

MULTIPLE LINEAR REGRESSION

***LUXURY BRAND MARKET IN INDIA & THE FACTORS DRIVING
THE MARKET***

Dec 2019

By

Anshu Somani (IPMX12055)

Manish Yadav (IPMX12021)

Sunanda Ghosh (IPMX12099)

Chitrangana Singh (IPMX12063)

Gourab Biswas (IPMX12065)

Sourav Debnath (IPMX12096)

INDEX

Abstract.....	3
Introduction.....	4
Explanatory variables.....	5
Assumptions.....	7
Steps in a Multiple Linear Regression.....	8
Multi-collinearity.....	10
Final Model.....	11
Conclusion.....	12
References.....	13

Abstract

The objective of this study is to understand how luxury brand consumption is quantitatively related to various economic and non-economic forces. We have gathered data from E-databases available in Gyanodyaya and also ministry databases like WorldBank and RBI.

We have analysed luxury brand market consumption for the past 14 years, ie., from FY 06 onwards. The fiscal year 2006 has been an instrumental year and has marked entry of renowned western brands in the Indian markets and also previously existing brands have widened their presence in the country. This is the year when Indian markets opened formally for FDIs and globalization took place in a massive form. Also, in the year 2008 onwards, e-commerce started off in India with a small base. Hence we have considered data from 2006 onwards to derive a meaningful insight on the luxury brand consumption in the Indian market.

We have used tools like R, Stata & E-views in our analysis.

Introduction

Luxury good (or upmarket good) is defined as one whose demand increases in a greater proportion with the increase in income.

Let us understand few economic concepts that can also be categorized as behavioural or psychological influences guiding luxury brand market.

The Veblen effect – perceived conspicuous value. Veblenian consumers attach greater importance to price as an indicator of prestige, because their primary objective is to impress others.

The snob effect – perceived unique value. Snob consumers perceive price as an indicator of exclusivity, and avoid using popular brands to experiment with inner-directed consumption

The bandwagon effect – perceived social value. Relative to snob consumers, bandwagon consumers attach less importance to price as an indicator of prestige, but will place greater emphasis on the effect they make on others while consuming prestige brands.

The hedonic effect – perceived emotional value. Hedonist consumers are more interested in their own thoughts and feelings, and thus will place less emphasis on price as an indicator of prestige.

The perfectionism effect – perceived quality value. Perfectionist consumers rely on their own perception of the product's quality, and may use price as further evidence of quality.

The above concepts are associated with the luxury good demand which states that quantity demanded for such goods increase with the increase in price (an exception to the law of demand). Luxury goods, hence, have high income elasticity of demand.

In our analysis, we seek to deep dive into the factors that actually cause such a behavior when the good in question is a luxury good.

The data we have considered is a time- series data (and not cross-sectional data) to predict how luxury brands consumption will move in the future.

Explanatory variables

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \beta_{10} X_{10}$$

Luxury Brand Consumption:

The 'Y' variable includes Total revenue of Shoppers Stop Limited (listed in NSE) which is taken as a proxy for spending on luxury brands. Monthly revenue has been calculated assuming the yearly revenue is equally spread.

Shoppers Stop has established 86 stores across 40 cities in India and has an umbrella of brands under its ambit which includes clothing, accessories, handbags, shoes, jewelry, fragrances, cosmetics, health and beauty products, home furnishing and decor products. Hence Shoppers Stop's revenues will indicate how Indian luxury brand market has grown over the past 13 years from the time of its incorporation in 2006. An e-store with delivery across major cities in India was launched in 2008, with a smartphone app in 2016.

Per Capita Income:

A general notion supports the belief that as people grow richer & wealthier, the spends on luxury goods increases. It can be supposed that the demand for luxury goods would be positively related to income as a rise in average household income would lead to an increased demand for luxury goods holding all other variables constant. Therefore, household income might be a useful predictor of luxury good spending in the economic model.

