

Running SAS on the WRDS Cloud to obtain Daily TAQ or Monthly TAQ Data

F635 Market Microstructure Assignment 2: Due Next Class

Overview: The purpose of this assignment is to get you familiar with running SAS programs on the WRDS Cloud that: (1) access Daily TAQ or Monthly TAQ data, (2) clean the data, (3) compute the National Best Bid and Offer (NBBO), and (4) computing standard liquidity measures. It is much faster to run such programs on the WRDS Cloud, because this avoids downloading large raw data files. The WRDS servers are now fast and have the capacity to run big jobs. The alternative is PC-based SAS-CONNECT programs. This approach takes longer to run and ties-up you PC, but they may give you extra control and flexibility. The approach described below has you edit SAS code on the PC because this is a more user-friendly environment and then run the SAS code on the WRDS Cloud because this is more efficient.

Monthly TAQ spans 1/1/1993 to 12/31/2014 with trades and quotes timestamped to the second. Daily TAQ spans 9/10/2003 to the present with trades, quotes, and NBBOs timestamped to the millisecond (one-thousandth of a second) until mid-2015 and timestamped to the microsecond (one-millionth of a second) since then.

Needed code and programs:

1. Obtain the Holden and Jacobsen Daily TAQ and Monthly TAQ SAS code. Go to my web site www.kelley.iu.edu/cholden, click on the link that says **Code**, go to your PC's **Downloads** folder, right-click on the file "**Holden-and-Jacobsen-Daily-TAQ-and-Monthly-TAQ-Code-2016-07-21.zip**", click on **Extract All**, click on **Extract**, and you will obtain four versions of our DTAQ and MTAQ SAS code:
 - Daily TAQ version that runs in the WRDS Cloud
 - Monthly TAQ version that runs in the WRDS Cloud
 - Daily TAQ version that downloads raw data from WRDS via SAS-CONNECT and then runs on a PC
 - Monthly TAQ version that download raw data from WRDS via SAS-CONNECT and runs on a PC.
2. Download and install the free, open-source programs "**PuTTY**" and "**WinSCP**." At IU, both programs are available at **IUWare** (www.iuware.iu.edu/). Alternatively, you can obtain PuTTY at <http://www.putty.org/> and WinSCP at <https://winscp.net/eng/download.php>.

Steps:

1. **Configure PuTTY and then login to WRDS Cloud.** See Appendix 1 for step-by-step instructions.
2. **Configure WinSCP and then login to WRDS Cloud.** See Appendix 2 for step-by-step instructions.
3. **Make Changes to the Daily TAQ Code and Then Run It.** Launch SAS on your PC, open the Holden-and-Jacobsen-Daily-TAQ-WRDS-Cloud-Based-Code-2016-07-21.sas, and make the following changes:
 - **Adapt to your WRDS account.** The first line of code is:

```
libname project '/home/institution/username';
```

Replace “institution” with your WRDS institution and “username” with your WRDS username.

- **Select dates.** There are separate data files for each day, because they are so large. Below the **data** DailyNBBO step, there are requested data files in the “nbbo.nbbom_YYYYMMDD” format:

```
set nbbo.nbbom_20150727 nbbo.nbbom_20150728;
```

 Similarly, below the **data** DailyQuote step, there are requested data files in the “cq.cqm_YYYYMMDD” format:

```
set cq.cqm_20150727 cq.cqm_20150728;
```

 Similarly, Below the **data** DailyTrade step, there are requested data files in the format “ct.ctm_YYYYMMDD” format:

```
set ct.ctm_20150727 ct.ctm_20150728;
```

 Change the given YYYYMMDD dates to other dates that are trading days. Be sure to avoid weekends and holidays! Optionally add requested data files for additional trading days in all three locations using the corresponding format.
- **Select firms.** There is a list of requested stock symbols in three locations (below **data** DailyNBBO, below **data** DailyQuote, and below **data** DailyTrade):

```
where sym_root in ('AAPL', 'IBM')
```

 Change ‘AAPL’ to the stock symbol for some other NASDAQ firm and change ‘IBM’ to the stock symbol for some other NYSE firm in all three locations.
- **In SAS, save the updated file.**

