

RESSET高频数据库

培训内容

群聊：对外经贸锐思数据服务群



该二维码7天内(3月29日前)有效，重新进入将更新

01

RESSET高频数据

02

高频数据提取介绍

03

高频数据概览

04

数据词典介绍

05

学术应用

PART

1

RESSET高频数据





高频系列数据库



■ 深度处理，提取方便：

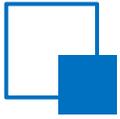
RESSET高频数据是按每个交易证券每年/每月存为一个SAS数据集，节省大量自己从交易所原始披露的每日数据集中，挨个提取、合并同一个证券数据的时间。我们也有提供csv、txt、stata、spss等不同格式批量导出的模板程序。

■ 衍生指标丰富（沪深高频）：

对买卖方向、绝对买卖报价差、相对买卖报价差、绝对有效买卖报价差、相对有效买卖报价差、深度1、深度2等市场买卖指标进行了计算。

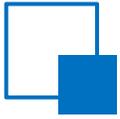
■ 频度全面：

除了分笔高频数据外，还对8种分时频度（1分钟、5分钟、10分钟、15分钟、20分钟、30分钟、40分钟、60分钟）进行了计算，大量节省需要使用分时高频数据时的处理时间。



高频系列数据库：特色与优势

- **严格的逻辑关系检测**：RESSET高频数据是对原始数据进行深度处理的高频数据，采用多种查验方法进行数据清洗，从分笔数据生成的频度齐全的分时数据，对数据各字段进行严格逻辑关系检测来确保数据的规范性和准确性。
- **按每月每个交易证券存储**：RESSET高频数据是按交易工具形式的SAS数据集进行存储，每月每个交易证券作为一个SAS数据集，数据集名含有证券类别、年份、月份、代码标识（分时间隔），用户直接按照表名规则提取该证券对应的一个数据集即可，从而大量节省从交易所原始披露的每日数据集中挨个提取、合并同一个证券的数据的时间。
- **提供大量经计算后的衍生指标**：对买卖方向、绝对买卖报价差、相对买卖报价差、绝对有效买卖报价差、相对有效买卖报价差、深度1、深度2等市场买卖指标进行了计算。
- **频度全面**：除了分笔高频数据外，还对8种分时频度（1分钟、5分钟、10分钟、15分钟、20分钟、30分钟、40分钟、60分钟）进行了计算，大量节省需要使用分时高频数据时的处理时间。
- **SAS处理数据的优势**：（1）资源消耗低；（2）运行效率更高；（3）独家支持SAS CONNECT 在线计算平台。使用RESSET高频数据无需将庞大的数据内容落地到本地客户端，可以利用SAS软件与服务器数据建立映射，远程调用直接下载所需要的数据，也可将SAS程序代码提交至服务器，再把最终的计算结果下载至本地，大大提高了数据的计算和使用效率。
- **批量转换其它数据格式**：SAS数据集可以批量导出成csv、txt、STATA、SPSS等多种数据格式，同时也可以提供不同格式批量导出的模板程序。



高频数据类型：Level I & Level II

Level I 级别数据

- 包括基础数据：证券代码、证券简称、前收盘价、最新成交价、开高低收、成交情况以及**五档行情**
- 数据间隔：5秒 -> 3秒

Level II 级别数据

- 包括证券代码、证券简称、前收盘价、最新成交价、开高低收、成交情况以及**十档行情**
- 还包括**指数数据、逐笔成交、至多50位长度的委托队列及逐笔委托数据（深交所）**
- 数据间隔：大约3秒

注：沪深交易所提供的变量并不完全一致，因此以上只是一个大概的介绍，详细的说明请参见数据字典。



命名规则

数据集命名规则

- 分笔数据集命名方式：证券类别缩写+HF+年份_证券代码+市场。
- 分时数据集命名方式：分笔数据集名称_分钟间隔，分为1分钟、5分钟、10分钟、15分钟、20分钟、30分钟、40分钟、60分钟。例如，*StkHF2018_600001SH_1*，代表上交所股票600001在2018年1月的市场五档行情表的Level1一分钟分时数据。
- 例如*StkHF2016_600000SH_5*，可以分为/ *Stk* / *HF* / *2016* / *600000* / *SH* / *5* / 其中每一部分的意义如下。

Stk	: 表示证券类别是“股票”，其他的证券类别还包括：债券 (Bond)，基金 (Fund)，指数 (Indx)，回购 (Repo)，资产支持证券 (Abs)
HF	: 表示高频数据
2016	: 表示数据时间区间为2016年整年的数据
600000	: 表示股票代码是600000
SH	: 表示交易所为上交所 (上海证券交易所)
5	: 表示分时频率为5分钟

PART **2**

数据提取示例





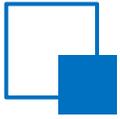
特别提醒

拷贝需谨慎：系统盘被装满可能会导致系统卡顿、不稳定，甚至崩溃、无法启动。



平台选择SAS的逻辑

	数据提取，速度慢 从服务器下载原始数据	远程执行，速度快 当你只有500G的硬盘却想处理2T的数据
其他方式	简单 要求会sql等	困难 有可能耗尽服务器资源
SAS数据集	中等 我们有提供相关工具	中等 要求会SAS



使用锐思数据提供的get_data_l1宏

只需自己写一行代码，即可实现获取指定种类、指定时间段的高频数据，并将其以指定的格式输出到指定目录中。

code	证券代码，纯数字，例如000001。
beginDate	起始时间，必须是'03Jan2023'd 这种格式，或者使用下面提供的yearmonthday宏
endDate	结束时间，必须是'03Jan2023'd 这种格式，或者使用下面提供的yearmonthday宏
exch	指定交易所，需要是SH SZ中之一，不区分大小写
secuCategory	指定证券类型，需要是STK BOND FUND INDX REPO ABS中之一，不区分大小写
select (可选)	保留那些变量，如果不填则保留全部变量
export	输出位置，路径中不能有特殊字符(以下均为英文，,.:&%' "#;~^空格等)，例如D:/out.csv
dbms	输出类型，需要是CSV EXCEL XLS XLSX STATA SPSS中之一，不区分大小写
minutes (可选)	填写需要的分时频率，需要是1 5 10 15 20 30 40 60中之一，留空表示分笔数据
encoding (可选)	输出编码，例如utf8 gbk等

■ yearmonthday宏

三个数值参数，分别是年、月、日



使用SAS代码生成器

■ SAS代码生成器

黑色方框的行，填入需要的内容

橙色方框的行，每一行只能选一个，底色改为其他颜色即可

第二张表“证券列表”中填入需要下载的股票，代码和交易所必填
完成后，点击“生成代码”

开始时间	2020/1/1	结束时间	2021/1/1						
证券类型（只能选一个）	STK	BOND	FUND	INDX	ABS	REPO			
频率（只能选一个）	逐笔	1分钟	5分钟	10分钟	15分钟	20分钟	30分钟	40分钟	60分钟
输出文件类型（只能选一个）	CSV	EXCEL	STATA	SPSS					
输出目录	D:/temp								
输出文件名前缀	stkhf								
宏文件位置	D:/temp								
生成代码					清空选择				

```

data stk;
infile cards;
informat startdt enddt yymmdd10.;
format startdt enddt yymmdd10.;
input stkcd $ exch $ startdt enddt;
cards;
600000 SH 2020-1-1 2021-1-1
;
run;
%include 'D:/temp/conn_11.sas';
%include 'D:/temp/get_data_11.sas';
filename tmp temp;
data _null_; file tmp; set stk;
cmd = compress("%get_data_11('||stkcd||', ' || startdt || ', ' || enddt || ', ' || exch || ',STK, D:/temp/stkhf || stkcd || '.csv, CSV, 5)');
put cmd; run;
%include tmp;
filename tmp clear;

```



使用Matlab、R或Python读取SAS数据集

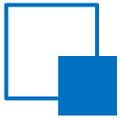
■ 使用Python读取SAS数据集

使用pyreadstat程序包

```
>>> import pyreadstat
>>> df = pyreadstat.read_sas7bdat('D:/temp/test.sas7bdat')
>>> df
(
  Exchflg  Code  Code_Mkt  Securm  QTime  ...  PE1 PE2  ChangePCT1  ChangePCT2  OpInterest
0      1.0  010303  010303.SH  03国债(3)  2022-01-04 09:29:00  ...  NaN NaN      NaN      NaN      NaN
1      1.0  010303  010303.SH  03国债(3)  2022-01-04 10:29:01  ...  NaN NaN      NaN      NaN      NaN
2      1.0  010303  010303.SH  03国债(3)  2022-01-04 11:25:00  ...  NaN NaN      NaN      NaN      NaN
3      1.0  010303  010303.SH  03国债(3)  2022-01-04 11:59:00  ...  NaN NaN      NaN      NaN      NaN
4      1.0  010303  010303.SH  03国债(3)  2022-01-04 12:59:00  ...  NaN NaN      NaN      NaN      NaN
...
1931    1.0  010303  010303.SH  03国债(3)  2022-12-30 11:59:52  ...  NaN NaN      NaN      NaN      NaN
1932    1.0  010303  010303.SH  03国债(3)  2022-12-30 12:59:52  ...  NaN NaN      NaN      NaN      NaN
1933    1.0  010303  010303.SH  03国债(3)  2022-12-30 13:59:52  ...  NaN NaN      NaN      NaN      NaN
1934    1.0  010303  010303.SH  03国债(3)  2022-12-30 14:59:52  ...  NaN NaN      NaN      NaN      NaN
1935    1.0  010303  010303.SH  03国债(3)  2022-12-30 15:29:46  ...  NaN NaN      NaN      NaN      NaN
[1936 rows x 70 columns], <pyreadstat._readstat_parser.metadata_container object at 0x000001AFF10A21A0>
>>>
```

使用sas7bdat程序包

```
>>> import sas7bdat
>>> from sas7bdat import SAS7BDAT
>>> with SAS7BDAT('D:/temp/test.sas7bdat', encoding='gbk') as reader:
...     df = reader.to_data_frame()
...
>>> df.head()
  Exchflg  Code  Code_Mkt  Securm  QTime  ...  PE1 PE2  ChangePCT1  ChangePCT2  OpInterest
0      1.0  010303  010303.SH  03国债(3)  2022-01-04 09:29:00  ...  None None      None      None      None
1      1.0  010303  010303.SH  03国债(3)  2022-01-04 10:29:01  ...  None None      None      None      None
2      1.0  010303  010303.SH  03国债(3)  2022-01-04 11:25:00  ...  None None      None      None      None
3      1.0  010303  010303.SH  03国债(3)  2022-01-04 11:59:00  ...  None None      None      None      None
4      1.0  010303  010303.SH  03国债(3)  2022-01-04 12:59:00  ...  None None      None      None      None
[5 rows x 70 columns]
>>>
```



使用Matlab、R或Python读取SAS数据集

■ 使用Python读取SAS数据集

使用pyreadstat程序包 vs 使用sas7bdat程序包

```
In [1]: import pyreadstat
        from sas7bdat import SAS7BDAT

In [2]: def read_st():
        return pyreadstat.read_sas7bdat('D:/temp/test.sas7bdat')

        def read_py():
        return SAS7BDAT('D:/temp/test.sas7bdat', encoding='gbk').to_data_frame()

In [4]: %%timeit
        read_st()

        20.4 ms ± 600 µs per loop (mean ± std. dev. of 7 runs, 10 loops each)

In [5]: %%timeit
        read_py()

        270 ms ± 10.5 ms per loop (mean ± std. dev. of 7 runs, 1 loop each)
```



使用Matlab、R或Python读取SAS数据集

■ 使用R读取SAS数据集

使用haven程序包

```
> library(haven)
> read_sas('D:/temp/test.sas7bdat')
# A tibble: 1,936 × 70
  Exchflg Code Code_Mkt Secunm QTime Qdate Stdtime Highpr Lowpr Begpr TVolume_accu TSum_accu
  <dbl> <chr> <chr> <chr> <dtm> <date> <time> <dbl> <dbl> <dbl> <dbl> <dbl>
1 1 010303 010303... 03国... 2022-01-04 09:29:00 2022-01-04 09:30 0 0 0 0 0
2 1 010303 010303... 03国... 2022-01-04 10:29:01 2022-01-04 10:30 101. 0 101. 18028 18274829
3 1 010303 010303... 03国... 2022-01-04 11:25:00 2022-01-04 11:30 101. 101. 101. 332 336583
4 1 010303 010303... 03国... 2022-01-04 11:59:00 2022-01-04 12:00 101. 101. 101. 2 2028
5 1 010303 010303... 03国... 2022-01-04 12:59:00 2022-01-04 13:00 101. 101. 101. 0 0
6 1 010303 010303... 03国... 2022-01-04 13:59:00 2022-01-04 14:00 101. 101. 101. 2791 2829256
7 1 010303 010303... 03国... 2022-01-04 14:59:00 2022-01-04 15:00 101. 101. 101. 1046 1060336
8 1 010303 010303... 03国... 2022-01-04 15:29:48 2022-01-04 15:30 101. 101. 101. 0 0
9 1 010303 010303... 03国... 2022-01-05 09:29:00 2022-01-05 09:30 0 0 0 0 0
10 1 010303 010303... 03国... 2022-01-05 10:29:53 2022-01-05 10:30 101. 0 101. 3921 3974086
```



使用Matlab、R或Python读取SAS数据集

■ 使用Matlab读取SAS数据集

文件列表:

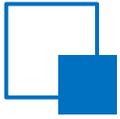
ReadStat.mexw64

read_sas.m

```
编辑器 - D:\Downloads\ReadStat\VS17\x64\Debug\read_sas.m
read_sas.m x +
1 function out_table = read_sas(path)
2 % 读取指定路径的SAS数据集文件。返回table类型的结果，表头为变量的名称，表格数据中日期和日期时间的数据类型统一采用datetime结构，时间数据类型采用duration:
3 % 最低支持: Matlab R2015b
4 % path: SAS数据集文件路径，长度不要超过1024个字符
5
6 % 返回类型: table。
7 [header, content, vartypes] = ReadStat(path);
8 [~, n] = size(content);
9 out_table = cell2table(content, 'VariableNames', header);
10
11 for i = 1: n
12     if strcmp(vartypes{1, i}, 'DATETIME') > 0
13         out_table(header{1, i}) = datetime(out_table(header{1, i}), 'ConvertFrom', 'epochtime', 'Epoch', '1960-01-01');
14     end
15     if strcmp(vartypes{1, i}, 'DATE') > 0
16         out_table(header{1, i}) = datetime(out_table(header{1, i}), 'ConvertFrom', 'datenum');
17
命令窗口
>> df = read_sas('D:/zy/programs/matlab_2015b/bin/test.sas7bdat');
>> df(1:5, :)

ans =

    Exchflg      Code      Code_Mkt      Securmn      QTime      Qdate      Stdtime      Highpr      Lowpr      Begpr      TVolume_a
    -----      -
    1          '010303'      '010303.SH'      '03国债(3)'      2022-01-04 09:29:00      2022-01-04 00:00:00      09:30:00      0          0          0          0
```



使用Python进行数据处理的一般性提示

1、谨慎使用Groupby。 由于迭代器每次只返回chunksize条观测，数据并不完全，直接使用groupby可能会导致结果不准确、变量缺失、重复观测等问题。

2、谨慎使用Pandas包。 虽然使用Pandas包会使数据分析变得方便很多，但是，Pandas包的速度较慢，并且内存使用较大。当数据集较小时，并不会遇到太大的问题。但是当数据集逐渐增大到接近甚至超过物理内存时，这些问题的不良影响将会被放大，甚至导致Out of Memory。

3、尝试使用numba或cython等工具进行加速。 Python本身的执行速度较慢，尤其是针对这种计算密集型的任务，所以需要设法引入速度更快的其他编程语言。

4、尽可能使用经过验证的第三方工具包。 Python的数据处理工具包虽然丰富，但是处理速度都谈不上有多快，但是近年来新出现了一些新的工具，速度更快，并且内存占用更低。我的建议是大家在正式使用之前先测试一下，以确保结果的正确性。如果结果不准确，无论处理速度再快也是不能用的。

5、考虑使用其他可用的工具。 例如Spark还有Python binding，可能在Python中进行调用。或者还可以将数据导入到sql数据库，然后使用sql查询语言来进行groupby等复杂的操作，不会出现内存或结果错误等问题，但是速度上就会慢一些了。

PART

3

数据概览





数据概览

■ 9:15 - 9:25

开盘集合竞价时间。根据买卖双方挂单，预先撮合，**不成交**，可申报。**由于此时并无真实的成交，所以仅前收盘更新，开高低都无数据。**

■ 9:25

集合竞价挂单统一成交，成交价作为开盘价。集合竞价阶段未成交的买卖申报，自动进入当日后续竞价交易阶段。

Security Name	Quotation Time	Quotation Date	Previous Close Price	Open Price	High Price	Low Price	CLPr	Trade Price	TVolume
平安银行	04JAN2022:09:25:00.00	2022-01-04	16.4800	16.4800	16.4800	16.4800	.	16.4800	1057909
平安银行	04JAN2022:09:26:00.00	2022-01-04	16.4800	16.4800	16.4800	16.4800	.	16.4800	0
平安银行	04JAN2022:09:27:00.00	2022-01-04	16.4800	16.4800	16.4800	16.4800	.	16.4800	0
平安银行	04JAN2022:09:28:00.00	2022-01-04	16.4800	16.4800	16.4800	16.4800	.	16.4800	0
Security Name	Modify Time	Quotation Time	Quotation Date	Standard Time	Highest Price in 5 Min	Lowest Price in 5 Min	Begin Price in 5 Min		
平安银行	03JAN2017:09:19:38	03JAN2017:09:19:39	2017/1/3	9:20					
平安银行	03JAN2017:09:24:56	03JAN2017:09:24:57	2017/1/3	9:25					
平安银行	03JAN2017:09:25:05	03JAN2017:09:25:06	2017/1/3	9:30	9.11	9.11	9.11	9.11	
平安银行	03JAN2017:09:34:57	03JAN2017:09:34:57	2017/1/3	9:35	9.11	9.09	9.11	9.11	
平安银行	03JAN2017:09:40:11	03JAN2017:09:39:54	2017/1/3	9:40	9.1	9.09	9.09	9.1	
平安银行	03JAN2017:09:45:18	03JAN2017:09:44:57	2017/1/3	9:45	9.11	9.09	9.09	9.09	



