**Vithika Agrawal**

ANALYSIS OF PIONEERING EFFECTS ON INDUSTRIAL GOODS INDUSTRIES

This research paper addresses the pioneering effect of order of entry on market share in industrial goods markets. It also suggests the factors that can lead to an order to entry advantages. An exploratory data analysis suggests that there is a direct relationship between order of entry of a brand and its market share. This research further suggests three sources that leads to market pioneering advantages:

1. Market pioneering leads to long lived marketing advantages, with relative direct costs held as constant
2. Market can lead to direct cost savings with respect to competition
3. Market pioneer can yield relative consumer information advantages by way of product experiences or familiarity, leading to market share advantages

Market share is modelled as a linear function of order of entry with respect to price, purchase frequency, time and advertising; product quality relative to competition, price relative to competition and direct costs. In estimating the model, both exogenous as well as the endogenous explanatory variables are taken in to consideration to reduce the risk of finding spurious relationships.

**Introduction**

The focus of this research paper is to answer the following questions: Does order of entry in an industrial goods industry have a major role to play in sustaining competitive advantages? Empirical evidence in various papers including the same analysis done for the consumer goods industries suggest that there is direct relationship between the order of entry and market share of the pioneers. Companies like Coca-Cola, Hallmark and Birds eyes have been market leaders for decades. This research paper caters to whether there is the same effect of this in industrial goods industries.

However, there are two theories that can limit the possible effect of this study. Firstly, Pioneer advantage can be realized only when the pioneer is successful in understanding the consumer preferences by continually improving the product and keeping their differentiating factors intact against its competitors. Secondly, there is a possibility that the late entrants keeps a watch on the mistakes of the pioneers and then take strategic advantage of this by investing much lesser time and money on their R&Ds. This can give a substantial advantage to the late entrants by adopting “imitate and improve” strategy.

In addition to this, adopting trends in the market can also help the pioneers define their strategy to remain in the market. For example, Artificial Intelligence can be a key trend for a music industry today. Companies like Spotify has to keep up with 20,000 new songs that gets uploaded everyday and categorize them based on genres and liking of a consumer and then suggest them recommendations. For companies like Sony Music Entertainment, which has been a pioneer in the industry since the 2000s are struggling to keep up with these digital trends but still manage to stay in the top game when it comes to records and labels. Companies like Spotify and Apple are the market pioneers for this online digital streaming. This particular example can cater to the issue of how different trends can lead to having different pioneers in the industry for those trends.

The analysis of this paper includes Robinson and Fornell’s paper validation with industrial markets. Market pioneer in the PIMS (Profit Impact of Market Strategies) data has its findings in line with the consumer goods industries. The data suggest that the pioneers have substantially high market shares than the late entrants. The market share here depends on both business and industry characteristics.

This research paper covers the model, data definitions, descriptive statistics, exploratory data analysis, results and findings, summary and conclusions. The appendix and the references are attached at the last of the paper. The model that fitted the most here was three step OLS regression which addresses the case of multicollinearity in the data.

**Market Pioneering Data and Definitions**

In the PIMS data, pioneer effect or the order of entry is based on entrance of business in the market. Businesses are classified into three categories:

* Pioneers ( Businesses that are the first in developing the product or service)
* Early Follower (Businesses which are early followers of the pioneers in still growing market)
* Late Entrant (Businesses that came late in the already established market)

The above three categories are mentioned as a categorical (binary) variable in the data. There are other variables as well which are used to define the model for studying the pioneering effect:

