# 1 Linear-Log Model

Intuitively, I think the relationship between the stock return volatility and the open interest of its options is linear-log because options outstanding are different by thousands or even millions across companies, but their returns' volatilities are relatively close.

Table 1: Linear-log Model Regression Results

| | Estimate | Std. Error | t value | Pr(>\|t\|) |
|---|---|---|---|---|
| (Intercept) | 0.0001 | 0.0004 | 0.28 | 0.7826 |
| data$ln_delta_30_total | -0.0000 | 0.0000 | -0.63 | 0.5277 |
| data$monthlychangeinVIX | -0.0001 | 0.0000 | -4.37 | 0.0000*** |
| data$changeinimpl_volatility_averag | 0.0052 | 0.0018 | 2.97 | 0.0030** |

From the table, we can see, first, opposite to my intuitition, the relationship between the stock return volatility and the open interest of its options is not log-linear because the the coefficient of the log change of the open interest of all options is insignificant. Second, the coefficient of the control variables, change of VIX, a proxy of market volatility, and the coefficient of change in average implied volatility of the stock's options, a proxy of idosyncratic risk, are both significant. Note that, there is multicolinearity between the two variables because in practice, implied volatility is effected by the market volatility. Thus, the change in VIX variable very likely picks up the varaition of the change in implied volatility variable thus the latter is less significant. Lastly, the insignificance of the intercept makes sense because it does not make sense that there should not be any volatility cannot be explained by the market and the individual volatility. Then, I move to linear-linear model.

Table 2: Linear-linear Model Regression Results

| | Estimate | Std. Error | t value | Pr(>\|t\|) |
|---|---|---|---|---|
| (Intercept) | -0.0001 | 0.0001 | -1.31 | 0.1903 |
| data$changeinTotal | -0.0000 | 0.0000 | -6.15 | 0.0000*** |
| data$monthlychangeinVIX | -0.0001 | 0.0000 | -4.25 | 0.0000*** |
| data$changeinimpl_volatility_averag | 0.0063 | 0.0018 | 3.58 | 0.0004*** |

Table 2 shows the result of linear-linear model. The coefficients of all independent variables are significant. However, the coefficient of data$changeinTotal

is very small. This is because the amount of option open interest is usually in millions whereas volatility is in percentage. If I multiply the coefficient by $10^6$, the coefficient becomes 0.00167. Since most of option contracts do not contain information. This explains why the coefficient is still small after I adjust the difference the order of magnitude of option open interest and volatility. However, different firms have different characteristics in terms of both volatility and option open interest. In detail, firms in different sectors have different level of risk, and popular firms have more option contracts out standing comparing to less known firms. Thus, it makes sense to use a panel data model with fixed effect– this model will tell me if the relationship between option open interest and the underlying stock return volatility is significant for each firm in my sample.

Table 3: Panel Regression Results

| | *Dependent variable:* |
|---|---|
| | changeinNext30daysVolatilit |
| changeinTotal | -0.000*** |
| | (0.000) |
| monthlychangeinVIX | -0.0001*** |
| | (0.00002) |
| changeinimpl_volatility_averag | 0.006*** |
| | (0.002) |
| Observations | 2,886 |
| $R^2$ | 0.020 |
| Adjusted $R^2$ | 0.002 |
| F Statistic | 18.857*** (df = 3; 2834) |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 |

Table 3 displays the result of fixed effect panel data regression result. The significant results confirm that the relationship between the stock return volatility and the open interest of its options is significant within each firm, controlling for the market and individual volatility. And the F-statistic shows

that the variables are jointly significant. However, a concerning fact is that the $R_{sqaure}$ is 0.02, which is too low.