Real Interest Rate:

Logical reasoning would suggest that an increase in the real interest rate would decrease the demand for luxury goods. In general, as real interest rates rise the marginal cost of consumption becomes relatively more expensive, ceteris paribus; therefore, consumers tend to defer consumption and substitute it for saving. Here Substitution effect plays a role & Luxury goods tends to be substituted by Necessities as interest rate increases. Therefore, a rise in the real interest rate would make the marginal cost of consuming a luxury good even higher.

Following two types of interest rates are useful:

- a) The gross real excess returns in the stock market between time t and $t + 1$ is used in the model. The excess returns are calculated as the return on the S&P 500 less the return on the 3-month Treasury bill at the end of the corresponding fiscal year. δ_1 is the ceteris paribus effect of return on equity on luxury consumption growth.
- b) The real long-term interest rate as calculated by the yield on long-term government bonds is included in the model. λ_1 is the ceteris paribus effect of the real long-term interest rate on luxury consumption growth.

We have considered Risk free rate on a 10-year T-Bond as a proxy for Real interest rates

Youth population % in India:

Conventionally, period from adolescence to middle age is termed as youth. The National Youth Policy initially (in 2003) defined the youth as in the age group 13-35. A market survey shows that major consumers of luxury brands are youth in the aforesaid age-gap. Hence this variable inclusion in the model will determine whether this variable is significant or not.

Advertising expenditures:

Advertising might have a significant effect on luxury good demand. It can be proposed that an increase in advertising spending by luxury good brands will have a positive impact on the demand for luxury goods, ceteris paribus.

We have considered the selling and administration expenses of Shoppers Stop Limited as a proxy for advertisement expenditure.

Real GDP:

Real GDP is a direct indicator of the health of the economy. The general proposition is that 2006 onwards, GDP grew manifold leading to increase in per capita income & accordingly increase in luxury brand consumption. With this model we seek to test the same.

We can expect correlation between two independent variables here- GDP & Per capita income, which will be discussed in depth later.

Recession or Boom:

A categorical variable is inserted to determine whether economy was in a recession or boom phase in corresponding years. Luxury goods' demand is expected to be positively related with boom years & negatively related with recession years. This variable is included in the model to determine the overall health of the economy.

Price Index:

The Veblen effect can be tested via price indices. Veblen effect indicates the presence of conspicuous consumption wherein consumers tend to buy more when the prices go up. Bandwagon effect, Snob effect and Veblen effect tend to explain the consumers' behavior and their so-called "status spending" on luxury goods. We shall test this by incorporating price index in our model as one of the explanatory variables.

Exchange Rate:

We have taken currency movements as a regressor variable impacting luxury brand consumption expenditures. The exchange rate is taken for Euro since maximum luxury brands like Gucci, Louis Vuitton, Zara, H&M have their origin in Europe.

The data shows how INR has depreciated against Euro over the periods in consideration. Through this model we seek to establish whether this is a significant factor driving luxury brand consumption.

Social Media Presence:

By far the most influential factor, as we are expecting, driving brand expenditures. Model will help us to understand how significant or otherwise it is.

