

EMBA Applied Regression Analysis - Final Project

Predicting IPOs Pricing Within or Above The Range

Richie Fleischman and AD King

July 9, 2016

Predicting IPO Pricing - Background

Project Idea

- ❑ Do certain variables predict whether IPOs will price below versus within / above the range


Hypothesis

- ❑ Our hypothesis is that a certain attributes of a Company outside of financial metrics drives investor sentiment and views on valuation

Experiment / Test

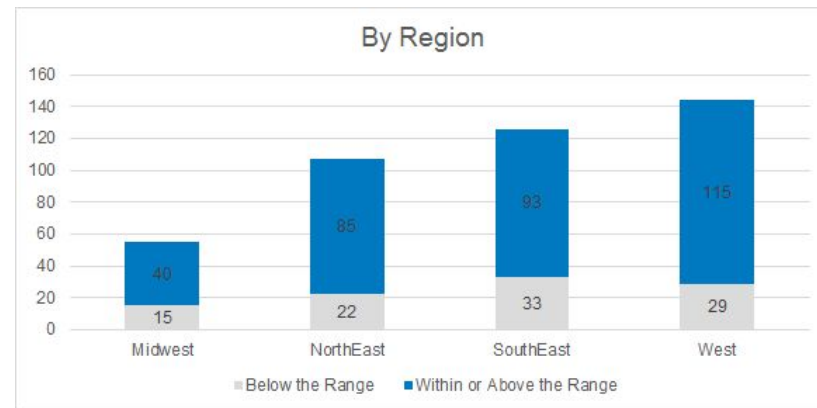
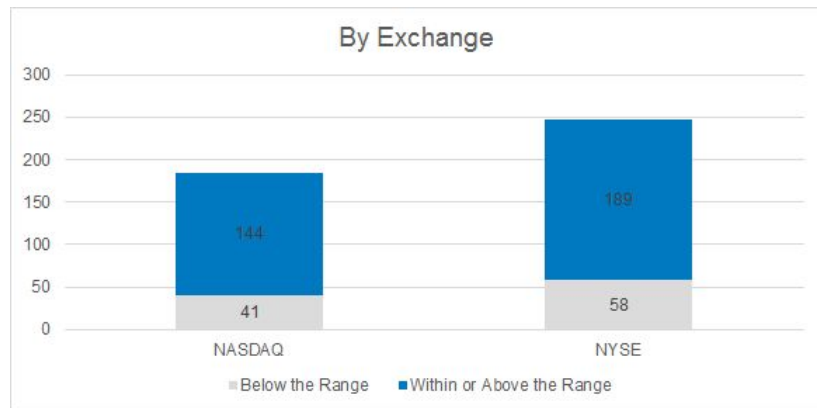
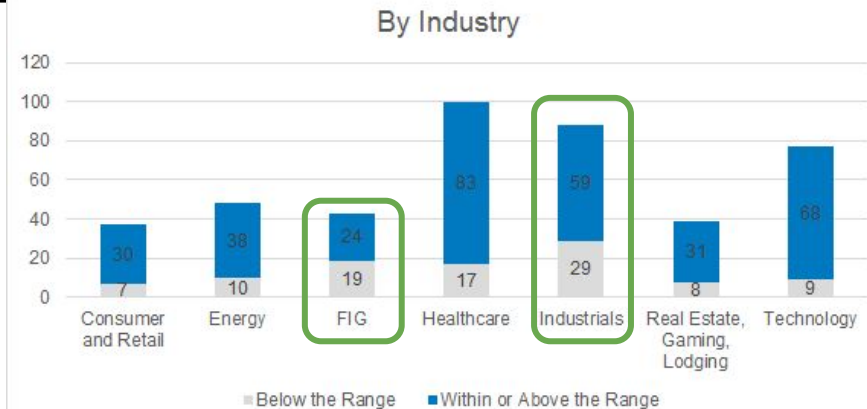
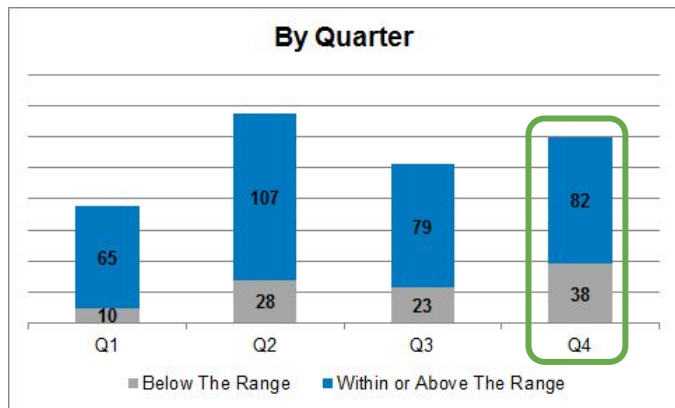
- ❑ Review the IPO pricing data for all deals completed in the past 5 years