数据概览

■ 9:25

集合竞价挂单统一成交，成交价作为开盘价。集合竞价阶段未成交的买卖申报，自动进入当日后续竞价交易阶段。

集合竞价时，成交价格的确定原则为：

- （一）可实现最大成交量的价格；
- （二）高于该价格的买入申报与低于该价格的卖出申报全部成交的价格；
- （三）与该价格相同的买方或卖方至少有一方全部成交的价格。

两个以上申报价格符合上述条件的，使未成交量最小的申报价格为成交价格；仍有两个以上使未成交量最小的申报价格符合上述条件的，其中间价为成交价格。

集合竞价的交易以同一价格成交。

* 集合竞价期间**不受价格笼子的影响。**



数据概览

■ 9:25-9:30

交易暂停。

Security Name	Quotation Time	Quotation Date	Previous Close Price	Open Price	High Price	Low Price	CLPr	Trade Price	TVolume
平安银行	04JAN2022:09:25:00.00	2022-01-04	16.4800	16.4800	16.4800	16.4800	.	16.4800	1057909
平安银行	04JAN2022:09:26:00.00	2022-01-04	16.4800	16.4800	16.4800	16.4800	.	16.4800	0
平安银行	04JAN2022:09:27:00.00	2022-01-04	16.4800	16.4800	16.4800	16.4800	.	16.4800	0
平安银行	04JAN2022:09:28:00.00	2022-01-04	16.4800	16.4800	16.4800	16.4800	.	16.4800	0
平安银行	04JAN2022:09:29:00.00	2022-01-04	16.4800	16.4800	16.4800	16.4800	.	16.4800	0
平安银行	04JAN2022:09:30:00.00	2022-01-04	16.4800	16.4800	16.5000	16.4800	.	16.4900	331500
平安银行	04JAN2022:09:30:03.00	2022-01-04	16.4800	16.4800	16.5000	16.4600	.	16.4900	925900
平安银行	04JAN2022:09:30:06.00	2022-01-04	16.4800	16.4800	16.5000	16.4600	.	16.4800	154900
Security Name	Modify Time	Quotation Time	Quotation Date	Standard Time	Highest Price in 5 Min	Lowest Price in 5 Min	Begin Price in 5 Min		
平安银行	03JAN2017:09:19:38	03JAN2017:09:19:39	2017/1/3	9:20					
平安银行	03JAN2017:09:24:56	03JAN2017:09:24:57	2017/1/3	9:25					
平安银行	03JAN2017:09:25:05	03JAN2017:09:25:06	2017/1/3	9:30	9.11	9.11	9.11		
平安银行	03JAN2017:09:34:57	03JAN2017:09:34:57	2017/1/3	9:35	9.11	9.09	9.11		
平安银行	03JAN2017:09:40:11	03JAN2017:09:39:54	2017/1/3	9:40	9.1	9.09	9.1		
平安银行	03JAN2017:09:45:18	03JAN2017:09:44:57	2017/1/3	9:45	9.11	9.09	9.09		



数据概览

■ 9:30 - 11:30

连续竞价时间。

撮合规则：价格优先、时间优先

涨停价买入如何成交（注册制前/后，创业板/其他板）

连续竞价时，成交价格的确定原则为：

（一）最高买入申报价格与最低卖出申报价格相同，以该价格为成交价格；

（二）买入申报价格高于即时揭示的最低卖出申报价格的，以即时揭示的最低卖出申报价格为成交价格；

（三）卖出申报价格低于即时揭示的最高买入申报价格的，以即时揭示的最高买入申报价格为成交价格。

主板

2023-04-10*

创业板*

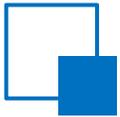
2020-06-12*

科创板

2019-07-22

价格笼子

* 发行的首只注册制股票上市首日



价格笼子

- 买卖股票的，在**连续竞价阶段**的限价申报，应当符合下列有效申报价格范围的要求：
 - （一）买入申报价格不得高于买入基准价格的102%和**买入基准价格以上十个申报价格最小变动单位**的孰高值；
 - （二）卖出申报价格不得低于卖出基准价格的98%和**卖出基准价格以下十个申报价格最小变动单位**的孰低值。
- （仅创业板，且全面注册制后无此条规定）超过有效竞价范围的限价申报不能即时参加竞价，暂存于交易主机；当价格波动使其进入有效竞价范围时，交易主机自动取出申报，参加竞价。

优先顺序	买入基准价格	卖出基准价格
1	即时揭示的最低卖出价格	即时揭示的最高买入价格
2	即时揭示的最高买入价格	即时揭示的最低卖出价格
3	最新成交价	最新成交价
4	前收盘价	前收盘价



数据概览

■ 11:30 - 13:00

休市。

平安银行	04JAN2022:11:29:51.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	10200
平安银行	04JAN2022:11:29:54.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	14200
平安银行	04JAN2022:11:29:57.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	8800
平安银行	04JAN2022:11:30:00.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	8800
平安银行	04JAN2022:11:31:00.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	0
平安银行	04JAN2022:11:32:00.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	0
平安银行	04JAN2022:11:33:00.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	0
平安银行	04JAN2022:11:34:00.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	0
平安银行	04JAN2022:11:35:00.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	0
平安银行	04JAN2022:11:36:00.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	0
平安银行	04JAN2022:11:37:00.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	0
平安银行	04JAN2022:11:38:00.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	0
平安银行	04JAN2022:11:39:00.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	0
平安银行	04JAN2022:11:40:00.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	0



数据概览

■ 13:00 - 14:57

连续竞价时间。

■ 14:57 - 15:00

收盘集合竞价时间。注意，此时交易主机不接受撤单申报。

平安银行	04JAN2022:14:56:57.00	2022-01-04	16.4800	16.4800	16.6600	16.1800	.	16.6500	67700
平安银行	04JAN2022:14:57:00.00	2022-01-04	16.4800	16.4800	16.6600	16.1800	.	16.6500	42400
平安银行	04JAN2022:14:57:09.00	2022-01-04	16.4800	16.4800	16.6600	16.1800	.	16.6500	0
平安银行	04JAN2022:14:57:18.00	2022-01-04	16.4800	16.4800	16.6600	16.1800	.	16.6500	0
平安银行	04JAN2022:14:57:27.00	2022-01-04	16.4800	16.4800	16.6600	16.1800	.	16.6500	0
平安银行	04JAN2022:14:57:36.00	2022-01-04	16.4800	16.4800	16.6600	16.1800	.	16.6500	0
平安银行	04JAN2022:14:57:45.00	2022-01-04	16.4800	16.4800	16.6600	16.1800	.	16.6500	0
平安银行	04JAN2022:14:57:54.00	2022-01-04	16.4800	16.4800	16.6600	16.1800	.	16.6500	0
平安银行	04JAN2022:14:58:03.00	2022-01-04	16.4800	16.4800	16.6600	16.1800	.	16.6500	0
Security Name	Modify Time	Quotation Time	Quotation Date	Standard Time	Highest Price in 1 Min	Lowest Price in 1 Min	Begin Price in 1 Min	Accumulat	
平安银行	03JAN2017:15:12:36	03JAN2017:14:48:57	2017/1/3	14:49	9.17	9.16	9.16	95509	
平安银行	03JAN2017:15:14:23	03JAN2017:14:49:57	2017/1/3	14:50	9.17	9.16	9.16	106768	
平安银行	03JAN2017:15:16:17	03JAN2017:14:50:57	2017/1/3	14:51	9.17	9.16	9.16	49988	
平安银行	03JAN2017:15:18:08	03JAN2017:14:51:57	2017/1/3	14:52	9.17	9.16	9.16	58575	
平安银行	03JAN2017:15:19:48	03JAN2017:14:52:57	2017/1/3	14:53	9.17	9.16	9.17	97287	
平安银行	03JAN2017:15:21:24	03JAN2017:14:53:57	2017/1/3	14:54	9.17	9.16	9.16	194021	
平安银行	03JAN2017:15:22:46	03JAN2017:14:54:57	2017/1/3	14:55	9.17	9.15	9.16	209470	
平安银行	03JAN2017:15:24:15	03JAN2017:14:55:57	2017/1/3	14:56	9.16	9.15	9.15	106300	
平安银行	03JAN2017:15:25:33	03JAN2017:14:56:57	2017/1/3	14:57	9.16	9.15	9.16	77480	
平安银行	03JAN2017:15:26:06	03JAN2017:14:57:54	2017/1/3	14:58	9.16	9.16	9.16	0	
平安银行	03JAN2017:15:26:40	03JAN2017:14:58:57	2017/1/3	14:59	9.16	9.16	9.16	0	
平安银行	03JAN2017:15:27:10	03JAN2017:14:59:51	2017/1/3	15:00	9.16	9.16	9.16	0	