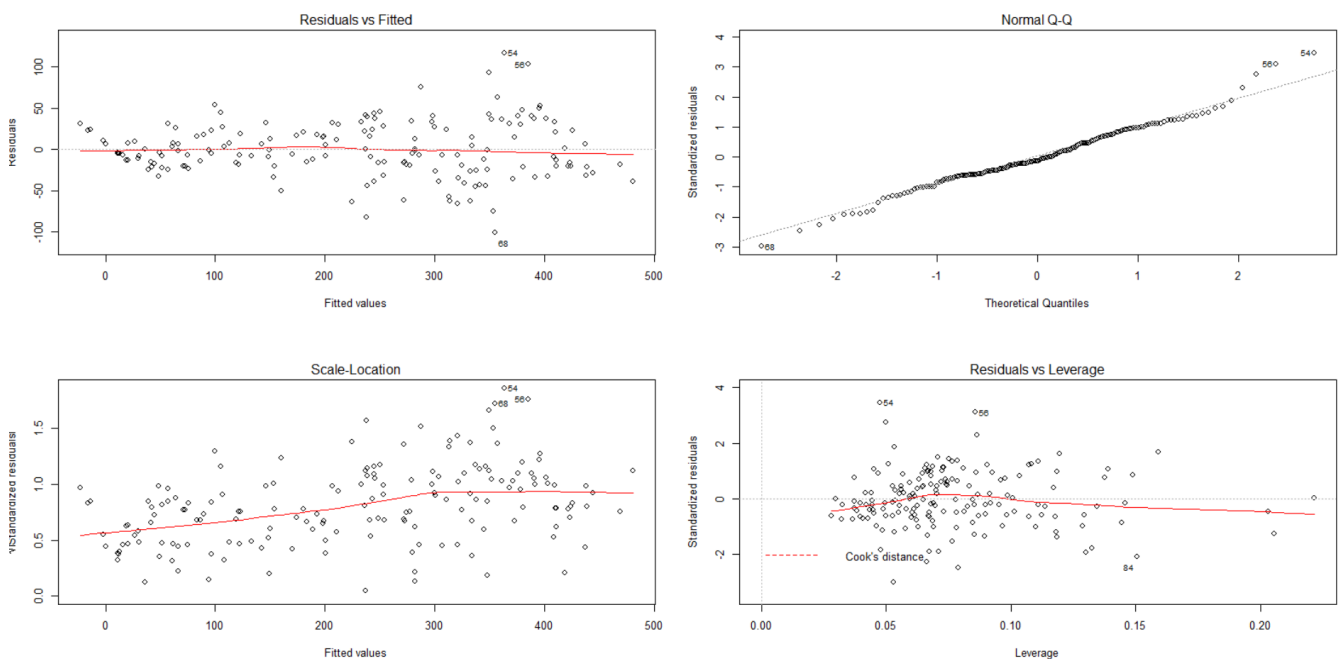
Imports (Non-Oil):

The brand expenditure is expected to have a strong dependency on imports. Brand itself is a western concept and it originated when “classy” goods were imported from western countries. The model will help us further our understanding on this variable.

Assumptions

Before applying the Multi-linear regression, we will first examine whether the assumptions of Linear regression hold good for the data in question.

Plotting the data gives the following graphs:



Normal Distribution- The Q-Q Plots show a fairly straight line with few outliers. Hence all in all data is normally distributed

Homoscedasticity- The Residual Vs Fitted graph is not showing a funnel shape and hence we can say that errors have equal variances

Linear relationship- The Residual Vs Fitted graph does not exhibit a parabolic shape or funnel- shape. Hence we can say that linearity exists

Errors are independent- We use Durbin Watson statistic to establish whether there exists any auto-correlation in error terms. Below is the snapshot of score:

Explanatory variables	Score
Real Interest rate	0.00155
Per Capita income	0.0487
Advertising expenditure	0.014
Consumer Price Index	0.0009
Imports- Non oil	0.0807
Recession/ Boom	NA
Social Media influence	NA

Durbin-Watson statistic between 0 & 2 indicates a positive auto-correlation and between 2 & 4 indicates a negative auto-correlation. In this case, it can be said that there exists a *positive auto correlation in the data*.

The nature of data is economic variables over a time period and these variables are correlated with previous years' data since many reforms may have taken place which lead to improvement this year as compared to previous years. Also, as economy grows, so do the variables like per capita income and alike. Hence Multi-linear regression cannot be used in such cases since either relationship is non-linear or models like ARIMA would apply and give better inferences.

However, we have still used Multi-linear regression assuming that *theoretically* as such there is no dependency of observations to other observations but since the nature of variables is such (economic variables) that it is bound to grow positively as we go ahead into future years (assuming no Depression takes place) and hence multi-linear model can justify the results.