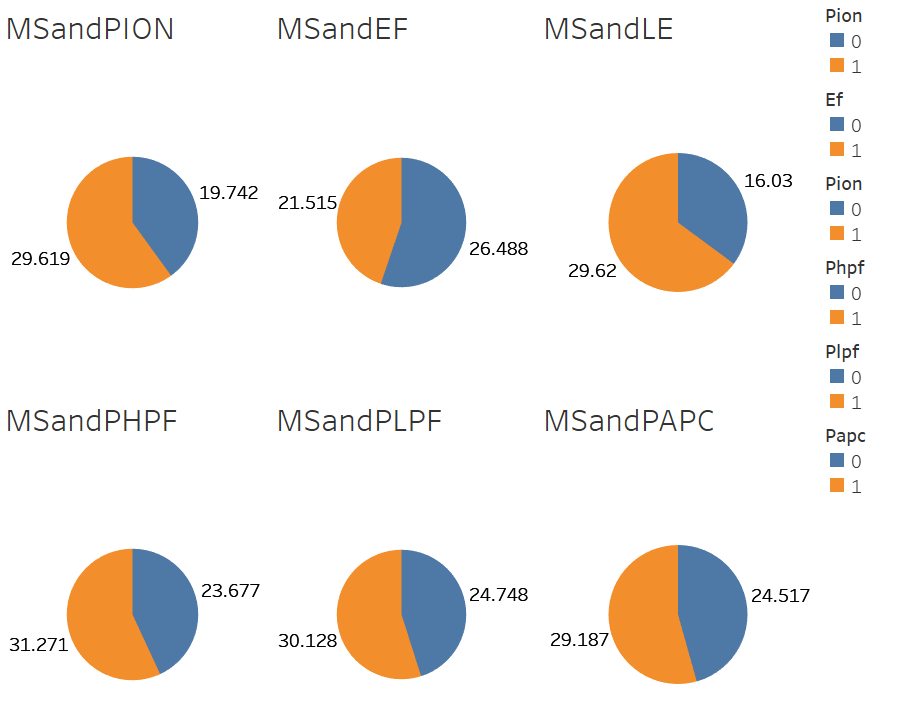
**Key Variables and Definitions:**

|  |  |  |
| --- | --- | --- |
| **Variables** | **Variable Type** | **Definitions** |
| Market Share | Explanatory (Numeric) | This is the share of the served market accounted for by this business |
| Order of Entry  Pioneer (PION)  20-Year (TYRP)  Early Follower(EF) | Categorical (Binary) | 1. Pioneers (1 if the business is a market follower, 0 otherwise) 2. 20-Year pioneer ( if the business is a market pioneer and has been in the market for 20 years or longer, 0 otherwise) 3. Early Follower (1, if the business is an early follower, 0 otherwise) |
| Relative Product Quality | Explanatory (Percentage) | Gives estimate of percentage of business’s sales volume accounted for by products and services from the perspective of the customer. The value is a percentage difference value between superior and inferior products. |
| Relative product line breadth | Categorical | Relative to the weighted average of the product lines of the three largest competitors, estimate the breadth of the product line of the business:  1 – Narrower  2 – Same  3 – Broader |
| Relative Price | Explanatory  ( Numeric) | Estimate the average level of selling prices of this business’s product and services (with respect to the competitors = 100%) |
| Relative Direct Costs | Categorical | Average level of this direct costs per unit of products and services relative to the average level of its three competitors.  1 – Less  2 – Similar  3 – More |
| Pioneer High Purchase Frequency (PHPF) | Categorical (Binary) | 1, if the business is a pioneer and the product is purchased once a month or more, 0, otherwise |
| Pioneer Low Purchase Frequency (PLPF) | Categorical (Binary) | 1, if the business is a pioneer and the product is purchased once a year or less, 0, otherwise |
| Pioneer Seasonal Product Change (PSC) | Categorical (Binary) | 1, if the business is a pioneer, and all or a part of the product line is changed on a seasonal basis, 0, otherwise |
| Pioneer annual/ periodic product change (PAPC) | Categorical (Binary) | 1, if the business is a pioneer and all or a part of the product line is changed on annual or periodic basis, 0 , otherwise |
| Number of competitors (NCOMP) | Categorical | Relative to last year’s data being entered, how many competing businesses were in the served market? Ignoring competitors with less than 1% of the served market |
| Relative advertising and promotion (MKTEXP) | Categorical | Relative to leading competitors, how much did the business spend, as a percentage of sales, on advertising and sales promotion?  1 – much less  2 – somewhat less  3 – about the same  4 – somewhat more  5 – much more |
| Percentage new products (PNP) | Explanatory | Percentage of the total sales for this business was accounted for by products introduced during the preceding 3 years? |
| Relative customer type (CUSTTYP) | Categorical | Estimate the breadth of this business’s served market, relative to the average of its 3 leading competitors:  1 - less than competitors  2 - same as competitors  3 - more than competitors |
| Relative number of customers (NCUST) | Categorical | Approximately how many immediate customers were served by the business?  1 – 3 or fewer 5 – 50-99  2 – 4-9 6 – 100-999  3 – 10-19 7 – 1000 – 9999  4- 20-49 8 – 10,000 or more |
| Relative customer size (CUSTSIZE) | Categorical | Same definition as that of Customer type |
| Plant and Equipment Newness (PENEW) | Explanatory (Numeric) | The net book value of plant and equipment divided by the gross book value |
| Capacity Utilization (CAP) | Explanatory (Percentage) | The average percentage of standard capacity utilized |
| Relative Backward Vertical Integration (RBVI) | Categorical | Comparing the degree of backward vertical integration of this business with that of its leading competitors  1 – less  2 – same  3 – more |
| Employee Productivity (EMPRODY) | Explanatory (Numeric) | The average level of value added per employee, where value added equals sales less purchases |
| Percentage Employees Unionized (UNION) | Explanatory (Percentage) | The percentage of the employees of this business who are unionized. |

**Exploratory Data Analysis**

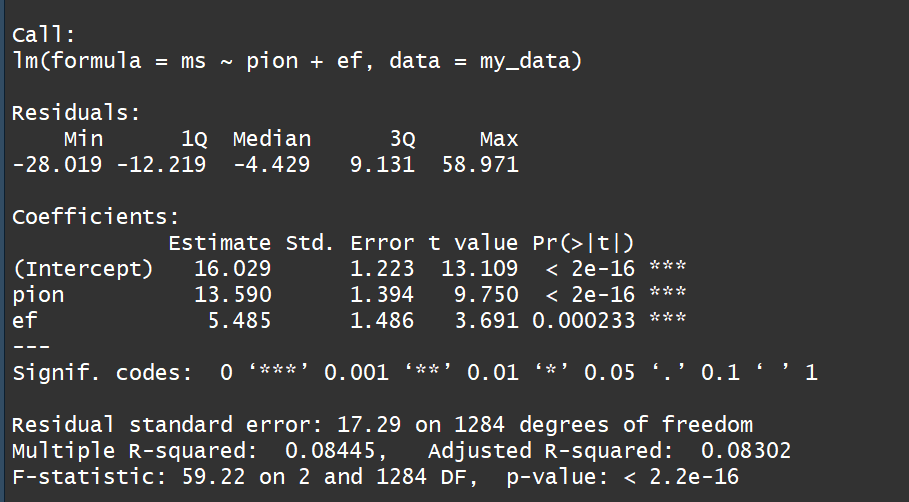
**Descriptive Statistics**

With the given data in hand, exploratory data analysis was conducted to understand the relationships between the variables and to understand the extent to which the variables were significant to include in the analysis. The PIMS data cover the established strategic business units. The sample used here is limited to mature industrial goods manufacturing businesses (n = 1287)