数据概览

■ 15:05 - 15:30

(科创板) 盘后固定价格交易时间, 出现的频率不高。

■ 15:00 -

收盘, 交易状态更新

平安银行	04JAN2022:14:59:06.00	2022-01-04	16.4800	16.4800	16.6600	16.1800	.	16.6500	0
平安银行	04JAN2022:14:59:15.00	2022-01-04	16.4800	16.4800	16.6600	16.1800	.	16.6500	0
平安银行	04JAN2022:14:59:24.00	2022-01-04	16.4800	16.4800	16.6600	16.1800	.	16.6500	0
平安银行	04JAN2022:14:59:33.00	2022-01-04	16.4800	16.4800	16.6600	16.1800	.	16.6500	0
平安银行	04JAN2022:14:59:42.00	2022-01-04	16.4800	16.4800	16.6600	16.1800	.	16.6500	0
平安银行	04JAN2022:14:59:51.00	2022-01-04	16.4800	16.4800	16.6600	16.1800	.	16.6500	0
平安银行	04JAN2022:15:00:03.00	2022-01-04	16.4800	16.4800	16.6600	16.1800	16.6600	16.6600	1407900

Security Name	Modify Time	Quotation Time	Quotation Date	Standard Time	Highest Price in 1 Min	Lowest Price in 1 Min	Begin Price in 1 Min	Accumulat
平安银行	03JAN2017:15:12:36	03JAN2017:14:48:57	2017/1/3	14:49	9.17	9.16	9.16	95509
平安银行	03JAN2017:15:14:23	03JAN2017:14:49:57	2017/1/3	14:50	9.17	9.16	9.16	106768
平安银行	03JAN2017:15:16:17	03JAN2017:14:50:57	2017/1/3	14:51	9.17	9.16	9.16	49988
平安银行	03JAN2017:15:18:08	03JAN2017:14:51:57	2017/1/3	14:52	9.17	9.16	9.16	58575
平安银行	03JAN2017:15:19:48	03JAN2017:14:52:57	2017/1/3	14:53	9.17	9.16	9.17	97287
平安银行	03JAN2017:15:21:24	03JAN2017:14:53:57	2017/1/3	14:54	9.17	9.16	9.16	194021
平安银行	03JAN2017:15:22:46	03JAN2017:14:54:57	2017/1/3	14:55	9.17	9.15	9.16	209470
平安银行	03JAN2017:15:24:15	03JAN2017:14:55:57	2017/1/3	14:56	9.16	9.15	9.15	106300
平安银行	03JAN2017:15:25:33	03JAN2017:14:56:57	2017/1/3	14:57	9.16	9.15	9.16	77480
平安银行	03JAN2017:15:26:06	03JAN2017:14:57:54	2017/1/3	14:58	9.16	9.16	9.16	0
平安银行	03JAN2017:15:26:40	03JAN2017:14:58:57	2017/1/3	14:59	9.16	9.16	9.16	0
平安银行	03JAN2017:15:27:10	03JAN2017:14:59:51	2017/1/3	15:00	9.16	9.16	9.16	0
平安银行	03JAN2017:15:27:20	03JAN2017:15:00:06	2017/1/3	15:01	9.16	9.16	9.16	579500



数据概览

■ 特殊时期处理

例如2016年初的熔断机制时期。

平安银行	04JAN2016:13:33:09.00	04JAN2016:13:33:27.00	2016-01-04	11.9900	12.0000	12.0300	11.3500	11.3600
平安银行	04JAN2016:13:33:12.00	04JAN2016:13:33:30.00	2016-01-04	11.9900	12.0000	12.0300	11.3500	11.3500
平安银行	04JAN2016:13:33:15.00	04JAN2016:13:33:33.00	2016-01-04	11.9900	12.0000	12.0300	11.3300	11.3300
平安银行	04JAN2016:13:33:17.00	04JAN2016:13:33:36.00	2016-01-04	11.9900	12.0000	12.0300	11.3300	11.3300
平安银行	04JAN2016:13:33:21.00	04JAN2016:13:33:39.00	2016-01-04	11.9900	12.0000	12.0300	11.3300	11.3400
平安银行	04JAN2016:13:33:26.00	04JAN2016:13:33:45.00	2016-01-04	11.9900	12.0000	12.0300	11.3200	11.3200
平安银行	04JAN2016:13:33:29.00	04JAN2016:13:33:48.00	2016-01-04	11.9900	12.0000	12.0300	11.3200	11.3300
平安银行	04JAN2016:13:33:33.00	04JAN2016:13:33:51.00	2016-01-04	11.9900	12.0000	12.0300	11.3200	11.3300
平安银行	04JAN2016:13:33:36.00	04JAN2016:13:33:54.00	2016-01-04	11.9900	12.0000	12.0300	11.3200	11.3200
平安银行	04JAN2016:13:33:39.00	04JAN2016:13:33:57.00	2016-01-04	11.9900	12.0000	12.0300	11.3100	11.3300
平安银行	04JAN2016:13:33:41.00	04JAN2016:13:34:00.00	2016-01-04	11.9900	12.0000	12.0300	11.3100	11.3200
平安银行	04JAN2016:13:33:45.00	04JAN2016:13:34:03.00	2016-01-04	11.9900	12.0000	12.0300	11.2300	11.3200
平安银行	04JAN2016:13:33:48.00	04JAN2016:13:34:06.00	2016-01-04	11.9900	12.0000	12.0300	11.2300	11.3000
平安银行	04JAN2016:15:00:13.00	04JAN2016:15:00:34.00	2016-01-04	11.9900	12.0000	12.0300	11.2300	11.3300
平安银行	04JAN2016:15:01:52.00	04JAN2016:15:02:13.00	2016-01-04	11.9900	12.0000	12.0300	11.2300	11.3300
平安银行	05JAN2016:09:05:11.00	05JAN2016:09:05:34.00	2016-01-05	11.3300
平安银行	07JAN2016:09:58:13.00	07JAN2016:09:58:33.00	2016-01-07	11.5300	11.4100	11.4100	10.9300	10.9400
平安银行	07JAN2016:09:58:17.00	07JAN2016:09:58:36.00	2016-01-07	11.5300	11.4100	11.4100	10.9300	10.9500
平安银行	07JAN2016:09:58:21.00	07JAN2016:09:58:39.00	2016-01-07	11.5300	11.4100	11.4100	10.9100	10.9300
平安银行	07JAN2016:09:58:24.00	07JAN2016:09:58:42.00	2016-01-07	11.5300	11.4100	11.4100	10.9100	10.9300
平安银行	07JAN2016:09:58:27.00	07JAN2016:09:58:45.00	2016-01-07	11.5300	11.4100	11.4100	10.9100	10.9400
平安银行	07JAN2016:09:58:32.00	07JAN2016:09:58:48.00	2016-01-07	11.5300	11.4100	11.4100	10.9100	10.9200
平安银行	07JAN2016:09:58:35.00	07JAN2016:09:58:51.00	2016-01-07	11.5300	11.4100	11.4100	10.9100	10.9500
平安银行	07JAN2016:09:58:44.00	07JAN2016:09:59:12.00	2016-01-07	11.5300	11.4100	11.4100	10.9100	10.9100
平安银行	07JAN2016:15:00:01.00	07JAN2016:15:00:30.00	2016-01-07	11.5300	11.4100	11.4100	10.9100	10.9400
平安银行	07JAN2016:15:00:57.00	07JAN2016:15:01:27.00	2016-01-07	11.5300	11.4100	11.4100	10.9100	10.9400
平安银行	08JAN2016:09:19:37.00	08JAN2016:09:20:06.00	2016-01-08	10.9400
平安银行	08JAN2016:09:19:40.00	08JAN2016:09:20:12.00	2016-01-08	10.9400
平安银行	08JAN2016:09:19:49.00	08JAN2016:09:20:21.00	2016-01-08	10.9400



数据概览

交易所接口说明书

需要注意的是，注册制下，交易规则的变更不仅仅是我刚刚提到的内容，还有更多的调整，比如交易所新增了订单类型等等，这些调整目前来看不会对之前已有数据的解读产生明显的影响，等到新的数据拿到之后，我们会继续进行评估。同样，我今天所讲的内容，随着时间的推移，当大家拿到新的数据，或许又需要进行新的调整，所以这就要求大家，对资本市场的政策要求以及历史的变化都要比较了解，这样，当大家面对高频数据的时候，才能够真正的理解这些数据背后到底发生了什么。

关于我们前面提到的Level 2数据变量的种种变动，我们可以去参考交易所官方的数据接口文档，文档的开头“文档变动说明”会给大家简要的说明每次的变化。比如我这里给大家截图的是上交所Level II数据接口文档中“文档变动说明”的一部分。从这里我们可以看到高频数据内容更改的一些端倪。

2018-08-13	1.1.1	<ol style="list-style-type: none">1. 调整 UA3202 消息说明2. 更新交易日时间表3. 更新交易状态（InstrumentStatus）字段取值，新增 CCALL，删除 BETW 和 BREAK
2018-10-30	2.0.0	<ol style="list-style-type: none">1. 取消虚拟集合竞价消息 UA3107，相关行情数据通过行情快照消息 UA3202 消息发送（机制同收盘集合竞价）2. 行情快照消息 UA3202 发布时间从 9:25 统一调整到 8:453. 行情快照消息 UA3202 的交易状态（Tag10135）会体现 START（开市）、OCALL（开盘集合竞价）4. 调整行情快照消息 UA3202 在停牌期间的发送机制及交易状态