Predicting IPO Pricing - Regression Model 1

Data Set	Obs.	Pricing Factors	Variables
Regression Model #1 <u>Other Data Constraint</u> <i>IPO Size > \$75 million</i>	433 	<input type="checkbox"/> Exchange <input type="checkbox"/> Date <input type="checkbox"/> Industry <input type="checkbox"/> Region of Headquarters	<input type="checkbox"/> Nasdaq (reference), NYSE <input type="checkbox"/> Q1 (reference), Q2, Q3, Q4 <input type="checkbox"/> Consumer & Retail (reference), FIG, Energy, Technology, Real Estate, Gaming Lodging, Energy, Healthcare, Industrials <input type="checkbox"/> Midwest (reference), Northeast, Southeast, West

We ran a binary logistic regression with the above observations and variables

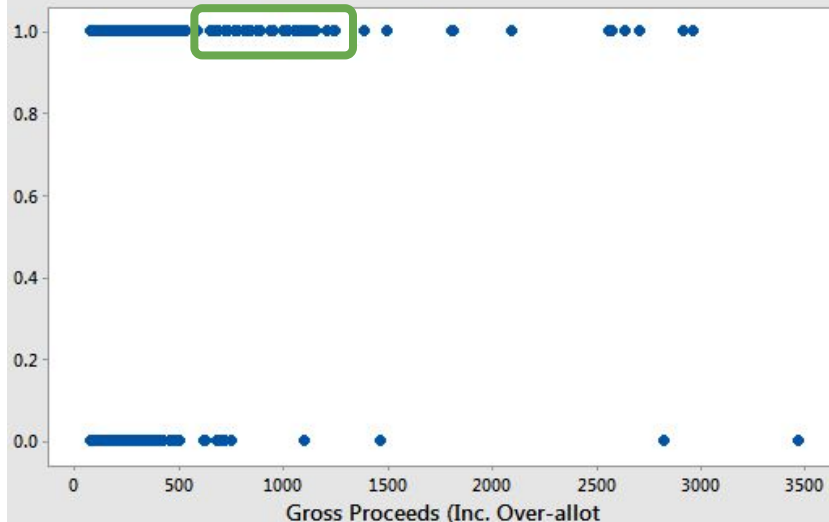
Predicting IPO Pricing - Model 1 Histograms



Predicting IPO Pricing - Regression Model 1 Outputs

Priced
within or
Above
Range

Scatterplot of Pricing Within Filing Ra vs Gross Proceeds (Inc. Ove



Priced
Below
Range

- There appears to be a some concentration in the amount of IPOs that price within or above the range in the \$600 million to \$1,500 million gross proceeds range