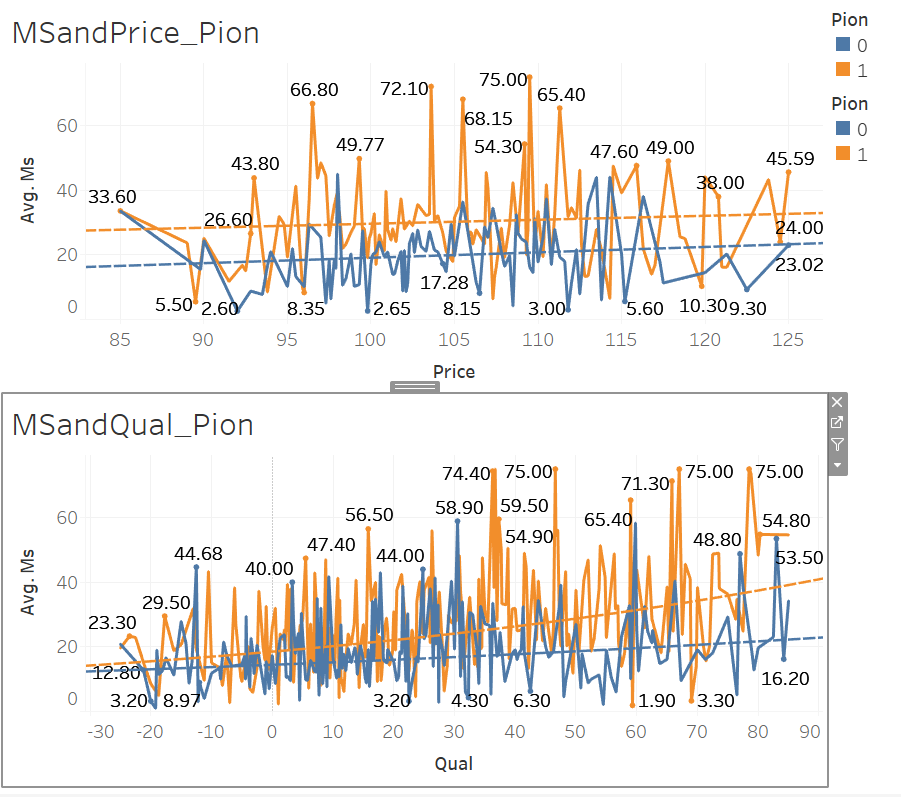
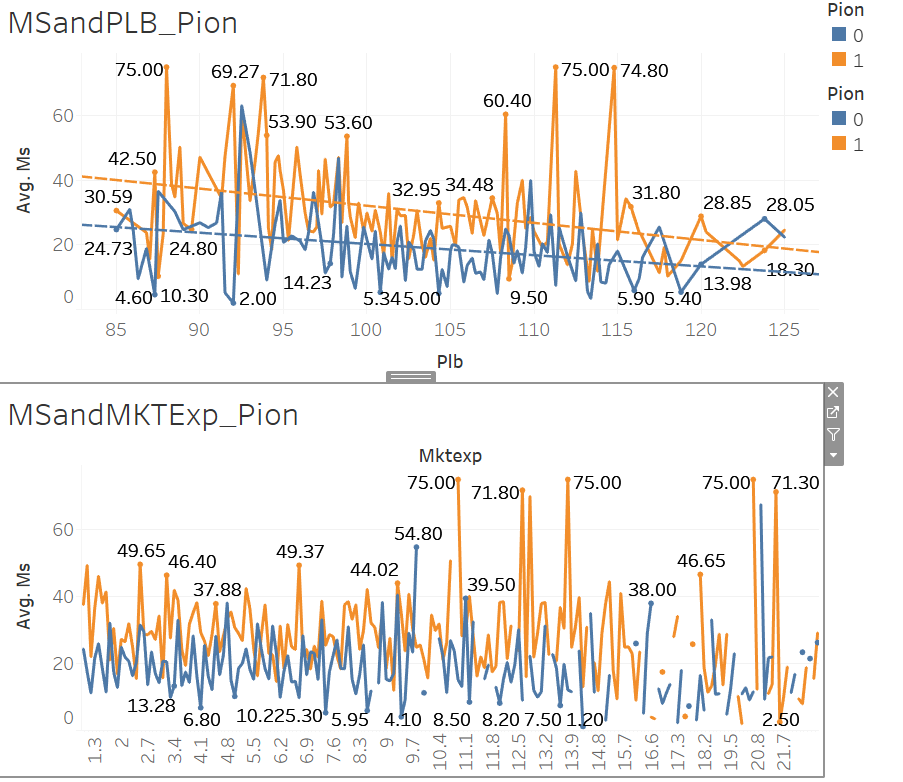
Firstly, the data was studied to understand the percentage of market share with respect to the categorical variables given in the data. The above exploratory analysis shows the percentage of market share with respect to the following variables:

* For mature industrial goods businesses, market pioneers (PION) have an average market share of 29.62%, early followers (EF) 21.52% and late entrants 16.03%
* Pioneer High Purchase Frequency (PHPF) with Market Share: This shows that the market pioneers who have high purchase frequency i.e. their product is purchased once a month, covers around 31.27% of market share. Hence this seems to be an important variable to consider in building a model.
* Pioneer Low Purchase Frequency (PLPF) with Market Share: This shows that the market pioneers who have low purchase frequency i.e. their product is purchased once a year, covers around 30.13% of the market share. Hence, we can observe a relationship between the purchase frequency and the market share.
* Pioneer Annual/ Periodic Change (PAPC) with Market Share: This shows that the market pioneers whose product line changes in annual or periodic basis, has 29.18% of market share.