4. **Upload the SAS file and run it.** In WinSCP, drag the updated SAS file from your local computer to your WRDS account folder. In PuTTY, type “LS” and you will see your SAS file. Also in PuTTY, type “**nohup qsas & yourfile.sas**”. To explain, “nohup” prevents the command from being aborted if you log out or exit the shell, “qsas” submits the sas job, and “&” runs the job in the background which returns you to the command prompt while it is running. You will see a notification that you job has been submitted and a job number. If you type “LS” again, you will see a new SAS log file “yourfile.log”. In WinSCP, open and view this file to make sure that everything ran correctly.
5. **Examine the results.** Once the job has finished running, type “LS” and you will see three new output files containing standard liquidity measures:
 - “QuotedSpreadsandDepths” contains Quoted Spreads and Depths
 - “EffectiveSpreads” contains Effective Spreads
 - “RealizedSpreadsandPriceImpacts” contains Realized Spreads and Price Impacts that are aggregated based three conventions:
Ave = simple average, DW = dollar-weighted, SW = share-weightedIn WinSCP, drag these three files from your WRDS account folder to your local computer and examine them in PC SAS.

F635 students: To verify your success in doing Assignment 2, please email the file “EffectiveSpreads” to me.

Additional Notes:

1. **The Monthly TAQ WRDS Cloud code works in the same manner as the Daily TAQ WRDS Cloud code explained above.**
2. **Consider examining the raw data files and intermediate data files.** At the very end of the Holden and Jacobsen code are commented out lines to copy six other files from the SAS temporary WORK folder to your WRDS account file. All six are relatively large files. If you remove the comment outs, then you will get the following:

Three raw data files containing DTAQ data for the firms and dates you selected:

"DailyNBBO" contains NBBO data

"DailyQuote" contains quote data

"DailyTrade" contains trade data

Importantly, the "DailyNBBO" file does NOT contain the complete NBBO. When one exchange has both the best bid and best offer it is only noted in the "DailyQuote" file, not the "DailyNBBO" file. Our code combines data from both files to construct the official complete NBBO (see file below).

Three intermediate data files:

"OfficialCompleteNBBO" contains the official complete NBBO

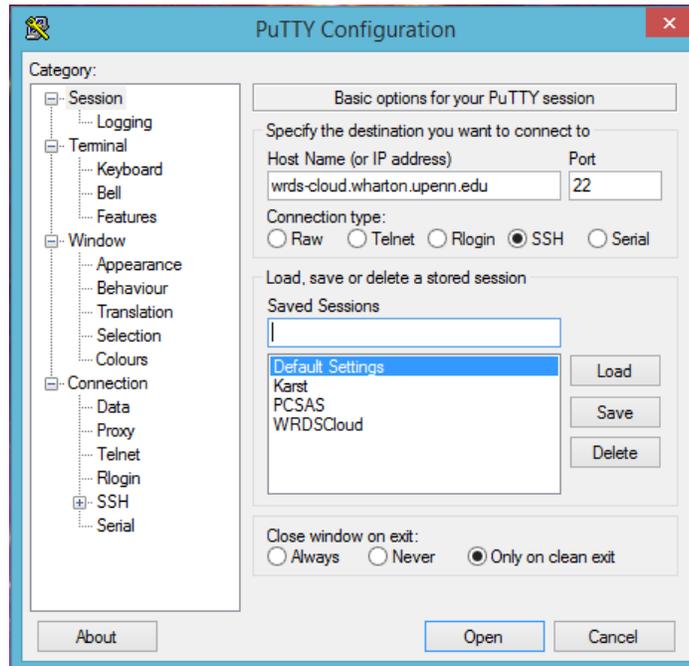
"TradesPriorNBBOandQuotes" contains trades, prior
microsecond NBBO, and quotes

"BuySellIndicators" adds buy/sell indicators based on three
conventions: LR = Lee & Ready (1991), EMO = Ellis, Michaely &
O'Hara (2000), CLNV = Chakrabarty, Li, Nguyen, & Van Ness (2006)