Predicting IPO Pricing - Regression Model 1 Outputs

Coefficients Table				Deviance Table					
Term	Coef	SE Coef	VIF	Source	DF	Adj Dev	Adj Mean	Chi-Square	P-Value
Constant	2.108	0.644		Regression	14	37.260	2.6615	37.26	0.001
Gross Proceeds (Inc. Over-allot	0.000414	0.000313	1.12	Gross Proceeds (Inc. Over-allot	1	1.998	1.9983	2.00	0.157
Q2				Q2	1	2.369	2.3688	2.37	0.124
1	-0.627	0.419	2.55	Q3	1	2.750	2.7505	2.75	0.097
Q3				Q4	1	10.853	10.8525	10.85	0.001
1	-0.707	0.436	2.41	NorthEast	1	0.200	0.1995	0.20	0.655
Q4				SouthEast	1	0.033	0.0331	0.03	0.856
1	-1.282	0.414	2.69	West	1	0.174	0.1737	0.17	0.677
NorthEast				Energy	1	0.106	0.1060	0.11	0.745
1	0.184	0.411	2.08	FIG	1	8.407	8.4070	8.41	0.004
SouthEast				Healthcare	1	0.001	0.0012	0.00	0.972
1	-0.071	0.392	2.33	Industrials	1	3.514	3.5142	3.51	0.061
West				Real Estate, Gaming, Lodging	1	0.331	0.3311	0.33	0.565
1	0.166	0.397	2.32	Technology	1	0.470	0.4703	0.47	0.493
Energy				NYSE	1	0.087	0.0874	0.09	0.768
1	-0.190	0.584	2.32	Error	417	427.803	1.0259		
FIG				Total	431	465.063			
1	-1.501	0.541	2.49						
Healthcare									
1	0.018	0.511	2.88						
Industrials									
1	-0.888	0.494	3.33						
Real Estate, Gaming, Lodging									
1	-0.346	0.603	2.03						
Technology									
1	0.393	0.570	2.25						
NYSE									
1	0.083	0.281	1.33						

Model Summary		
Deviance	Deviance	
R-Sq	R-Sq(adj)	AIC
8.01%	5.00%	457.80

Predicting IPO Pricing - Regression Model 1 Observations

Date

- ❑ Q4 has the lowest P-value amongst the dates and the highest negative coefficient suggesting that IPOs priced at the end of the year are more likely to price below the initial range

Industry

- ❑ FIG has the lowest P-value amongst the industries and highest negative coefficient suggesting FIG IPOs are more likely to price below the range


KKR-Backed First Data Dips in Debut After Below-Range IPO

October 15, 2015 – 10:15 AM EDT *Updated on* October 15, 2015 – 4:36 PM EDT

Conclusion

- ❑ With an Adjusted R-Squared of 5.0%, the model is not as predictive as we would hope. We removed Region and Exchange and re-ran the model as those were not as predictive

Predicting IPO Pricing - Regression Model 2

Data Set	Obs.	Pricing Factors	Variables
Regression Model #2 Other Data Constraint <i>IPO Size > \$75 million</i>	433 	<input type="checkbox"/> Date <input type="checkbox"/> Industry	<input type="checkbox"/> Q1 (reference), Q2, Q3, Q4 <input type="checkbox"/> Consumer & Retail (reference), FIG, Energy, Technology, Real Estate, Gaming Lodging, Energy, Healthcare, Industrials

Regression model #2 removes Exchange and Region Company Headquarters

Predicting IPO Pricing - Regression Model 2 Outputs

Coefficients Table				Deviance Table					
Term	Coef	SE Coef	VIF	Source	DF	Adj Dev	Adj Mean	Chi-Square	P-Value
Constant	2.213	0.563		Regression	10	36.350	3.6350	36.35	0.000
Gross Proceeds (Inc. Over-allot	0.000441	0.000308	1.08	Gross Proceeds (Inc. Over-allot	1	2.380	2.3805	2.38	0.123
Q2				Q2	1	2.439	2.4394	2.44	0.118
1	-0.633	0.417	2.53	Q3	1	2.897	2.8972	2.90	0.089
Q3				Q4	1	10.803	10.8035	10.80	0.001
1	-0.715	0.431	2.36	Energy	1	0.175	0.1747	0.17	0.676
Q4				FIG	1	8.462	8.4622	8.46	0.004
1	-1.273	0.412	2.67	Healthcare	1	0.006	0.0058	0.01	0.939
Energy				Industrials	1	3.559	3.5586	3.56	0.059
1	-0.235	0.564	2.17	Real Estate, Gaming, Lodging	1	0.355	0.3546	0.35	0.551
FIG				Technology	1	0.621	0.6212	0.62	0.431
1	-1.493	0.537	2.45	Error	421	428.713	1.0183		
Healthcare				Total	431	465.063			
1	0.038	0.504	2.81						
Industrials									
1	-0.878	0.486	3.24						
Real Estate, Gaming, Lodging									
1	-0.352	0.592	1.97						
Technology									
1	0.443	0.558	2.16						