PART

4

数据词典介绍





五档行情

变量名	中文全称	英文全称	类型	长度	单位	注释
标识与日期						
Exchflg	交易所标识	Exchange Flag	数值	1		1—上交所；2—深交所；9—申万指数。
Code	代码	Code	字符	6		
Code_Mkt	代码与市场标识	Code and Market Flag	字符	10		
Secunm	证券名称	Security Name	字符	8		
MTime	修改时间	Modify Time	时间	22.2		修改时间，精确到秒。
QTime	行情时间	Quotation Time	时间	22.2		交易时间，精确到秒。
Qdate	行情日期	Quotation Date	时间	10		交易日期，精确到日。
价格与成交量						
PrevClPr	前收盘价	Previous Close Price	数值	10.4	元	
Oppr	开盘价	Open Price	数值	10.4	元	
Hipr	最高价	Highest Price	数值	10.4	元	当日成交的最高价。
Lopr	最低价	Lowest Price	数值	10.4	元	当日成交的最低价。
TPrice	成交价	Trade Price	数值	10.4	元	当笔交易的成交价。
TVolume	成交量	Trade Volume	数值	16	股	当笔交易的成交量。
TSum	成交额	Trade Sum	数值	16.4	元	当笔交易的成交金额。
TDeals	成交笔数	Turnover Deals	数值	20		仅深交所公布。
TVolume_accu	累计成交量	Accumulated Trade Volume	数值	23	股	截至现时所有交易的累计成交量。
TSum_accu	累计成交额	Accumulated Trade Sum	数值	23.2	元	截至现时所有交易的累计成交额。
TDeals_accu	累计成交笔数	Accumulated Turnover Deals	数值	12		截至现时所有交易的累计成交笔数。



五档行情

Secunm	证券名称	Security Name	字符	8	
MTime	修改时间	Modify Time	时间	22.2	修改时间，精确到秒。
QTime	行情时间	Quotation Time	时间	22.2	交易时间，精确到秒。
Qdate	行情日期	Quotation Date	时间	10	交易日期，精确到日。
价格与成交量					
PrevClPr	前收盘价	Previous Close Price	数值	10.4元	
Oppr	开盘价	Open Price	数值	10.4元	
Hipr	最高价	Highest Price	数值	10.4元	当日成交的最高价。
Lopr	最低价	Lowest Price	数值	10.4元	当日成交的最低价。
TPrice	成交价	Trade Price	数值	10.4元	当笔交易的成交价。
TVolume	成交量	Trade Volume	数值	16股	当笔交易的成交量。
TSum	成交额	Trade Sum	数值	16.4元	当笔交易的成交金额。

累计量

Quotation Time	Quotation Date	Previous Close Price	Open Price	High Price	Low Price	ClPr	Trade Price	TVolume
04JAN2022:11:29:54.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	14200
04JAN2022:11:29:57.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	8800
04JAN2022:11:30:00.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	8800
04JAN2022:11:31:00.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	0
04JAN2022:11:32:00.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	0
04JAN2022:11:33:00.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	0
04JAN2022:11:34:00.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	0
04JAN2022:11:35:00.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	0
04JAN2022:11:36:00.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	0
04JAN2022:11:37:00.00	2022-01-04	16.4800	16.4800	16.5000	16.1800	.	16.4100	0

交易停止，累计量不再变化



五档行情

StdTime	标准时间	Standard Time	时间			分时标准时间，精确到分。
期间						
Highpr	期间最高价	Highest Price	数值	10.4	元	
Lowpr	期间最低价	Lowest Price	数值	10.4	元	
Begpr	期初成交价	Begin Price	数值	10.4	元	
TVolume_accu1	期间累计成交量	Accumulated Trade Volume	数值	16	股	分时期间内所有交易的累计成交量。
TSum_accu1	期间累计成交额	Accumulated Trade Sum	数值	20.2	元	分时期间内所有交易的累计成交金额。
TDeals_accu1	期间累计成交笔数	Accumulated Turnover Deals	数值	12		分时期间内所有交易的累计成交笔数。

期间量

Security Name	Modify Time	Quotation Time	Quotation Date	Standard Time	Highest Price in 1 Min	Lowest Price in 1 Min	Begin Price in 1 Min	Accumulat
平安银行	03JAN2017:15:12:36	03JAN2017:14:48:57	2017/1/3	14:49	9.17	9.16	9.16	95509
平安银行	03JAN2017:15:14:23	03JAN2017:14:49:57	2017/1/3	14:50	9.17	9.16	9.16	106768
平安银行	03JAN2017:15:16:17	03JAN2017:14:50:57	2017/1/3	14:51	9.17	9.16	9.16	49988
平安银行	03JAN2017:15:18:08	03JAN2017:14:51:57	2017/1/3	14:52	9.17	9.16	9.16	58575
平安银行	03JAN2017:15:19:48	03JAN2017:14:52:57	2017/1/3	14:53	9.17	9.16	9.17	97287
平安银行	03JAN2017:15:21:24	03JAN2017:14:53:57	2017/1/3	14:54	9.17	9.16	9.16	194021
平安银行	03JAN2017:15:22:46	03JAN2017:14:54:57	2017/1/3	14:55	9.17	9.15	9.16	209470
平安银行	03JAN2017:15:24:15	03JAN2017:14:55:57	2017/1/3	14:56	9.16	9.15	9.15	106300
平安银行	03JAN2017:15:25:33	03JAN2017:14:56:57	2017/1/3	14:57	9.16	9.15	9.16	77480
平安银行	03JAN2017:15:26:06	03JAN2017:14:57:54	2017/1/3	14:58	9.16	9.16	9.16	0
平安银行	03JAN2017:15:26:40	03JAN2017:14:58:57	2017/1/3	14:59	9.16	9.16	9.16	0
平安银行	03JAN2017:15:27:10	03JAN2017:14:59:51	2017/1/3	15:00	9.16	9.16	9.16	0
平安银行	03JAN2017:15:27:20	03JAN2017:15:00:06	2017/1/3	15:01	9.16	9.16	9.16	579500



五档行情-成交笔数

价格与成交量						
PrevClPr	前收盘价	Previous Close Price	数值	10.4	元	
Oppr	开盘价	Open Price	数值	10.4	元	
Hipr	最高价	Highest Price	数值	10.4	元	当日交易的最高价。
Lopr	最低价	Lowest Price	数值	10.4	元	当日交易的最低价。
TPrice	成交价	Trade Price	数值	10.4	元	当笔交易的成交价。
TVolume	成交量	Trade Volume	数值	16	股	当笔交易的成交量。
TSum	成交额	Trade Sum	数值	16.4	元	当笔交易的成交金额。
TDeals	成交笔数	Turnover Deals	数值	20		仅深交所公布。
TVolume_accu	累计成交量	Accumulated Trade Volume	数值	23	股	截至现时所有交易的累计成交量。
TSum_accu	累计成交额	Accumulated Trade Sum	数值	23.2	元	截至现时所有交易的累计成交额。
TDeals_accu	累计成交笔数	Accumulated Turnover Deals	数值	12		截至现时所有交易的累计成交笔数。

限价单簿 (买)		
投资者	价格	手
老王	10.8	1
老王	10.7	2
小张	10.8	5
大橘	10.5	3

BidPr1	申买价一	10.8
BidVol1	申买量一	6
BidDeals1	申买笔数1	2



五档行情-市场买卖指标

市场买卖指标部分，是锐思数据根据主流计算方法计算出来供大家参考的指标。除此部分指标外，其他所有的变量都是由交易所直接提供，我们并没有做任何修改。

	市场买卖指标					
Trdirec	买卖方向	Trade Direction	字符	1		B—买方主动成交； S—卖方主动成交； F—无法确定买卖方向。 计算方法见注3。
Absspread	绝对买卖报价差	Absolute Spread	数值	12.4	元	计算方法：买价1-卖价1。
Respread	相对买卖报价差	Relative Spread	数值	12.8		计算方法： $(\text{买价1}-\text{卖价1}) / (0.5 * (\text{买价1} + \text{卖价1}))$ 。
Abseffspread	绝对有效买卖报价差	Absolute Effective Spread	数值	12.4	元	计算方法： $2 * \text{成交价} - 0.5 * (\text{买价1} + \text{卖价1}) $ 。
Reeffspread	相对有效买卖报价差	Relative Effective Spread	数值	12.8		计算方法： $2 * \text{成交价} - 0.5 * (\text{买价1} + \text{卖价1}) / (0.5 * (\text{买价1} + \text{卖价1}))$ 。
Depth1	深度1	Depth1	数值	18.2	元	计算方法： $\text{买价1} * \text{买入数量1} + \text{卖价1} * \text{卖出数量1}$ 。
Depth2	深度2	Depth2	数值	18.2	元	计算方法： $\sum (\text{买价} * \text{买入数量} + \text{卖价} * \text{卖出数量})$ 。

PART

5

学术应用



锐思高频数据库助力教学与科研

300

300余所
国内外高校用户

95%

财经类院校
用户覆盖率超95%

8000

每年8000余篇
期刊、论文引用量
(不完全统计)

