

HW 2.

Background information

We are interested in the relation between the average bid-ask spreads on stocks and the characteristics of the corresponding companies.

The data file lab2.xls contains information for the 100 stocks in the S&P 100 index.

Our variable of interest (the Y variable) is the bid-ask spread (constructed as an average over the day) - or tradecost - of the

S&P100 stocks. The explanatory, or X , variables are:

1. log volatility - The log of the daily return standard deviation
2. log size - The log of the size of the stock. Size is total outstanding
3. number of shares multiplied by share price. Size is measured in thousands of dollars
4. log trades - This is the log of the average number of trades per day
5. log turn - This is the log of the ratio of the average number of shares traded per day (in dollars) over the number of shares outstanding (in dollars)
6. NumberAnalysts - This is the number of analysts following the stock

Consider the following theories of the determinants of bid-ask spreads.

1. {Asymmetric information}. Stocks with greater degrees of asymmetry in information (regarding their fundamental value) tend to have wider bid-ask spreads. The number of analysts following a stock is viewed as an asymmetric information proxy. The larger it is, the lower private information, the smaller the spreads. Log turn-over is, also, seen as an asymmetric information proxy. The larger it is, the larger private information, the larger the spreads. Trading rates in excess of this proportion should be associated with informed trading.)

2. {Liquidity}. Stocks that trade more frequently and have larger market capitalization (i.e., more liquid stocks) tend to have lower bid-ask spreads. The larger log trades and log size, the larger liquidity, the smaller the spreads. Log turn-over is, also, sometimes seen as a liquidity proxy. The larger it is, the larger liquidity, the smaller the spreads.

3. {Fundamental volatility}. Stocks that have a higher volatility of fundamental values tend to have larger bid-ask spreads. Higher uncertainty about the underlying stock's value implies higher potential for adverse price moves and, hence, higher inventory risk, mostly in the presence of large imbalances to offset.

The greater the degree of asymmetric information, the wider the spreads should be as the market makers (who are not fully informed) charge a higher price when selling (raising the ask) and a lower price when buying(lowering the bid) to insulate themselves from losing money to potentially better informed traders. We do not get to see the degree of asymmetric information between market participants but we do know the following:

1) The larger the number of analysts following a stock, the lower asymmetric information we expect to be for that stock. Analysts provide information about the stock, thereby uncovering its fundamental value.

2) The greater the turn-over of the stock, the higher we expect asymmetric information to be for that stock. Intuitively, the greater the turnover, the faster individuals are getting in and out of investment positions in the stock. They do this more when they believe the current price does not accurately reflect the fundamental value of the stock - that is, when they believe they possess asymmetric (superior) information.

As indicated, turn-over is also viewed, by some, as a liquidity proxy.

*****8

Instructions:

- You will answer a set of questions that will require to analyze data and run/set-up multivariate regressions.
- The file is organized in block, with text setting up a question, then an area where you insert your code, and then a text area where you insert your comments and answers
- Once you complete the analysis, you will submit a .pdf file that reports your code, your results, and your answers to the questions (marked with Q).

Q1 (2 points). Import the data set and set up variable names

```
% import dataset lab2.xls (using Import Dataset)
```

Q2. (2 points) Generate a histogram of the bid-ask spread and the log of the bid-ask spread (tradedcost).

What are their differences?

Answer:

Q3.(5 points) Run a regression of the log of the bid-ask spread on the 5 explanatory variables.

Interpret the sign of the estimated coefficients of the 5 explanatory variables in light of the theories outlined in the background information.

Answer:

Q4. (4 points) Organize the data in matrix format, and compute the regression coefficients using the OLS formulas for these coefficients in matrix form. Check that you obtain the same results of those obtained with the fitlm command.

Answer:
