1. **Load the Raw Data**

Note: This is panel data. The statistical software must be told it is a panel data set.

Variables will have the notation X\_it with i being the specific firm and t being the year.

1. **Create Regression Variables of differenced lag K (flexible)**
   1. Create dependent variable, either market (ML) or book leverage (BL), of differenced lag k (Name Variable D\_ML)

*This is done by taking ML\_it - ML\_it - k. For example if k = 4 then the differenced variable is ML\_i6 - MLi2*

* 1. Create Independent Variables (Ind. V) of differenced lag k

This is done by taking X\_it-1 - X\_it-k-1. For example, if k = 4 and X\_it is X\_i6 then the differenced variable is X\_i5 - X\_i1 (Where X represents the independent variables)

Create the following independent variables in the same manner:D\_ signifies difference in variable name.

* + 1. Q: Q\_it-1 - Q\_it-k-1 (Name Variable D\_Q\_it-1)
    2. Research & Development: RnD\_it-1 - RnD\_it-k-1 (Name Variable D\_RnD)
    3. Capital Expenditures: CAPEX\_it-1 - CAPEX\_it-k-1 (Name Variable D\_CAPEX)
    4. Net Sales: SALE\_it-1 - SALE\_it-k-1 (Name Variable D\_SALE)
    5. Operating Income Before Depreciation: OIBD\_it-1 - OIBD\_it-k-1 (Name Variable D\_OIBD)
    6. Property Plant & Equipment: TANG\_it-1 - TANG\_it-k-1 (Name D\_TANG)

The following variables are macroeconomic variables. They are the same for each year for each company, hence there is only a t in the notation for the year and not an i.

* + 1. Equity Risk Premium: ERP\_t-1 - ERP\_t-k-1 (D\_ERP)
    2. Real Interest Rate: RIR\_t-1 - RIR\_t-k-1 (D\_RIR)
    3. Default Spread: DSP\_t-1 - DSP\_t-k-1 (D\_DSP)
    4. Term Spread: TSP\_t-1 - TSP\_t-k-1 (D\_TSP)

The following two macroeconomic variables start at t

* + 1. Statutory Corporate Tax Rate: TAXR\_t - TAXR\_t-1 (D\_TAXR)
    2. Real Gross Domestic Product: RGDP\_t - RGDP\_t-1 (D\_GDP)

The following are Dummy Variables and are not differenced, but simply matched with the correct year t

* + 1. Research & Development Dummy: Name RnDD
    2. Economic Recession Dummy: Name ERD
    3. Great Economic Recession Dummy (Financial Crisis): Name GERD

1. **Make sure the variables are matched to the correct year for the regressions.**

For example if k = 4 and t = 1970 then the variables with X\_t - X\_t-1 would be X\_1970 - X\_1966 and the variables X\_t-1 - X\_t-k-1 would be X\_1969 - X\_1965.

ML\_it-1 - ML\_it-k-1 is correlated with the error terms. It must be replaced by an instrumental variable. ML\_it-k-1 is a valid instrument for ML\_it-1.

1. **Using pooled OLS run the following regression: ML\_it-1=B\_0+B\_1 ML\_it-k-1+ε\_it**
2. **Use B\_0 and B\_1 to estimate new ML\_it-1 that are not correlated with the error term.**

This part is not a regression, but addition and multiplication of the following: β\_0+β\_1 ML\_it-k-1+ε\_it which produce new ML\_it-1.

1. **Call this new variable 1ML\_it-1**
2. **Take the long difference 1ML\_it-1 - 1ML\_it-k-1 to create Instrumental Variable 1 (Name variable IV1)**

Run the full long differencing regression for the first time to get the coefficients necessary to produce a better instrumental variable

1. **Run the following Regression using pooled OLS:**

D\_ML = lamda(IV1) + B\_1(D\_Q) + B\_2(D\_RnDD) + B\_3(D\_RnD) + B\_4(D\_CAPEX) + B\_5(D\_SALE) + B\_6(D\_OIBD) + B\_7(D\_TANG)+ B\_8(D\_ERP) + B\_9(D\_DSP) + B\_10(D\_TSP) + B\_11(D\_TAXR) + B\_12(D\_RGDP) + B\_13(ERD) + B\_14(GERD) + error term

1. **Save all the coefficients from the regression (lambda(hat), B\_0(hat), B\_1(hat) ... B\_14(hat)) to be used in the second instrumental variable regression.**
2. **Estimate a new and better instrumental varible (IV) for the leverage ratio ML** This better instrumental variable will be known as the second instrumental variable (IV2)

The valid instruments for ML\_it-1 are: ML\_it-k-1, L\_it-k-lambda(hat)L\_it-k-1 ...

1. **Simplify the following residuals into one number –**

**R1** = ML\_it-1 – lambda\_hat(ML\_it-2) – B\_1\_hat(Q\_it-2) + B\_2(RnDD\_it-2) + B\_3(RnD\_it-2) + B\_4(CAPEX\_it-2) + B\_5(SALE\_it-2) + B\_6(OIBD\_it-2) + B\_7(TANG\_it-2)+ B\_8(ERP\_it-2) + B\_9(DSP\_it-2) + B\_10(TSP\_it-2) + B\_11(TAXR\_it-2) + B\_12(RGDP\_it-2) + B\_13(ERD\_it-2) + B\_14(GERD\_it-2)

If k = 4, then go until R4, which would be

**R1** = ML\_it-4 – lambda\_hat(ML\_it-5) – B\_1\_hat(Q\_it-5) +…

1. **Run the following regression using pooled OLS:**

ML\_it-1=B\_0+B\_1 ML\_it-k-1 + B\_2(R1) + B\_3(R2) + B\_4(R3) + B\_5(R4)+ ε\_it

1. **Use the coefficients estimated in the regression make the new ML\_it-1. Again addition, subtraction, and multiplication are used not a regression.**

The new leverage ratio variable is called 2ML\_it-1

1. **Take the long difference 2ML\_it-1 - 2ML\_it-k-1 to create instrumental variable 2 (Name variable IV2)**
2. **Run the full long differencing regression for the second time using pooled OLS and IV2**

D\_ML = lamda(IV2) + B\_1(D\_Q) + B\_2(D\_RnDD) + B\_3(D\_RnD) + B\_4(D\_CAPEX) + B\_5(D\_SALE) + B\_6(D\_OIBD) + B\_7(D\_TANG) + B\_8(D\_ERP) + B\_9(D\_DSP) + B\_10(D\_TSP) + B\_11(D\_TAXR) + B\_12(D\_RGDP) + B\_13(ERD) + B\_14(GERD) + error term

**End of 1st iteration**

1. **Save all the estimated coefficients and name them with an underscore of 1 to signify they are the end of the first iteration.**

Example: lambda(hat)\_1.

1. **Estimate a new and better IV for ML. This will be called IV3.**

The coefficients from the end of the 1st iteration will be used in the residuals.

1. **Simplify the following residuals into one number –**

**R1** = ML\_it-1 – lambda\_hat(ML\_it-2) – B\_1\_hat(Q\_it-2) + B\_2(RnDD\_it-2) + B\_3(RnD\_it-2) + B\_4(CAPEX\_it-2) + B\_5(SALE\_it-2) + B\_6(OIBD\_it-2) + B\_7(TANG\_it-2)+ B\_8(ERP\_it-2) + B\_9(DSP\_it-2) + B\_10(TSP\_it-2) + B\_11(TAXR\_it-2) + B\_12(RGDP\_it-2) + B\_13(ERD\_it-2) + B\_14(GERD\_it-2)

If k = 4, then go until R4, which would be

**R1** = ML\_it-4 – lambda\_hat(ML\_it-5) – B\_1\_hat(Q\_it-5) +…

1. **Run the following regression using pooled OLS:**

ML\_it-1=B\_0+B\_1 ML\_it-k-1 + B\_2(R1) + B\_3(R2) + B\_4(R3) + B\_5(R4)+ ε\_it

1. **Use the coefficients estimated in the regression make the new ML\_it-1.**

Again addition, subtraction, and multiplication are used not a regression.

The new leverage ratio variable is called 3ML\_it-1

1. **Take the long difference 3ML\_it-1 - 3ML\_it-k-1 to create instrumental variable 3 (Name variable IV3)**
2. **Run the full long differencing regression for the third time using pooled OLS and IV3**

# D\_ML = lamda(IV3) + B\_1(D\_Q) + B\_2(D\_RnDD) + B\_3(D\_RnD) + B\_4(D\_CAPEX) + B\_5(D\_SALE) + B\_6(D\_OIBD) + B\_7(D\_TANG)

# + B\_8(D\_ERP) + B\_9(D\_DSP) + B\_10(D\_TSP) + B\_11(D\_TAXR) + B\_12(D\_RGDP) + B\_13(ERD) + B\_14(GERD) + error term

**End of 2nd iteration**

1. **Save all the estimated coefficients and name them with an underscore of 2 to signify they are the end of the first iteration.**

Example: lambda(hat)\_2.

1. **Estimate a new and better IV for ML. This will be called IV4.**

The coefficients from the end of the 2nd iteration will be used in the residuals.

1. **Simplify the following residuals into one number –**

**R1** = ML\_it-1 – lambda\_hat(ML\_it-2) – B\_1\_hat(Q\_it-2) + B\_2(RnDD\_it-2) + B\_3(RnD\_it-2) + B\_4(CAPEX\_it-2) + B\_5(SALE\_it-2) + B\_6(OIBD\_it-2) + B\_7(TANG\_it-2)+ B\_8(ERP\_it-2) + B\_9(DSP\_it-2) + B\_10(TSP\_it-2) + B\_11(TAXR\_it-2) + B\_12(RGDP\_it-2) + B\_13(ERD\_it-2) + B\_14(GERD\_it-2)

If k = 4, then go until R4, which would be

**R1** = ML\_it-4 – lambda\_hat(ML\_it-5) – B\_1\_hat(Q\_it-5) +…

1. **Run the following regression usign pooled OLS:**

ML\_it-1=B\_0+B\_1 ML\_it-k-1 + B\_2(R1) + B\_3(R2) + B\_4(R3) + B\_5(R4)+ ε\_it

1. **Use the coefficients estimated in the regression make the new ML\_it-1.**

Again addition, subtraction, and multiplication are used not a regression.

The new leverage ratio variable is called 4ML\_it-1

1. **Take the long difference 4ML\_it-1 - 4ML\_it-k-1 to create instrumental variable 4 (Name variable IV4)**
2. **Run the full long differencing regression for the fourth time using pooled OLS and IV4**

D\_ML = lamda(IV4) + B\_1(D\_Q) + B\_2(D\_RnDD) + B\_3(D\_RnD) + B\_4(D\_CAPEX) + B\_5(D\_SALE) + B\_6(D\_OIBD) + B\_7(D\_TANG) + B\_8(D\_ERP) + B\_9(D\_DSP) + B\_10(D\_TSP) + B\_11(D\_TAXR) + B\_12(D\_RGDP) + B\_13(ERD) + B\_14(GERD) + error term

**End of 3rd iteration**

1. **Save all the estimated coefficients and name them with an underscore of 3 to signify they are the end of the first iteration.**

Example: lambda(hat)\_3.

Huang and Ritter used the 3rd iteration coefficients as their final results. We will also try using a larger number of iterations to see if there is any significant improvement or difference.