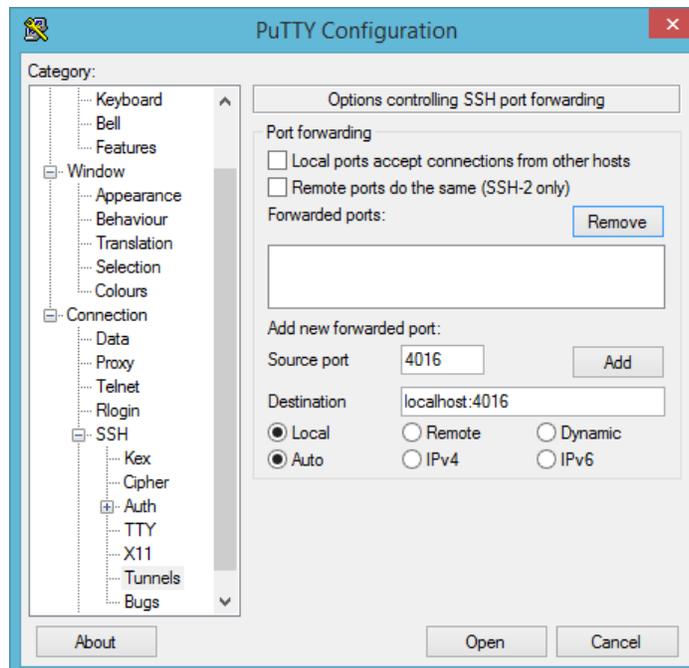
3. **What to do if your program runs out of space in the WRDS Cloud.** TAQ files are relatively large. If you select a large number of firms and long time period, then it is possible for SAS to run out of available work space in the WRDS Cloud and your program will crash. If this happens to you, cut it into a series of smaller programs. For example, you might divide a program using a year of data into four programs using a quarter worth of data each. Similarly, a program using a quarter of data might be divided into three programs using a month worth of data each. The number of trades and quotes is growing about 30% per year compounded year after year. So a program using recent data will have much bigger files than a program using older data.

Appendix 1: How to Configure and Run PuTTY to Connect to WRDS

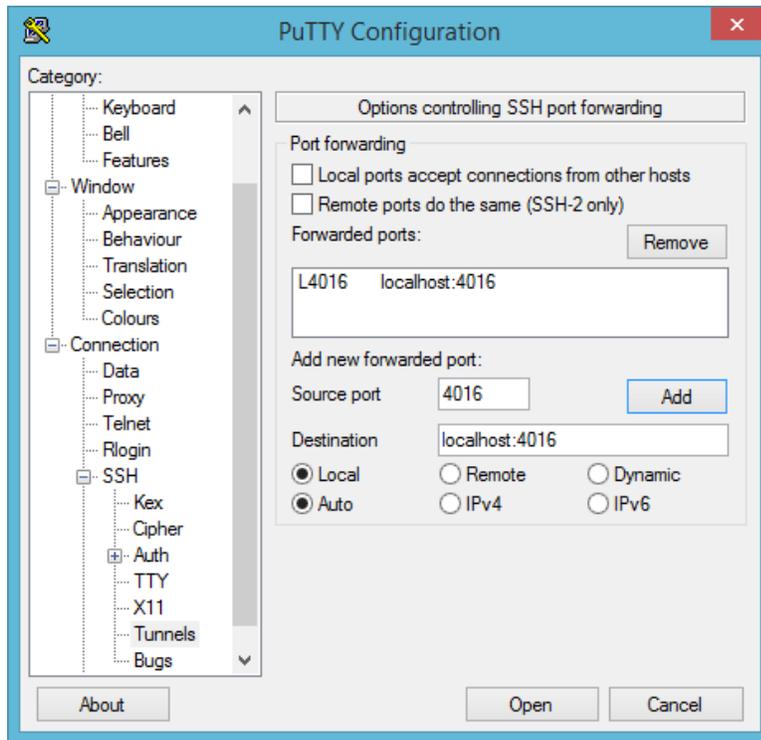
1. After you have downloaded and installed **PuTTY**, then:
 - a. Launch **PuTTY**
 - b. Type in the Host Name, **wrds-cloud.wharton.upenn.edu**
Default port should be **22**