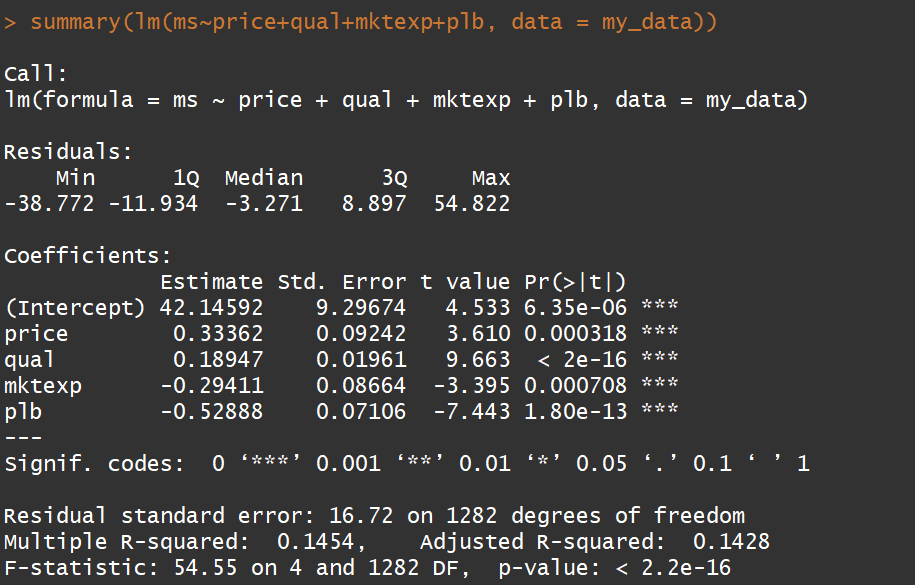


Order of market entry (Pioneers as well as the early followers) alone explains 8.5% of the variation as per data. Hence, the empirical association between order of entry as well as the market suggest further analysis.

**Understanding the distribution of the data: (Normality, Variance and Linearity of the sample data) – To check if there are any departures from the assumptions or not.**

The graphs above shows the relationship between the following given the company is a pioneer or not: Market Share and the Price of the product, Market Share and the Product line breadth, Market share and the relative quality of the product and Market Share and the Marketing expenditure on the product. When we closely observe all the graphs mentioned above, we notice that for each individual graphs the variations are huge and we can observe non normality among each relationships. Hence, we can conclude that given the variations, we can’t consider only one variable in the regression to study its effect on the market share.

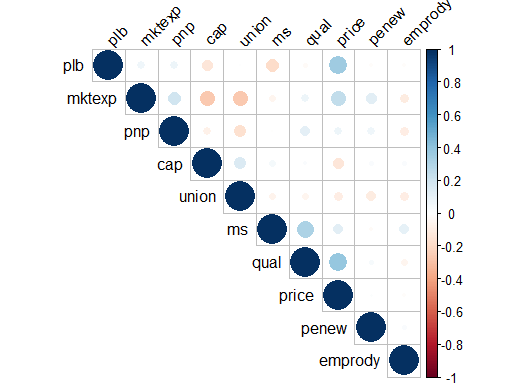


Doing the simple linear regression of price, qual, mktexp and plb on market share, we observed that the p-value of all the variables is less than 0.001. Hence, all the variables above are statistically significant. This suggests that these all variables need to be considered in the model to study the market share.

Also, it is observed that the F-value is huge, this suggests that the variation in this model is huge and significant. Hence, it is important to understand the reason for this variation and check whether there exists a case of correlation among the explanatory variables i.e. whether there is multicollinearity in the data.

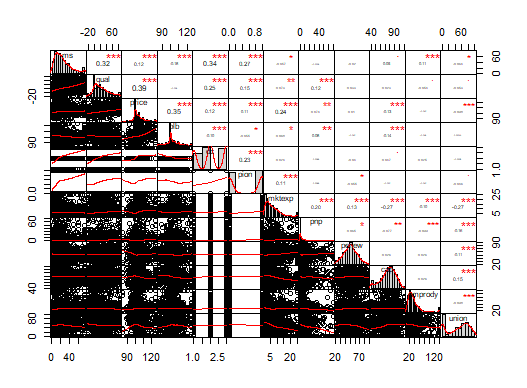
**Checking the property of independence among the variables:**

In order to check whether there is any independence in the data, correlation matrix was created in R to understand the effect of multicollinearity and weather this exists or not.



We can see the existence of multi collinearity within the data among several variables. Given the existence of multicollinearity in the model, we cannot simply run Ordinary least square method as it will not give accurate results. In order to deal with multicollinearity, two step or three step least square method would be more suitable.

For understanding the seriousness of the multicollinearity, correlation chart was created to understand the significance level of each correlations.



All the explanatory variables which are not categorical are taken into consideration except for Pioneering and Direct Cost variables are these are the primarily variables to be considered. Other categorical variables in the model are included as per their practical significance with these models and hypothesis.

