241784
04feb1985	CL1985H	26.52	-.00826138
05feb1985	CL1985H	26.78	.00975618
06feb1985	CL1985H	27.07	.01077073
07feb1985	CL1985H	27.21	.00515843
08feb1985	CL1985H	27.59	.01386887
11feb1985	CL1985H	28.04	.0161787
12feb1985	CL1985H	27.36	-.02454998
13feb1985	CL1985J	27.06	.0186503
14feb1985	CL1985J	27.04	-.00073932
15feb1985	CL1985J	27.38	.0124955
19feb1985	CL1985J	27.29	-.00329242
20feb1985	CL1985J	27.18	-.00403895
21feb1985	CL1985J	27.14	-.00147279
22feb1985	CL1985J	26.76	-.01410039
25feb1985	CL1985J	26.44	-.01203021
26feb1985	CL1985J	26.79	.01315068
27feb1985	CL1985J	26.69	-.00373973
28feb1985	CL1985J	26.73	.00149753
01mar1985	CL1985J	27.20	.01743049
04mar1985	CL1985J	27.74	.01965841
05mar1985	CL1985J	27.55	-.0068729
06mar1985	CL1985J	27.77	.00795381

```
07mar1985    CL1985J      28.08    .01110126
08mar1985    CL1985J      27.74   -.01218217
11mar1985    CL1985J      27.57   -.00614719
12mar1985    CL1985J      27.92    .01261507
13mar1985    CL1985J      28.06    .00500178
14mar1985    CL1985J      28.19    .00462227
15mar1985    CL1985J      28.32    .00460093
18mar1985    CL1985K      28.25    .02182948
19mar1985    CL1985K      28.19   -.00212613
20mar1985    CL1985K      27.99   -.00712003
21mar1985    CL1985K      28.32    .01172096
22mar1985    CL1985K      28.24   -.00282885
25mar1985    CL1985K      28.09   -.00532576
26mar1985    CL1985K      28.45    .01273454
27mar1985    CL1985K      28.16   -.01024566
28mar1985    CL1985K      28.25    .00319093
29mar1985    CL1985K      28.29    .00141496
```

You should note in this recipe how business-daily data have been handled: no use has been made of the time series operators (such as `L.`), as the trading-day data are not evenly spaced in calendar time. Also note that the multiple price quotations per trading day have been transformed into simple time series of prices and returns in the process.