50000

50000多名
个人注册用户

引用锐思高频数据发表的论文示例

题名	来源	年/期	作者	第一作者单位
基于金融高频数据的LASSO-CDRD协方差矩阵预测模型	统计研究	2022/09	刘广应; 包悦妍; 林金官	南京审计大学
基于新型指标的股票价格操纵识别模型研究	江淮论坛	2020/05	刘振清; 王旭; 姚远	河南大学管理科学与工程研究所
知情交易概率的贝叶斯估计	金融发展研究	2019/11	郇钰	中国工商银行博士后科研工作站
基于高频交易数据的“羊群行为”测度模型	数量经济技术经济研究	2019/08	朱菲菲; 唐涯; 徐建国; 李宏泰	北京大学光华管理学院
短期羊群行为的影响因素与价格效应——基于高频数据的实证检验	金融研究	2019/07	朱菲菲; 李惠璇; 徐建国; 李宏泰	北京大学光华管理学院
盈余公告、分析师推荐与伪羊群行为——基于高频数据的实证检验	经济学(季刊)	2019/04	李惠璇; 朱菲菲; 唐涯; 李宏泰	北京工商大学经济学院
A股波动率长记忆模型的探索研究	中国商论	2018-01	丁宇; 陈熙哲	广东外语外贸大学会计学院
利用高频数据预测沪深300指数波动率——基于Realized GARCH模型的实证研究	世界经济文汇	2014/10	王天一; 赵晓军; 黄卓	对外经济贸易大学金融学院



RESSET高频数据的学术应用

《中国股票市场波动率的高频估计与特性分析》

■ 研究问题

■ 研究背景

■ 数据选取

■ 频率选择

■ 研究结论

经济研究 2003年第2期

中国股票市场波动率的高频估计与特性分析*

黄后川

陈浪南

(南方基金管理公司 510200) (中山大学岭南学院与中山大学经济所 510275)

内容提要:本文旨在应用高频数据估计中国股市的已实现波动率。我们发现股票指数与个股的高频交易数据中的微观摩擦影响正好相反,使用极高频的数据会大大增加个股的波动率估计值,相反则会大大降低指数的波动率估计值。在计算各种频率的已实现波动率的基础上,本文构造了一种较为精确的估计波动率的方法,可以更好地平衡测量误差与微观结构误差。基于已实现波动率,本文研究了中国股市波动率不对称性和长期记忆特性。

关键词:波动率 高频估计 特征

一、引言与先前该领域研究述评

近二十年来,对波动率模型的研究已成为金融经济学领域研究的重要内容之一。自Eagle于1982年提出ARCH模型以来,经济学界已经发表了数千篇关于条件异方差或波动率的论文。特别是最近十年,一些学者提出用高频分时数据估计波动率的方法,这种方法可以得到比较准确的波动率估计值,称为“已实现波动率”(Realized Volatility)。以此为基础,众多学者在波动率的特性和预测两方面进行了更深入的研究,大大拓展了这个研究领域。

Andersen、Bollerslev、Diebold、Ebens(1998, 2001)等金融经济学家对这种高频估计方法以及“已实现波动率”的特性与预测进行了一系列研究,他们得出了如下几个主要结论(计算的波动率都是日波动率):

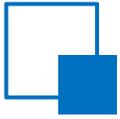
(1) 如果价格遵循普通的扩散过程,用此方法计算的已实现波动率,是无偏的。而且,当高频数据的时间间隔趋近于0时,已实现波动率的测量误差也趋于0。因此可以把已实现波动率当作一个观测值,它没有经典算法所带来的时间滞后。

(2) 通过对外汇市场和道·琼斯工业股票的实证研究,发现:① 股票市场中,正收益对后续波动率的影响不如负收益明显,即波动率具有杠杆效应。② 已实现波动率的对数具有明显的长期记忆特性。③ 虽然已实现波动率明显向右倾斜,但已实现波动率的对数呈现正态分布。④ 虽然原始的收益率数据有明显的高峰和大尾巴,但收益率除已实现波动率呈现正态分布。⑤ 股票市场的波动率与相关度呈相同方向运动,降低了资产组合分散化在高波动率时的作用。

(3) 依据(2)中的结论,用体现长期记忆的分综合一移动平均自回归(Auto Regression Fractional Integrated Moving Average, ARFIMA)方法可以得到更好的波动率预测。使用正态一对数正态混合分布可以得到很好的概率密度和分位数估计(例如VaR)。

已实现波动率的一个重要用途是作为对以前各类模型进行评价的基准。它的另一个更重要的用途是用于检验波动率的各种特性,并对未来波动率进行预测,因为已实现波动率可以直接当作波动率的一种观测值,因此可以采用一般的时间序列方法,无须像ARCH模型一样通过模拟收益率序

* 本文是国家自然科学基金课题(79800010, 70042005)、上交所2002年联合研究计划课题、教育部社科“十五”课题(01j790026)及2002年厦门大学校级课题成果之一。



RESSET高频数据的学术应用

《中国股票市场波动率的高频估计与特性分析》

- 研究问题：个股和指数的已实现波动率的估计和比较。

- 研究背景

- 近二十年来，对波动率模型的研究已成为金融经济学领域研究的重要内容之一。自 Engle 于1982 年提出 ARCH 模型以来，经济学界已经发表了数千篇关于条件异方差或波动率的论文。特别是最近十年，一些学者提出用高频分时数据估计波动率的方法，这种方法可以得到比较准确的波动率估计值，称为“已实现波动率”（Realized Volatility）。以此为基础，众多学者在波动率的特性和预测两方面进行了更深入的研究，大大拓展了这个研究领域。



RESSET高频数据的学术应用

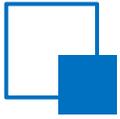
《中国股票市场波动率的高频估计与特性分析》

■ 数据选取

2000到2002年间上海证券交易所股票的日内分时交易数据。

右图为股票5分钟分时数据的截图。

Quotation	Standard T	Highest Pri	Lowest Pri	Begin Price	Accumulat	Accumulat	Accumulat	Previous Cl	Open Price	High Price	Low Price
2017/1/3	9:35	9.11	9.09	9.11	1754100	15966972	477	9.1	9.11	9.11	9.09
2017/1/3	9:40	9.1	9.09	9.1	812973	7396443	242	9.1	9.11	9.11	9.09
2017/1/3	9:45	9.11	9.09	9.09	2300124	20935114	373	9.1	9.11	9.11	9.09
2017/1/3	9:50	9.14	9.1	9.1	4486269	40915185	797	9.1	9.11	9.14	9.09
2017/1/3	9:55	9.17	9.14	9.14	4129157	37802704	1126	9.1	9.11	9.17	9.09
2017/1/3	10:00	9.17	9.15	9.16	1774416	16253361	540	9.1	9.11	9.17	9.09
2017/1/3	10:05	9.18	9.16	9.17	2778388	25480342	693	9.1	9.11	9.18	9.09
2017/1/3	10:10	9.17	9.16	9.16	1444560	13241895	480	9.1	9.11	9.18	9.09
2017/1/3	10:15	9.17	9.16	9.17	885724	8120767	360	9.1	9.11	9.18	9.09
2017/1/3	10:20	9.17	9.16	9.17	872220	7992628	336	9.1	9.11	9.18	9.09
2017/1/3	10:25	9.17	9.15	9.16	407360	3732072	216	9.1	9.11	9.18	9.09
2017/1/3	10:30	9.17	9.16	9.17	590975	5418081	177	9.1	9.11	9.18	9.09
2017/1/3	10:35	9.17	9.16	9.16	997220	9142955	396	9.1	9.11	9.18	9.09
2017/1/3	10:40	9.18	9.16	9.17	314200	2879915	178	9.1	9.11	9.18	9.09
2017/1/3	10:45	9.17	9.16	9.17	594600	5449837	246	9.1	9.11	9.18	9.09
2017/1/3	10:50	9.17	9.16	9.17	463420	4247062	207	9.1	9.11	9.18	9.09
2017/1/3	10:55	9.17	9.15	9.16	830280	7606611	332	9.1	9.11	9.18	9.09
2017/1/3	11:00	9.16	9.15	9.16	646478	5916442	263	9.1	9.11	9.18	9.09
2017/1/3	11:05	9.16	9.15	9.15	418300	3829022	157	9.1	9.11	9.18	9.09
2017/1/3	11:10	9.16	9.15	9.16	293660	2688441	144	9.1	9.11	9.18	9.09
2017/1/3	11:15	9.16	9.15	9.16	527600	4828733	191	9.1	9.11	9.18	9.09
2017/1/3	11:20	9.16	9.14	9.16	693421	6340112	260	9.1	9.11	9.18	9.09
2017/1/3	11:25	9.15	9.13	9.14	618560	5655539	220	9.1	9.11	9.18	9.09
2017/1/3	11:30	9.15	9.13	9.14	448504	4098280	143	9.1	9.11	9.18	9.09
2017/1/3	11:35	9.14	9.14	9.14	3100	28334	2	9.1	9.11	9.18	9.09
2017/1/3	13:05	9.14	9.13	9.14	639592	5844002	264	9.1	9.11	9.18	9.09
2017/1/3	13:10	9.15	9.13	9.14	518800	4741797	175	9.1	9.11	9.18	9.09
2017/1/3	13:15	9.15	9.13	9.14	370979	3389762	126	9.1	9.11	9.18	9.09