STEPS in a MULTIPLE LINEAR REGRESSION

Multiple Linear Regression involves the following:

- Framing of the model with identification of various explanatory variables
- Determining the line of best fit and examining its fitness
- Hypothesis testing to understand significance of the sample to the overall population

a) Model:

With the help of R, we have framed the model as-

Luxury Goods consumption = $-(3.785e+02) + (5.437e-03)$ Per Capita income + $(1.327e)$ Real interest rate + $(3.352e)$ Youth population % in India + $(3.686e)$ Advertising expenditure $-(7.237e-06)$ GDP – $(2.730$

e+01) Recession/Boom [Yes] + (2.029e) Consumer Price Index –(9.283e-01) Exchange Rate – (1.300e+02) Social Media [Highly strong] + (3.924e+01) Social Media [No influence] – (8.096e+01) Social Media [Strong] – (1.648e-04) Imports

b) Examining Fitness:

Alternate 1

Null Hypothesis H0: Model is not a best fit

Alternate Hypothesis H1: Model is good fit

F-statistic= 216.13

p-value~ 0

Hence Model is a good fit

Alternate 2

Another factor by which we examine the goodness of fit is via Adjusted R square.

Adjusted R square= 93.93%

c) Significance of the factors in explaining the population:

We will examine the significance of each factor to understand if its inclusion in the model is significant enough

Null Hypothesis H0: Beta coefficient is insignificant

Alternate Hypothesis H1: Beta coefficient is significant

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-3.785e+02	1.186e+02	-3.192	0.00171	**
`Per capita income (nominal)`	5.437e-03	1.807e-03	3.009	0.00306	**
`Real Interest rate`	1.327e+00	5.315e+00	0.250	0.80314	
`Youth population % in India`	3.352e+00	3.004e+00	1.116	0.26615	
`Advertising expenditure`	3.686e+00	2.906e-01	12.684	< 2e-16	***
`GDP (current prices)`	-7.237e-06	2.410e-05	-0.300	0.76440	
`Recession/ Boom`Yes	-2.730e+01	8.541e+00	-3.197	0.00168	**
`Consumer Price Index`	2.029e+00	3.109e-01	6.525	9.14e-10	***
`Exchange rate (INR against Euro)`	-9.283e-01	6.911e-01	-1.343	0.18117	
`Social Media Presence`Highly Strong	-1.300e+02	2.506e+01	-5.189	6.53e-07	***
`Social Media Presence`No influence	3.924e+01	1.431e+01	2.742	0.00682	**
`Social Media Presence`Strong	-8.096e+01	1.644e+01	-4.926	2.14e-06	***
`Imports (Non- Oil)`	-1.648e-04	8.321e-05	-1.980	0.04945	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

i. Per capita income-

P- value is 0.003 & t-statistic is 3.009

Thus the aforesaid factor is significant at alpha = 1%

Hence we reject Null hypothesis at above alpha level

ii. Real interest rate-

P- value is 0.803 & t-statistic is 0.25

Thus the aforesaid factor is not significant at alpha = 10%

Hence we do not reject Null hypothesis at above alpha level

Similarly, the above result will be used for other factors and the following factors are not significant:

Real Interest rate
Youth population
GDP (Current prices)
Exchange Rate

MULTI-COLLINEARITY

The present model requires a check for presence of multi-collinearity wherein two explanatory variables are correlated in a linear fashion and which impacts the model.

For this purpose, we run a *Variance Inflationary Factor* check to if Multi collinearity exists. In general, a VIF of more than 5 is considered as a cut-off for determining existence of multi-collinearity.

Below is the summary of VIF results. Clearly, presence of multi-collinearity is confirmed.

```
> vif(g)
```

	GVIF	Df	GVIF ^{1/(2*Df)}
`Per capita income (nominal)`	3.799962	1	1.949349
`Real Interest rate`	1.827222	1	1.351748
`Youth population % in India`	1.085512	1	1.041879
`Advertising expenditure`	7.686159	1	2.772392
`GDP (current prices)`	12.684297	1	3.561502
`Recession/ Boom`	1.623496	1	1.274165
`Consumer Price Index`	51.418378	1	7.170661
`Exchange rate (INR against Euro)`	5.444675	1	2.333383
`Social Media Presence`	87.884273	3	2.108556
`Imports (Non- Oil)`	4.948902	1	2.224613