**Model 1: Market Share = Quality + Price + Product Line Breadth + Direct Costs + Pioneering Effect +Mktexp+ Employee Productivity + Early Followers + TYRP + PHPF +PLPF+ PSC+ PAPC**

From the above correlation matrix, we can see that Market Share (MS) has high correlation with quality, price, product line breadth, direct costs, pioneering effect, and employee productivity. We see the graph of Direct Cost as well as pioneering to be showing that the variables here are categorical. Apart from the above variables, taking the variable Early Follower (EF) as well as 20 year market Pioneers (TYRP) would make sense to understand the impact of pioneers on the model better, Given the above correlation, we deduce the following equations for multi-step OLS regression.

After deducing this model, in order to deal with multicollinearity, the following models would be appropriate to run multi-step OLS regression.

**Model 2: Quality of the Product = Price + DC + Pion + MKTEXP + PNP + EF + TYRP**

We see with the correlation matrix above that Quality of the Product has high correlation with the variables such as Price, DC, Pioneering, Marketing Expenditure, and Percentage of New Products. The categorical variables Early Followers (EF) as well as TYRP (Twenty Year Pioneers) has been included in the model for studying the pioneering effect on the quality of the product better.

**Model 3: Price = MS + Qual + PLB + DC + Pion + MKTEXP + PNP + CAP + UNION**

We see from the correlation matrix above that the Price of the product is highly correlated with the market share, quality of the product, product line breadth, relative direct costs, pioneering effect, marketing expenditure, percentage of new products, capacity utilization as well as on the percentage of employee unions.

**Model 4: PLB = Direct Costs + Pion + Mktexp + PNP + CAP + EF + TYRP + CUSTTYP + NCUST + CUSTSIZE**

We see from the correlation matrix above that the product line breadth is highly correlated with direct costs, pion, marketing expenditure, percentage of new products and capacity utilization. The binary variables EF and TYRP is added to study their impact on the pioneering effect. Apart from these variables, the categorical variables of the customers can also have an impact on the product line breadth, hence, these variables are also added into the equation.

**Model 5: DIRECT COSTS = Market Share + Quality + Price + PLB + Pion + CAP + EF + TYRP**

We see from the correlation matrix above that the relative direct cost of the variable is highly correlated with market share, quality of the product, price of the product, Pion and Capacity utilization. The variables EF and TYRP are added here to study their impact on the pioneering effect.

These are the five models that has been considered as per the exploratory data analysis performed above.

**Modelling**

With the above data, there are five equations that were felt suitable to study the conceptual relationships between the variables.

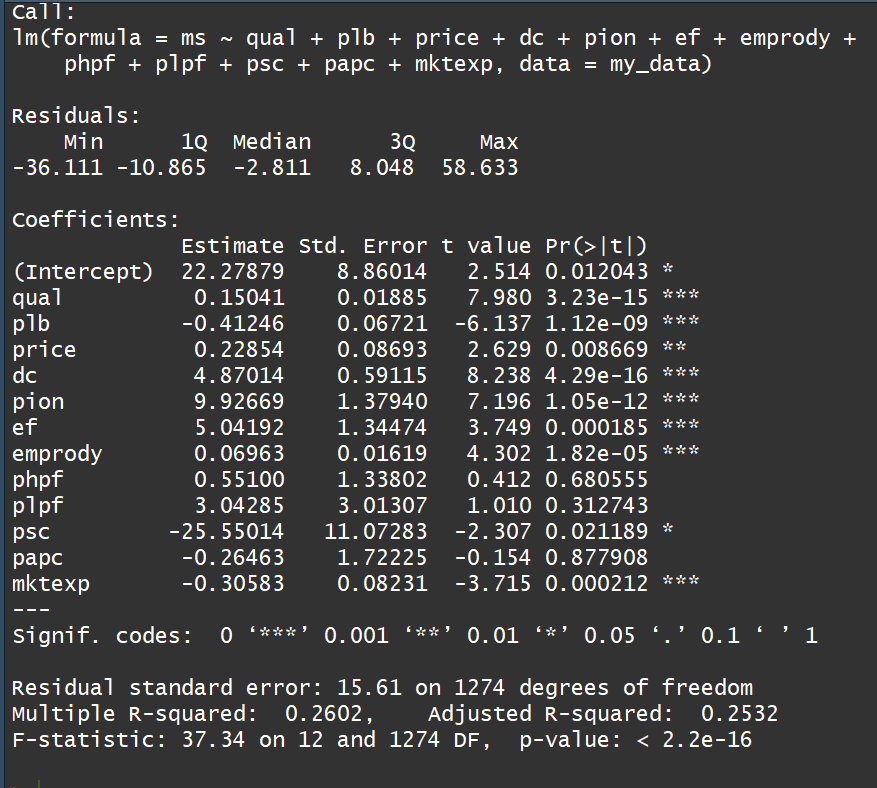
Model 1: Market Share = Quality + Price + Product Line Breadth + Direct Costs + Pioneering Effect +Mktexp+ Employee Productivity + Early Followers + TYRP + PHPF +PLPF+ PSC+ PAPC