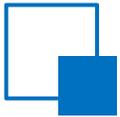


RESSET高频数据的学术应用

《中国股票市场波动率的高频估计与特性分析》

■ 分时频率的选择

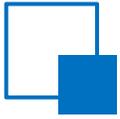
- 由于现实当中的价格并不是严格服从正态发散过程的，因此在估计已实现波动率时，高频数据的时间间隔并不是越小越好，用极高频数据估计波动率往往会因为微观摩擦而产生较大误差。
- Torben G .Andersen 等人在对道·琼斯 30 只工业股票做实证分析时，直接选取了 5 分钟数据，而在其它文献中他们又使用“标记图”方法得出较佳时间间隔。
- 因此，用于估计“已实现波动率”的最佳频率是与市场微观特性相关的，在不同市场、不同时期可能会有所不同。



《中国股票市场波动率的高频估计与特性分析》

■ 研究结论

1. 首先，我们发现股票指数与个股的高频交易数据中的微观摩擦影响正好相反，使用极高频的数据会大大增加个股的波动率估计值，相反却会大大降低指数的波动率估计值。
 - 因为测量误差与微观摩擦误差在各个时期有较大变化，所以直接选用 5 分钟或 15 分钟交易数据并不可取，凭主观进行判断的“标记图”方法也有一定偏差。
 - 本文中构造了一种较为精确的估计波动率的方法，得到了较为准确的上证A、B 指数的波动率序列。



《中国股票市场波动率的高频估计与特性分析》

■ 研究结论

2. 其次，基于已实现波动率，本文研究了中国股市波动率的不对称性和长期记忆特性。在波动率的不对称特性方面，上证 A 、 B 指均有明显对上期收益和本期收益的不对称性，折线回归可以很好地模拟这种不对称性；

- 在波动率的长期记忆特性方面，上证 A 、 B 指波动率及其相关度都有明显的持续性和长期记忆特征，上证 B 指波动率持续性较弱，上证 A 指波动率和两个指数的相关度持续性较强些



《存款准备金率调整对股票市场的影响 ——基于中国股市高频数据的实证研究》

- 研究问题
- 创新之处
- 研究假设
- 模型构建
- 研究结论

理论综合

存款准备金率调整对股票市场的影响

——基于中国股市高频数据的实证研究

李平¹ 高洁¹ 廖静池²

1.电子科技大学经济与管理学院, 成都 611731; 2.深圳证券交易所综合研究所, 广东 深圳 518038

摘要: 本文采用中国股票市场的日内高频数据实证检验了2006-2011年期间28次存款准备金率上调事件和4次下调事件对股票市场收益率的短期影响。与现有基于低频数据的文献认为存款准备金率调整对中国股票市场没有影响的研究结论不同, 本文的实证结果表明: 第一, 存款准备金率调整在短期内会导致股票市场总体的价格发生显著变化, 但这种影响在持续一段时间后会消失, 并且全天的收益情况与市场行情(牛市或熊市)密切相关; 第二, 相对于其他行业, 存款准备金率上调对房地产行业的影响更为明显, 导致其指数的累积收益率在大部分时段显著为负。总体而言, 存款准备金率调整虽然短期内向市场准确地传递了信号, 但受市场行情的影响, 长期来看并未对股票市场造成实质性影响。

关键词: 存款准备金率; 事件研究法; Bootstrap方法; 高频数据

Abstract: Using the intraday high frequency data from Chinese stock market, this paper empirically investigates the impact of 28 events of the increase and 4 events of the decrease in deposit reserve requirement occurred during the period 2006-2011 on the return of stock market in the short term. Existing literature argue that the adjustment of deposit reserve requirement ratio in China has no effect on stock market. However, we find that the adjustment of deposit reserve requirement ratio leads to significant change in the overall stock price in the short term, and this effect will continue and then disappear after a period of time. Moreover, whether the daily return on stock is positive or not depends on the market sentiment (bull market or bear market). We also find that the increase of deposit reserve requirement ratio produces greater effect upon the stock of real estate industry than stock of other industries, that is, the cumulative returns on stock of real estate industry are significantly negative in most of time. Thus, we conclude that the deposit reserve requirement ratio adjustment in China releases an accurate signal in the short term, but has no material impact on stock markets in the long term because of the effect of market sentiment.

Keywords: deposit reserve requirement ratio, event study, bootstrap approach, high frequency data

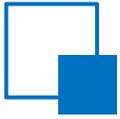
作者简介: 李平, 电子科技大学经济与管理学院副教授, 研究方向: 市场微观结构。高洁, 女, 电子科技大学经济与管理学院硕士生, 研究方向: 市场微观结构。廖静池, 深圳证券交易所综合研究所研究员, 研究方向: 市场微观结构。

中图分类号: F830.91 文献标识码: A

引言

2000年以来, 随着我国国际贸易顺差增大, 固定资产投资活跃, 信贷投放增长过快, 国内经济出现过热苗头, 通货膨胀日趋明显(陆斌峰和高攀, 2011)^[20]。出于收缩流动性、减轻通胀压力、缓解房地产价格上涨势头过快等原因(李阳, 2007; 杨琳, 2010)^{[7][22]}, 中国人民

银行于2006年6月16日开始频繁小幅调整存款准备金率(以下简称存准率)。截止2008年6月7日, 人民银行共18次上调存准率, 累计上调幅度达10%。然而, 2008年开始的国际金融危机对国内经济的冲击又迫使人民银行不得不迅速转变货币政策调控方向, 从2008年9月25日起连续4次下调存准率。同时, 国务院也出台“四万亿救市计划”以实现经济的“软着陆”。巨大的信贷供给虽然减



RESSET高频数据的学术应用

《存款准备金率调整对股票市场的影响——基于中国股市高频数据的实证研究》

- 研究问题：存款准备金率的调整对股票短期收益率的影响。

- 背景

- 现有文献均基于低频的日交易数据，只能检验存准率调整对股票市场的日间效应和长期影响，无法考察存准率调整在宣告后首个交易日的日内变化情况和短期影响。
- 事实上，随着通讯技术的不断发展，市场对信息的获取和消化速度大大加快，事件对市场造成的短期影响很可能在日内就已经消化了。



RESSET高频数据的学术应用

《存款准备金率调整对股票市场的影响——基于中国股市高频数据的实证研究》

■ 创新之处

- 采用高频数据实证检验了存准率调整对中国股票市场的影响，提高检验的可信度。
- 现有基于低频数据的文献认为存准率调整对股票市场没有影响，但并没有进一步证实造成这一现象的内在原因，而本文的研究结论则表明存准率调整仍然具有正向的信号传递作用，只是这种宣告效应可能在日内已被市场完全消化，并且受市场行情的影响，长期来看并未对市场造成实质影响。
- 此外，不同行业的股票对存准率调整的反应也不尽相同。相对而言，对宏观经济政策更加敏感的房地行业的反应更加明显。



RESSET高频数据的学术应用

《存款准备金率调整对股票市场的影响——基于中国股市高频数据的实证研究》

■ 研究假设

- 假设1：总体而言，存准率上调(下调)在短期内会导致股票市场的价格下跌(上涨)，但全天的收益情况则与市场行情(牛市或熊市)有关。换言之，存准率调整仍然具有正向的信号传递作用，但受市场行情的影响，长期来看可能并不会产生实质性的影响。
- 假设2：相对于其他行业，存准率上调(下调)对房地产行业的股票影响更为明显，具体表现为在短期内会导致房地产行业指数的收益率显著下跌(上涨)。