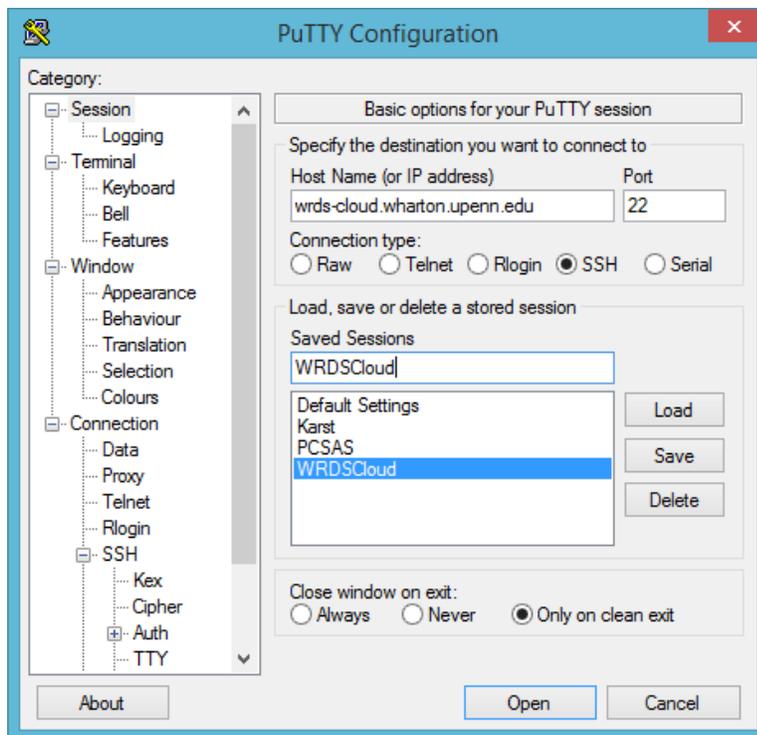
- c. Select **Connection->SSH->Tunnels**
 - d. Enter Source Port **4016** and Destination **localhost:4016**



- e. Click “Add”; It should now look like this:

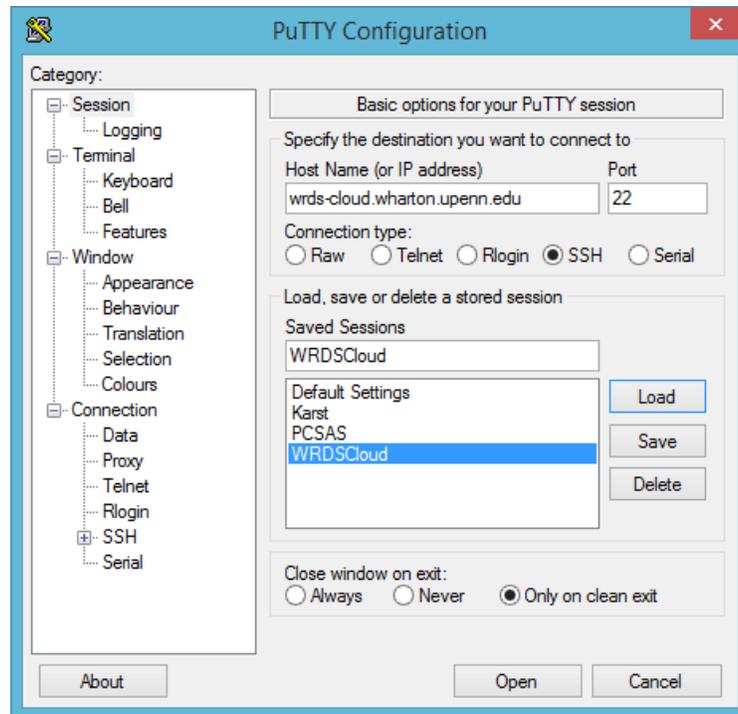


- f. Go back to **Session**. Enter **WRDSCloud** (or whatever you want) under **Saved Sessions**