The next step is to remove redundant variables from the model. We will use *Step function* for the same. The results of the same as run in "R" is presented below:

```
Start: AIC=1202.4
`Luxury Goods consumption` ~ `Per capita income (nominal)` +
  `Real Interest rate` + `Youth population % in India` + `Advertising expenditure`
+ `GDP (current prices)` + `Recession/ Boom` + `Consumer Price Index` +
  `Exchange rate (INR against Euro)` + `Social Media Presence` +
  `Imports (Non- Oil)`
```

	Df	Sum of Sq	RSS	AIC
- `Real Interest rate`	1	74	184744	1200.5

- `GDP (current prices)`	1	107	184777	1200.5
- `Youth population % in India`	1	1484	186154	1201.7
- `Exchange rate (INR against Euro)`	1	2150	186820	1202.3
<none>			184670	1202.4
- `Imports (Non- Oil)`	1	4672	189342	1204.6
- `Per capita income (nominal)`	1	10784	195455	1209.9
- `Recession/ Boom`	1	12177	196847	1211.1
- `Social Media Presence`	3	34425	219095	1225.1
- `Consumer Price Index`	1	50728	235398	1241.2
- `Advertising expenditure`	1	191673	376343	1320.0

Step: AIC=1200.46

`Luxury Goods consumption` ~ `Per capita income (nominal)` +
`Youth population % in India` + `Advertising expenditure` +
`GDP (current prices)` + `Recession/ Boom` + `Consumer Price Index` +
`Exchange rate (INR against Euro)` + `Social Media Presence` +
`Imports (Non- Oil)`

	Df	Sum of Sq	RSS	AIC
- `GDP (current prices)`	1	107	184851	1198.6
- `Youth population % in India`	1	1480	186224	1199.8
- `Exchange rate (INR against Euro)`	1	2076	186820	1200.3
<none>			184744	1200.5
- `Imports (Non- Oil)`	1	4630	189375	1202.6
- `Per capita income (nominal)`	1	10732	195476	1208.0
- `Recession/ Boom`	1	12717	197461	1209.7
- `Social Media Presence`	3	37472	222216	1225.5
- `Consumer Price Index`	1	50846	235591	1239.3
- `Advertising expenditure`	1	208546	393291	1325.4

Step: AIC=1198.56

`Luxury Goods consumption` ~ `Per capita income (nominal)` +
`Youth population % in India` + `Advertising expenditure` +
`Recession/ Boom` + `Consumer Price Index` + `Exchange rate (INR against Euro)`
+
`Social Media Presence` + `Imports (Non- Oil)`

	Df	Sum of Sq	RSS	AIC
- `Youth population % in India`	1	1421	186272	1197.8
- `Exchange rate (INR against Euro)`	1	2019	186870	1198.4
<none>			184851	1198.6
- `Imports (Non- Oil)`	1	4525	189376	1200.6
- `Per capita income (nominal)`	1	10762	195613	1206.1
- `Recession/ Boom`	1	12852	197703	1207.8
- `Social Media Presence`	3	38130	222981	1224.1
- `Consumer Price Index`	1	68145	252995	1249.3
- `Advertising expenditure`	1	209096	393947	1323.7

Step: AIC=1197.85

`Luxury Goods consumption` ~ `Per capita income (nominal)` +
`Advertising expenditure` + `Recession/ Boom` + `Consumer Price Index` +
`Exchange rate (INR against Euro)` + `Social Media Presence` +
`Imports (Non- Oil)`

	Df	Sum of Sq	RSS	AIC
- `Exchange rate (INR against Euro)`	1	1970	188241	1197.6
<none>			186272	1197.8
- `Imports (Non- Oil)`	1	4818	191090	1200.1
- `Per capita income (nominal)`	1	11031	197303	1205.5
- `Recession/ Boom`	1	12149	198421	1206.5
- `Social Media Presence`	3	37415	223686	1222.6
- `Consumer Price Index`	1	68323	254594	1248.3
- `Advertising expenditure`	1	211458	397729	1323.3