Model Summary

Deviance

R-Sq

7.82%

Deviance

R-Sq(adj)

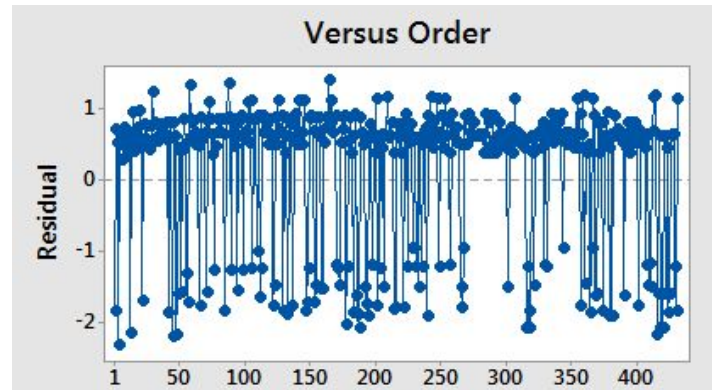
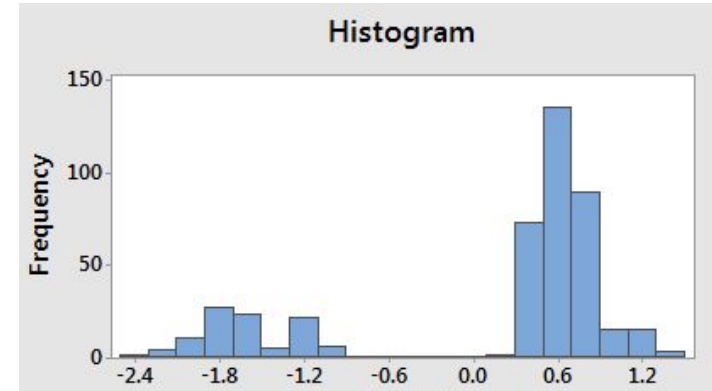
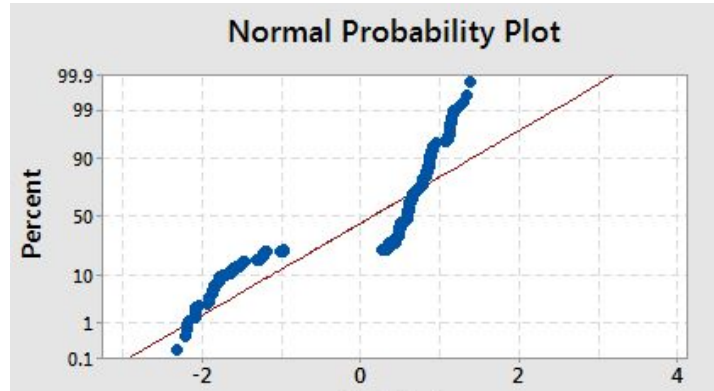
5.67%

AIC

450.71

Small Improvement in Adjusted R-Square!

Predicting IPO Pricing - Regression Model 2 Residual Plots



Predicting IPO Pricing - Conclusion & Key Takeaways

Key Project Takeaways

Regression Models

- ❑ Require trial and error and don't always present the findings you would expect

Data Sets

- ❑ Are difficult to collect and clean...

MiniTab

- ❑ Friend or foe?

Binary Logistic Regression: Pricing With versus Mgmt Year Ex, Gross Procee, Mgmt Avg Age, ...

- * WARNING * When the data are in the Response/Frequency format, the Residuals versus fits plot is unavailable.
- * ERROR * The model could not be fit. Maximum likelihood estimates of parameters may not exist due to quasi-complete separation of data points. Please refer to help or StatGuide for more information about quasi-complete separation.