Model 2: Quality of the Product = Price + DC + Pion + MKTEXP + PNP + EF + TYRP

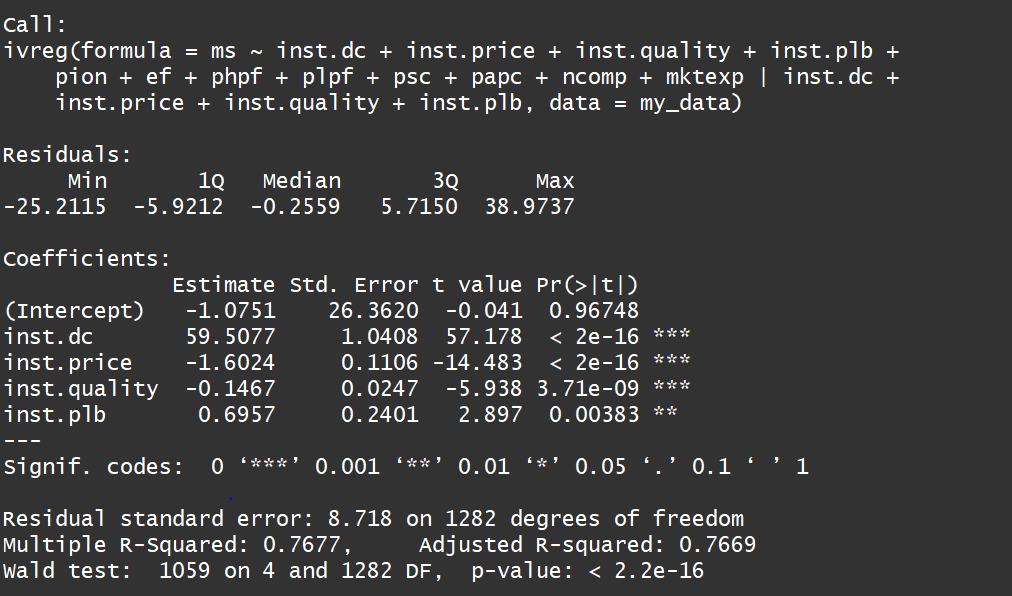
Model 3: Price = MS + Qual + PLB + DC + Pion + MKTEXP + PNP + CAP + UNION

Model 4: PLB = Direct Costs + Pion + Mktexp + PNP + CAP + EF + TYRP + CUSTTYP + NCUST + CUSTSIZE

Model 5: DIRECT COSTS = Market Share + Quality + Price + PLB + Pion + CAP + EF + TYRP



With the simple OLS for Model 1, we can see that if the company is a pioneer, the market share of the company increases by 9.92%. However, as the variables are highly correlated with each other, TSLS or 2SLS needs to be done to deal with multicollinearity and endogenity.

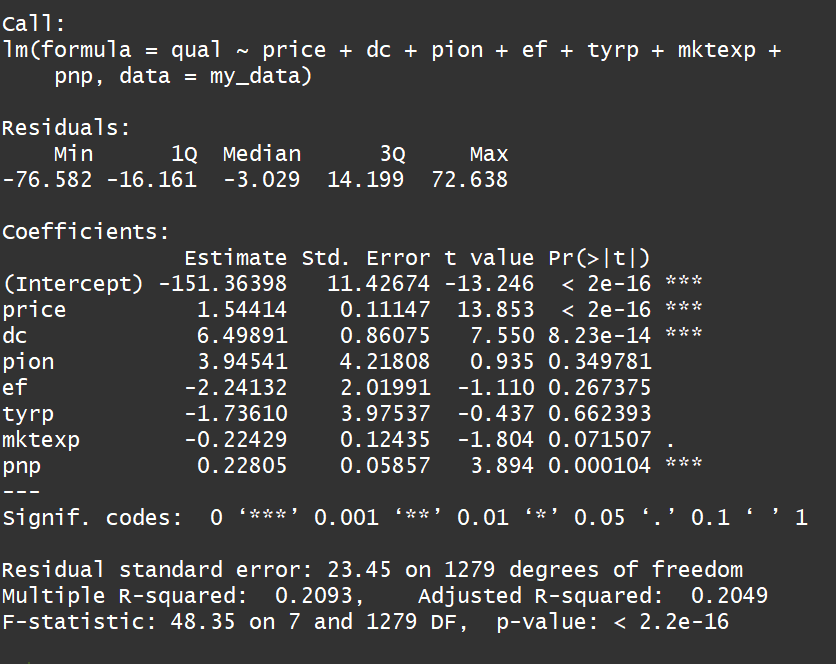
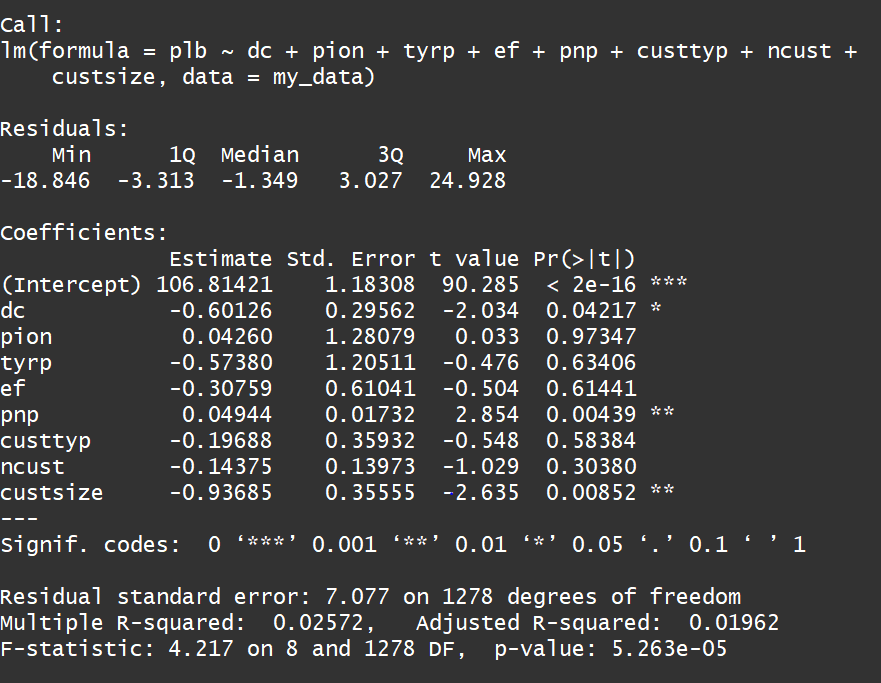