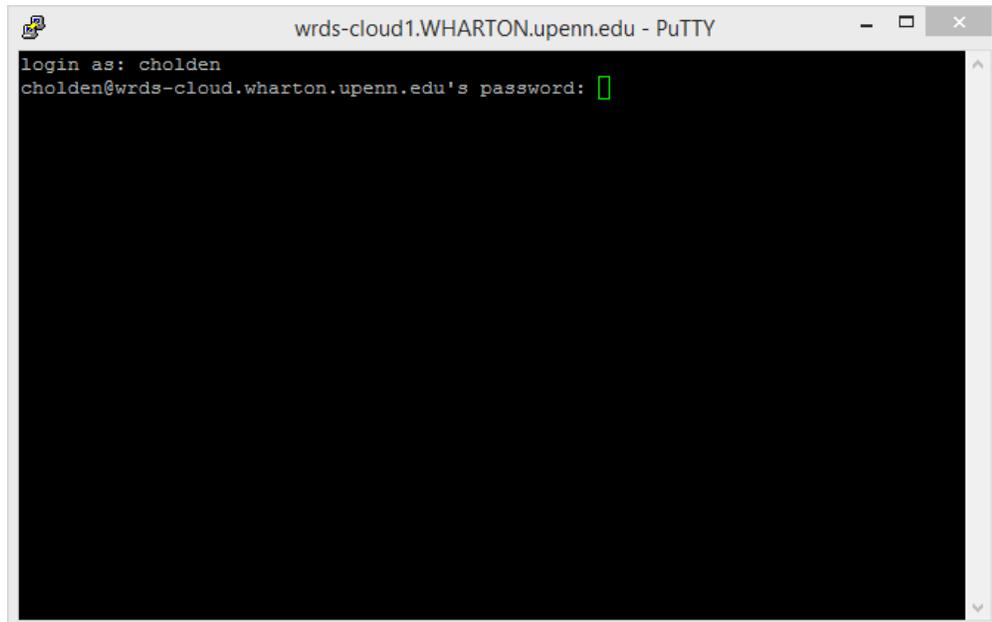


Click “Save”. You can now select a **WRDSCloud** session from **Saved Sessions** in the future.

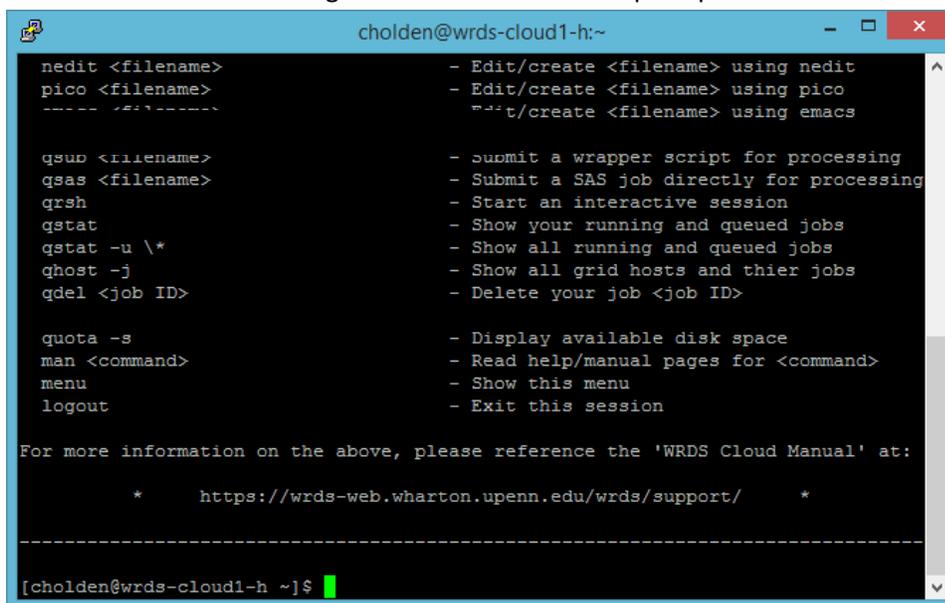
2. Launch **PuTTY**,
 - a. click “**WRDSCloud**” then “**Load**” and then “**Open**” (or else just double-click “**WRDSCloud**”) to open a new session.