RESSET高频数据的学术应用

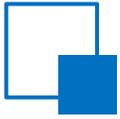
《存款准备金率调整对股票市场的影响——基于中国股市高频数据的实证研究》

■ 数据选取

2006到2011年间上海证券交易所和深圳证券交易所股票和指数的日内5分钟分时交易数据。

右图为股票5分钟分时数据的截图。

Quotation	Standard T	Highest Pri	Lowest Pri	Begin Price	Accumulat	Accumulat	Accumulat	Previous Cl	Open Price	High Price	Low Price
2017/1/3	9:35	9.11	9.09	9.11	1754100	15966972	477	9.1	9.11	9.11	9.09
2017/1/3	9:40	9.1	9.09	9.1	812973	7396443	242	9.1	9.11	9.11	9.09
2017/1/3	9:45	9.11	9.09	9.09	2300124	20935114	373	9.1	9.11	9.11	9.09
2017/1/3	9:50	9.14	9.1	9.1	4486269	40915185	797	9.1	9.11	9.14	9.09
2017/1/3	9:55	9.17	9.14	9.14	4129157	37802704	1126	9.1	9.11	9.17	9.09
2017/1/3	10:00	9.17	9.15	9.16	1774416	16253361	540	9.1	9.11	9.17	9.09
2017/1/3	10:05	9.18	9.16	9.17	2778388	25480342	693	9.1	9.11	9.18	9.09
2017/1/3	10:10	9.17	9.16	9.16	1444560	13241895	480	9.1	9.11	9.18	9.09
2017/1/3	10:15	9.17	9.16	9.17	885724	8120767	360	9.1	9.11	9.18	9.09
2017/1/3	10:20	9.17	9.16	9.17	872220	7992628	336	9.1	9.11	9.18	9.09
2017/1/3	10:25	9.17	9.15	9.16	407360	3732072	216	9.1	9.11	9.18	9.09
2017/1/3	10:30	9.17	9.16	9.17	590975	5418081	177	9.1	9.11	9.18	9.09
2017/1/3	10:35	9.17	9.16	9.16	997220	9142955	396	9.1	9.11	9.18	9.09
2017/1/3	10:40	9.18	9.16	9.17	314200	2879915	178	9.1	9.11	9.18	9.09
2017/1/3	10:45	9.17	9.16	9.17	594600	5449837	246	9.1	9.11	9.18	9.09
2017/1/3	10:50	9.17	9.16	9.17	463420	4247062	207	9.1	9.11	9.18	9.09
2017/1/3	10:55	9.17	9.15	9.16	830280	7606611	332	9.1	9.11	9.18	9.09
2017/1/3	11:00	9.16	9.15	9.16	646478	5916442	263	9.1	9.11	9.18	9.09
2017/1/3	11:05	9.16	9.15	9.15	418300	3829022	157	9.1	9.11	9.18	9.09
2017/1/3	11:10	9.16	9.15	9.16	293660	2688441	144	9.1	9.11	9.18	9.09
2017/1/3	11:15	9.16	9.15	9.16	527600	4828733	191	9.1	9.11	9.18	9.09
2017/1/3	11:20	9.16	9.14	9.16	693421	6340112	260	9.1	9.11	9.18	9.09
2017/1/3	11:25	9.15	9.13	9.14	618560	5655539	220	9.1	9.11	9.18	9.09
2017/1/3	11:30	9.15	9.13	9.14	448504	4098280	143	9.1	9.11	9.18	9.09
2017/1/3	11:35	9.14	9.14	9.14	3100	28334	2	9.1	9.11	9.18	9.09
2017/1/3	13:05	9.14	9.13	9.14	639592	5844002	264	9.1	9.11	9.18	9.09
2017/1/3	13:10	9.15	9.13	9.14	518800	4741797	175	9.1	9.11	9.18	9.09
2017/1/3	13:15	9.15	9.13	9.14	370979	3389762	126	9.1	9.11	9.18	9.09



RESSET高频数据的学术应用

《存款准备金率调整对股票市场的影响——基于中国股市高频数据的实证研究》

■ 模型构建

- 由于样本数量太少，t检验的结果可能存在偏误。为了解决小样本的假设检验和统计推断问题，Efron(1979) 提出了一种非参数方法——自助法(Bootstrap)。对收益率序列进行有放回的再抽样，得到一组样本容量为N 的新样本。根据这组新样本进行重新计算，得到新的收益率指标。
- 采用事件研究法的思想，使用累计收益率 (CR) 和累计异常收益率 (CAR)，深入考察存准率调整事件发生后股票市场的表现与在非事件日的差异。



RESSET高频数据的学术应用

《存款准备金率调整对股票市场的影响——基于中国股市高频数据的实证研究》

■ 研究结论

1. 研究结论表明，虽然存准率调整并未对市场造成长期影响，但短期内，下调会导致股票市场总体的价格显著下跌，而上调对股票市场的影响与市场行情密切相关。
 - 在熊市行情中，无论是累积收益率，还是累积异常收益率，无论上调存准率，还是下调存准率，上证指数均处于全天下跌的状态中，并且在大部分时段都显著小于零。
 - 在牛市行情中，可能出现存准率上调反而导致全天收益率(日收益率)下跌反涨的“异常现象”。



RESSET高频数据的学术应用

《存款准备金率调整对股票市场的影响——基于中国股市高频数据的实证研究》

■ 研究结论

2. 此外，无论是牛市还是熊市，存准率上调都会导致房地产行业的指数收益率在大部分时段显著为负，但对其他行业指数收益率的影响则不明显。总体而言，存款准备金率调整虽然短期内向市场准确地传递了信号，但受市场行情的影响，长期来看并未对股票市场造成实质性影响。

问题：如何批量下载多个股票代码的高频数据

回答：假设，您想要下载数据集中hf2018.class_2018所列示所有股票的五档行情数据集，数据集内容如下所示。使用的列包括Exchflg Code ClassCd，分别代表交易所编号（上交所1，深交所2）、股票代码、证券类型（1代表股票）。您可以使用如下操作。

	ISIN	Exchflg	Code	code_mkt	SecuAbbr	ClassCd	Class	SubClass
1	CNE000000040	2	000001	000001.SZ	平安银行	1	Stk	A Share
2	CNE0000000T2	2	000002	000002.SZ	万科A	1	Stk	A Share
3	CNE0000000Y2	2	000004	000004.SZ	国农科技	1	Stk	A Share
4	CNE0000001L7	2	000005	000005.SZ	世纪星源	1	Stk	A Share
5	CNE000000164	2	000006	000006.SZ	深振业A	1	Stk	A Share
6	CNE0000000P0	2	000007	000007.SZ	全新好	1	Stk	A Share
7	CNE0000001C6	2	000008	000008.SZ	神州高铁	1	Stk	A Share
8	CNE000000222	2	000009	000009.SZ	中国宝安	1	Stk	A Share
9	CNE000000156	2	000010	000010.SZ	*ST 美丽	1	Stk	A Share
10	CNE000000248	2	000011	000011.SZ	深物业A	1	Stk	A Share
11	CNE0000002R2	2	000012	000012.SZ	南玻A	1	Stk	A Share
12	CNE0000003G3	2	000014	000014.SZ	沙河股份	1	Stk	A Share
13	CNE0000002T8	2	000016	000016.SZ	深康佳A	1	Stk	A Share
14	CNE0000002Q4	2	000017	000017.SZ	深中华A	1	Stk	A Share
15	CNE0000002S0	2	000018	000018.SZ	*ST 神城	1	Stk	A Share
16	CNE0000002V4	2	000020	000020.SZ	深华发A	1	Stk	A Share
17	CNE000000FK4	2	000021	000021.SZ	深科技	1	Stk	A Share
18	CNE000000834	2	000023	000023.SZ	深天地A	1	Stk	A Share
19	CNE0000008N8	2	000025	000025.SZ	特力A	1	Stk	A Share
20	CNE000000875	2	000026	000026.SZ	飞亚达A	1	Stk	A Share
21	CNE000000933	2	000027	000027.SZ	深圳能源	1	Stk	A Share
22	CNE0000009N6	2	000028	000028.SZ	国药一致	1	Stk	A Share
23	CNE0000009Z0	2	000030	000030.SZ	富奥股份	1	Stk	A Share

以下程序分股票下载2018年1月27日至2018年3月3日的五档行情SNAP数据集，并按股票保存。

```
filename tmp temp;
data _null_;
file tmp;
set hf2018.class_2018 (obs=3);
where ClassCd = 1;
if exchflg = 1 then mkt = 'SH';
else mkt = 'SZ';
cmd = compress('%get_data_l1(' || code || ", '27Jan2018'd, '03Mar2018'd, " ||
              mkt || ", stk, ,D:/pytest/stk" || code || ".csv,csv,);");
put cmd;
run;

%include tmp;
filename tmp clear;
```

同样，大家还可以使用我们提供的SAS代码生成工具

Q:请问Level 1数据库的样本更新规则是怎样的？是每3秒更新一条盘口快照数据还是tick-level（以每笔交易和订单报撤的变化为准）？

A:Level 1 数据为快照数据，不是tick数据。Level 2 的SNAP表数据也是快照数据。

Q:请问这些订单的买卖方向是如何判定的？

A：“买卖方向”指标的计算方法在数据词典中有说明。现摘录如下：

Trdirec 买卖方向计算方法（根据 Lee and Ready(1991)）：

第一步，如果 成交价 $>0.5*($ 上一条观测买一价+上一条观测卖一价)，则为买方主动成交；如果 成交价 $<0.5*($ 上一条观测买一价+上一条观测卖一价)，则为卖方主动成交；如果 成交价 $=0.5*($ 上一条观测买一价+上一条观测卖一价)，则进行第二步。

第二步，如果 成交价 $>0.5*($ 向前第二条观测买一价+向前第二条观测卖一价)，则为买方主动成交；如果 成交价 $<0.5*($ 向前第二条观测买一价+向前第二条观测卖一价)，则为卖方主动成交；如果 成交价 $=0.5*($ 向前第二条观测买一价+向前第二条观测卖一价)，则进行第三步。

第三步，如果 成交价 $>0.5*($ 向前第三条观测买一价+向前第三条观测卖一价)，则为买方主动成交；如果 成交价 $<0.5*($ 向前第三条观测买一价+向前第三条观测卖一价)，则为卖方主动成交；如果 成交价 $=0.5*($ 向前第三条观测买一价+向前第三条观测卖一价)，则判定为无法确定买卖方向。

Q:是否有方法根据锐思数据库的逐笔交易数据恢复订单簿信息，比如区分市价订单和限价订单呢？

A:逐笔交易数据需要使用沪深高频增强版的TRADE表格数据，恢复订单簿以及区分市价订单和限价订单需要使用沪深高频增强版的ORDER表格数据。

高频数据分享

END

Q & A

也可在交流群中提问