The 2SLS or TSLS model was created with the models above. The IVREG package gives the effect of instrumental variables on the dependent variable. The interpretations are as follows:

* With one unit increase in the price of the product, the market share of that product decreases by 1.60%
* With one percent increase in the quality of the product, the market share of that product decreases by 0.14%
* If direct cost of a pioneer business is lesser than their competitors, then the market share of the product increases by 59.5%
* If the product line breadth of a pioneer business is broader than their competitors, then the market share of the product increases by 0.69%

The above variables in TSLS shows direct impact of direct costs, price, quality and product line breadth on the market share of the product. However, we need to understand how pioneering effect can impact the market share of the product.

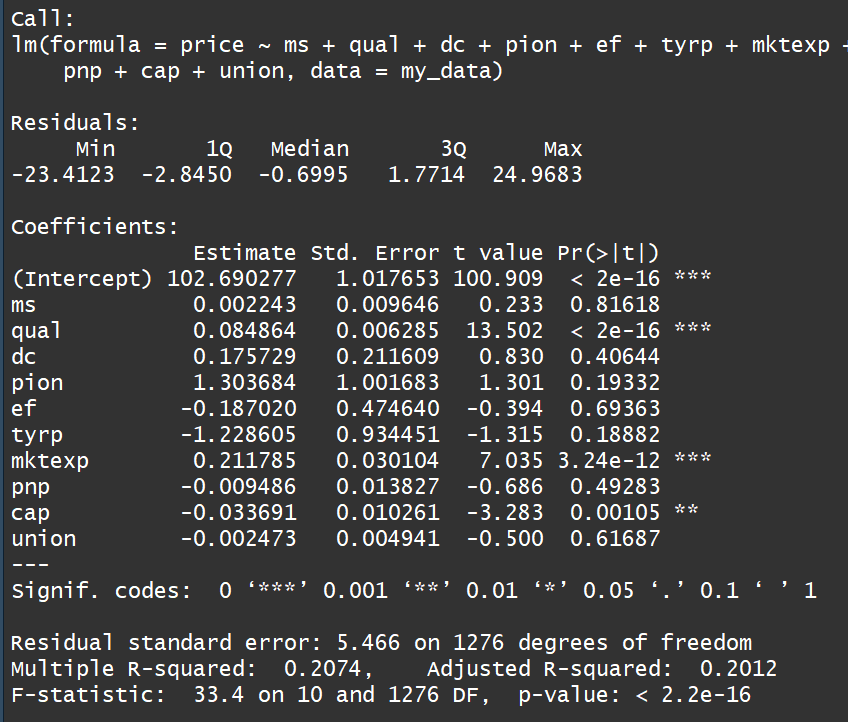
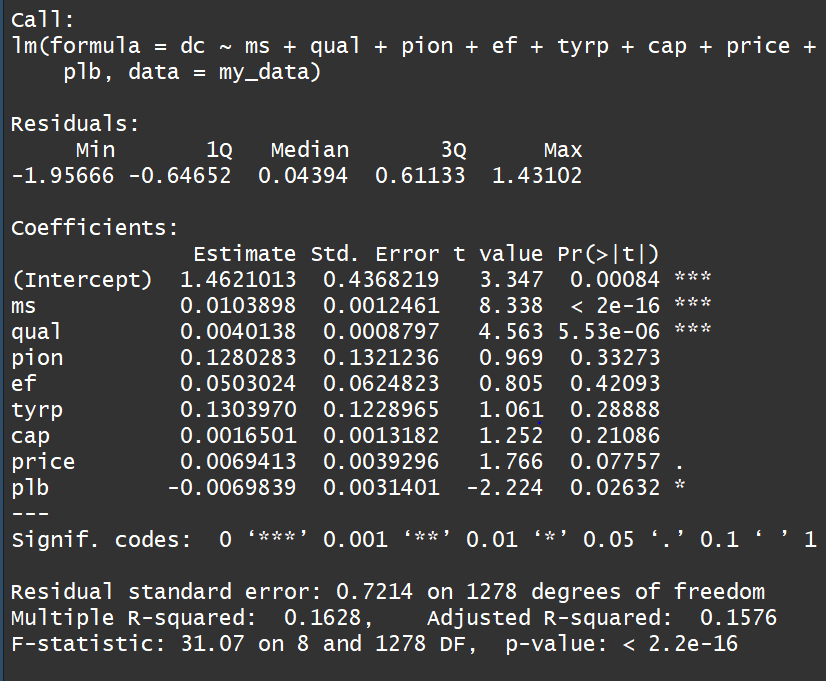
**Pioneering Effect with respect to Instrumental Variables:**

**Figure 1**  **Figure 2**

In Figure 1, it can be seen that if a company is a pioneer, the relative quality of the product increases by 3.94%.

In Figure 2, it can be seen that if a company is a pioneer, the average product line breadth is 4% broader than the other competitors.