- b. When you open a session, it will say: “login as:”. Type your WRDS username and press enter. Then it will ask for your password. Type your WRDS password and press enter.



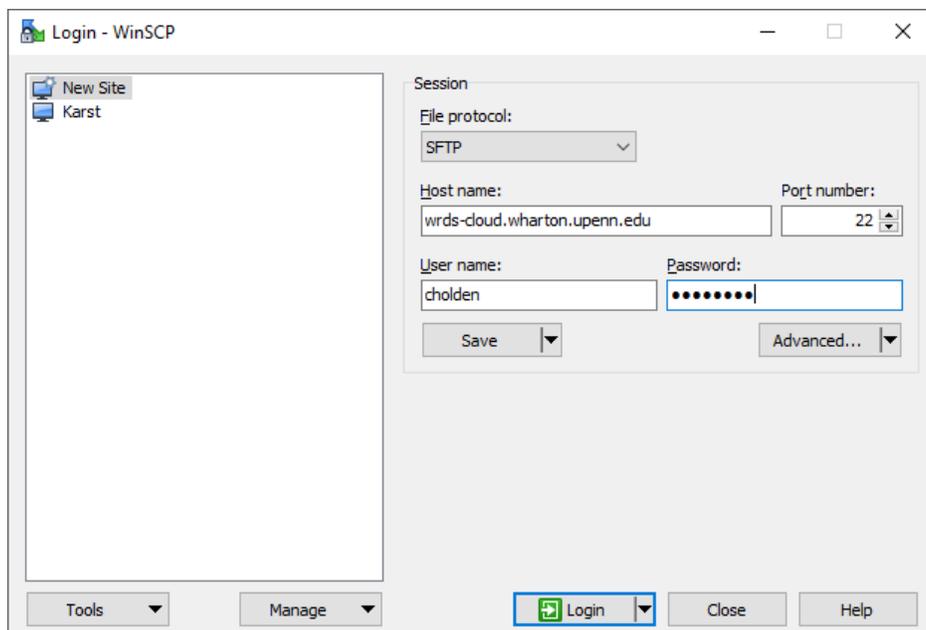
An established connection generates a “wrds-cloud” prompt that looks like this:



```
cholden@wrds-cloud1-h:~  
nedit <filename>          - Edit/create <filename> using nedit  
pico <filename>          - Edit/create <filename> using pico  
-----  
edit/create <filename> using emacs  
  
qsub <filename>          - Submit a wrapper script for processing  
qsas <filename>          - Submit a SAS job directly for processing  
qssh                    - Start an interactive session  
qstat                   - Show your running and queued jobs  
qstat -u \*             - Show all running and queued jobs  
qhost -j                - Show all grid hosts and thier jobs  
qdel <job ID>           - Delete your job <job ID>  
  
quota -s                - Display available disk space  
man <command>           - Read help/manual pages for <command>  
menu                    - Show this menu  
logout                  - Exit this session  
  
For more information on the above, please reference the 'WRDS Cloud Manual' at:  
  
*      https://wrds-web.wharton.upenn.edu/wrds/support/      *  
-----  
[cholden@wrds-cloud1-h ~]$
```

Appendix 2: How to Configure and Run WinSCP to Connect to WRDS

1. After you have downloaded and installed WinSCP, then:
 - a. Launch **WinSCP**
 - b. Click on the **New Site** button in the upper left column
 - c. Type in the Host name, **wrds-cloud.wharton.upenn.edu**
 - d. Enter your WRDS user name
 - e. Enter your WRDS password
 - f. Click on **Save** and accept the default name or enter new name for it



2. Launch **WinSCP**,

- a. Double-click the “**username@wrds-cloud.wharton.upenn.edu**” button (or whatever you named it) in the left column to login
- b. Confirm that this is what you want to do
- c. Enter your WRDS password
- d. A WinSCP session will open with two windows: (1) the local folder in one window and (2) the host folder in the other window – thus making it very easy to transfer files in either direction