Step: AIC=1197.61

`Luxury Goods consumption` ~ `Per capita income (nominal)` +
`Advertising expenditure` + `Recession/ Boom` + `Consumer Price Index` +
`Social Media Presence` + `Imports (Non- Oil)`

	Df	Sum of Sq	RSS	AIC
<none>			188241	1197.6
- `Imports (Non- oil)`	1	5522	193764	1200.5
- `Per capita income (nominal)`	1	10205	198446	1204.5
- `Recession/ Boom`	1	11340	199581	1205.4
- `Social Media Presence`	3	35868	224110	1220.9
- `Consumer Price Index`	1	70868	259110	1249.3
- `Advertising expenditure`	1	211643	399884	1322.2

Call:
lm(formula = `Luxury Goods consumption` ~ `Per capita income (nominal)` +
`Advertising expenditure` + `Recession/ Boom` + `Consumer Price Index` +
`Social Media Presence` + `Imports (Non- Oil)`)

Coefficients:

(Intercept)	-2.802e+02	`Per capita income (nominal)`	5.254e-03
`Advertising expenditure`	3.644e+00	`Recession/ Boom`	-2.493e+01
`Consumer Price Index`	1.847e+00	`Social Media Presence`	Highly Strong
`Social Media Presence`	No influence		-1.247e+02
	3.525e+01	`Social Media Presence`	Strong
			-8.152e+01
`Imports (Non- oil)`	-1.760e-04		

Here we compare the AIC values in order to determine the variables to exclude to make the model accurate.

Final reduced model

Luxury Goods Consumption = $-(2.802e+02) + (5.254e-03)$ Per Capita income + $(3.644e)$ Advertising expenditure – $(2.493e+01)$ Recession/ Boom [Yes] + $(1.847e)$ Consumer Price Index – $(1.247e+02)$ Social Media presence [Strong] + $(3.525e+01)$ Social Media presence [No influence] – $(8.152e+01)$ Social Media [Strong] – $(1.760e-04)$ Imports

Conclusion

Hence we can check that demand for luxury goods is dependent on per capita income of people, advertisements & publicities done by respective companies to establish “brand image”, economy’s condition, spending habits and social media involvements.

However, our model has the following limitations:

1. Auto-correlation factor is not completely ruled out and hence ARIMA model or so could give better inferences.
2. Luxury goods consumption and advertisements expenditures have been considered for a single company (Shoppers Stop Ltd) and taken as a proxy for the population as a whole. The inference could differ if we could take average expenditures for all the companies which sells a reputable

brand. But then brand itself is a subjective factor and what is a brand to one may not be to the other and hence due to this subjectivity model is not comprehensive totally

3. A limited timeframe of 2006 to 2019 is considered (since India has its history of brand spending s kicking off from year 2006 in substantial numbers and hence years' prior do not really hold much significance)

In spite of above limitations, model do gives us a fair view of the factors that impact the demand for branded products in India.

References & Sources

Data collection:

1. CPI- <http://www.epwrfits.in/>
2. 10-year T bill rate- <https://data.imf.org/regular.aspx?key=61545867>
3. Currency- Oanda site
4. Per capita income- api.worldbank.org › indicator › NY.GDP.PCAP.CD › downloadformat=excel
5. GDP- <https://data.gov.in/resources/year-wise-gross-domestic-product-gdp-2004-05-2017-18-ministry-statistics-and-programme>
6. Import data- <http://www.epwrfits.in/ExternalSectorDisplayData.aspx>
7. Youth as a % of total population- http://mospi.nic.in/sites/default/files/publication_reports/Youth in India-2017.pdf
8. Papers from Researchgate- Author: Kimberly Caserta "Luxury Goods Demand" <https://pdfs.semanticscholar.org/e912/bafe7738e3f454d6a284f2b73a5dd59949cb.pdf>
9. Data pertaining to Recession and Social Media have been simulated
10. Also, yearly data was obtained from the above sites and monthly contribution have been simulated keeping in mind the yearly numbers