 \

**Figure 3**  **Figure 4**

In Figure 3, we can see that if a company is a pioneer, the price of the product increases by 1.3% .

In Figure 4, we can see that if a company is a pioneer, the relative direct cost is 0.12% lesser than the competitors.

Hence we observe that pioneering effect indirectly effect on the market share of the product. There are other factors such as price, quality, product line breadth as well as the relative direct costs that directly affect the market share of the company.

Thus, overall, we can observe that indirectly, market pioneers have advantage in market share over the late entrants. The findings indicate that pioneers in industrial markets can have sustainable competitive advantage than the others as they can seize opportunities to develop them and consequently maintain a high market share for decades.

**Analysis and Results**

The model are estimated by multi-stage OLS method. The model suggested that market pioneers tend to have substantially high market shares than the late entrants. A product which is superior to the competitor’s products and have certain other industry characteristics such as high purchase frequency, seasonal product change, annual/ periodic product change helped to explain the pioneer share advantages. Although, in a practical environment, pioneering, especially in the industrial market is not easy, but the findings suggest that many pioneers in this area develop huge sustainable and competitive advantages.

**References and Appendix**

**Sources:**

1.Sources of Market Pioneer Advantages in Consumer Goods Industries by William T. Robinson and Claes Fornell

2. https://www.emeraldinsight.com/doi/full/10.1108/02634509610121505?fullSc=1&

**Codes for reference :**

library(data.table)

my\_data <- read.table(file = "clipboard", sep = "\t", header = TRUE)

str(my\_data)

View(my\_data)

**#Exploratory Data Analysis**

library(ggplot2)

**#Wider Descriptive Statistics Graphs done in Tableau**

**#Exploratory - See the amount of variation explained by the variables Early Followers, Pioneers and Late Entrants**

summary(lm(ms~pion+ef, data = my\_data))

summary(lm(ms~price+qual+mktexp+plb, data = my\_data))

**#Correlation matrix for the data**

res <- cor(my\_data[,c(1,2,3,4,13,15,19,20,22,23)])

round(res, 2)

library(Hmisc) # pvalue significance levels of correlation matrix

res2 <- rcorr(as.matrix(my\_data[,c(1,2,3,4,5,13,15,19,20,22,23)]))

res2

res2$r **# Extracting the correlation coefficients**

res2$P **# Extract P values**

library(corrplot)

corrplot(res, type = "upper", order = "hclust",

tl.col = "black", tl.srt = 45)

# Insignificant correlation are crossed

corrplot(res2$r, type="upper", order="hclust",

p.mat = res2$P, sig.level = 0.01, insig = "blank")

# Insignificant correlations are leaved blank

corrplot(res2$r, type="upper", order="hclust",

p.mat = res2$P, sig.level = 0.01, insig = "blank")

library(PerformanceAnalytics)

chart.Correlation(my\_data[,c(1,2,3,4,5,6,13,15,19,20,22,23)], histogram=TRUE, pch=19)

**#In the above plot:**

**#The distribution of each variable is shown on the diagonal.**

**#On the bottom of the diagonal : the bivariate scatter plots with a fitted line are displayed**

**#On the top of the diagonal : the value of the correlation plus the significance level as stars**

**#ach significance level is associated to a symbol : p-values(0, 0.001, 0.01, 0.05, 0.1, 1) <=> symbols("\*\*\*", "\*\*", "\*", ".", " ")**

DF <- my\_data

DF[] <- lapply(DF,as.integer)

library(sjPlot)

sjp.corr(DF)

sjt.corr(DF)

?sjp.corr

**#Modelling**

**#First modelling**

olsmodel <- lm(ms~qual+plb+price+dc+pion+ef+emprody+phpf+plpf+psc+papc+mktexp, data = my\_data)

summary(olsmodel)

first.stage.1 <- lm(qual~price+dc+pion+ef+tyrp+mktexp+pnp, data = my\_data)

summary(first.stage.1)

my\_data$inst.quality <- first.stage.1$fitted.values

first.stage.2 <- lm(plb~dc+pion+tyrp+ef+pnp+custtyp+ncust+custsize, data = my\_data)

summary(first.stage.2)

my\_data$inst.plb <- first.stage.2$fitted.values

first.stage.3 <- lm(price~ms+qual+dc+pion+ef+tyrp+mktexp+pnp+cap+union, data = my\_data)

summary(first.stage.3)

my\_data$inst.price <- first.stage.3$fitted.values

first.stage.4 <- lm(dc~ms+qual+pion+ef+tyrp+cap+price+plb, data = my\_data)

summary(first.stage.4)

my\_data$inst.dc <- first.stage.4$fitted.values

install.packages("AER")

library(AER)

tslsmodel<-ivreg(ms~inst.dc+inst.price+inst.quality+inst.plb+pion+ef+phpf+plpf+psc+papc+ncomp+mktexp|inst.dc+inst.price+inst.quality+inst.plb, data = my\_data)

summary(tslsmodel)

install.packages("estimatr")

library(estimatr)

rob\_model<-iv\_robust(ms~inst.dc+inst.price+inst.quality+inst.plb+pion+ef+phpf+plpf+psc+papc+ncomp+mktexp|inst.dc+inst.price+inst.quality+inst.plb, data = my\_data)

summary(rob\_model)

model\_fin<-lm(ms~inst.dc+inst.price+inst.quality+inst.plb+pion+ef+phpf+plpf+psc+papc+ncomp+mktexp, data = my\_data)

summary(model\_fin)