

Predicting Income and Feasible Loan Amount for a Household Unit (Expenditure Analysis of Badulla District, Sri Lanka)

K.W.S.N. Kumari and H.L.D.K. Jayarathna

Faculty of Science and Technology, Uva Wellassa University of Sri Lanka, Passara Road, Badulla, Sri Lanka.

Abstract

Expenditures of a household are growing due to many reasons which came up with complexity of basic needs and downturn of economy. Emerging body of knowledge suggest the necessity of expenditure studies in Sri Lanka. It is difficult to carry out an island wide study. Therefore Badulla was selected since it covers variety of living conditions in a limited extent. Six divisional secretariats were selected randomly and those were Badulla, Passara, Uva-Paranagama, Haldummulla, Soranathota and Meegahakivula. Respondents (household heads) were selected by systematic random sampling. They were interviewed by a structured questionnaire. 95% confidence intervals were calculated for mean costs of basic needs. These means were compared by living areas using one way ANOVA. Result implies that living area can influence on amount of cost. The main objective of this study is predicting total income and feasible loan amount per a household unit. Multiple regression analysis showed 70.9% of the variation of total income was accounted by the explanatory variables. Cost for foods, cost for transport and cost for clothing had a significant ($p\text{-value}<0.05$) effect on total income. Moreover, fitted model for predicting feasible loan amount explained 87.0% of total variation. Cost for electricity, cost for water, cost for education, cost for health and cost for social activities were significant ($p\text{-value}<0.05$) variables of the best fitted model. It can be concluded that multiple linear regression model which was fitted for predicting feasible amount of loan performs better than fitted model for predicting total income. This would be useful for administrative divisions and financial sector. Because before a person go for a higher amount of loan it can be explained by taking basic details of him or her what is the feasible amount of loan with his or her economic status. It would be good for loan receiver and loan providers.

Key Words: Expenditure, household income, stratified random sampling, multiple linear regression

DOI: 10.7176/DCS/9-6-11

Publication date: June 30th 2019

1. Introduction

The slogan “No man is rich whose expenditure exceeds his means; and no one is poor whose incoming exceed his outgoings” was said by famous author and great politician Thomas Chandler Haliburton and it is still valid even after several decades. As a citizen of a developing country we should be more aware of managing income and expenditures in a proper manner.

Income and expenditure are major concerns of the study as they currently affect to Sri Lankan living conditions with existing economic downturn. The study attempts to analyze and predict income and feasible financial assistance (loan) of a household in Badulla. It was difficult to select a study area that represents all living conditions. But Badulla district was recognized as most appropriate area to carry out the study. If once you have visited to Passara via Badulla-Colombo road you will feel the variation of living conditions from Haldummulla.

Badulla is the administrative district of Uva province covers 2,861 km² areas by land and 837,000 people live within this area. The main economy of district is based on agricultural farming and livestock. Badulla is partitioned in to two regions by climatic and geographical characteristics as upper and lower. Upper region is famous for tea plantation and vegetable cultivation while lower region is prominent for paddy cultivation.

According to reports of Department of Census and Statistics, Sri Lanka, the minimum expenditure per person per month to fulfill the basic needs is Rs. 4659 in average and it is Rs. 4565 (<Rs 4659) for Badulla district. The poverty head count ratio (percentage of proportion of a population below the poverty line) of Badulla was decreased from 13.3% to 12.3% between census studies of 2009/10 and 2012/13. The average values of those

two studies were 8.9 and 6.7 respectively. All these information concludes that economic level of Badulla is not adequate even to approach average level of Sri Lankan economy. Under these circumstances this study attempts to study how expenditure patterns behave in different living conditions of Badulla. Moreover, household income and feasible loan amount for a household are predicted using multiple linear regression. There are fifteen divisional secretariats in Badulla district. Among of them six divisional secretariats were randomly selected. They were Badulla, Passara, Uva-Paranagama, Haldummulla, Soranathota and Meegahakivula. Number of respondents which should be taken by each divisional secretariat was decided proportionately to sample size by stratified random sampling. Then structured questionnaires were distributed among respondents (household heads) who were selected using systematic random sampling. Figure 1.1 shows images of visited areas of Soranathota, Passara and Haldummulla respectively (left to right).

It is necessary to mention that Tamil translated questionnaire was used in estate areas with people those who are weak in Sinhala communication. Collected data was entered into Excel sheets and data analysis was carried out by SPSS 22 statistical software (Allen et al. 2014).

2. Literature Review

There were several literature found that has been completed expenditure analysis. Majority of studies are based on specific state or town in a developing country (Breuer et al. 2010). There was not a past literature based on expenditure analysis in Badulla district, Sri Lanka. According to a study of Lagos state, Nigeria, household income, tribe, household size and composition of household were significant factors that effect on food expenditure (Babalola et al. 2014).

Change of the expenditure patterns have been studied in Haryana state, India. There are four main factors that increase consumption expenditure. They are increase in urbanization, breaking up of the traditional joint family system, desire for quality food and lack of time which translates in to an increased need for convenience. The above mentioned study reveals that the factors such as advertisement, packaging and nutritive values has influenced the consumption behavior of the households to a greater extend (Mor et al. 2014).

There were some studies in literature which had investigated the role of woman towards change in household expenditure. This study attempts to explore the dynamic changes in status of the household, due to wives earnings. Research work was carried out in newly developed city of Udham Singh Nagar in Kumaon region of Uttarakhand. U.S.Nagar was chosen purposively as it truly represents the geography and economy of the state. According to study, average family expenditure on food items is around Rs 4000 per month for both types of households, where income is augmented by household participation of women in family income or not (Kumar et al. 2014).

Some expenditure analysis was mainly focused on life after retirement. A study was found that target group was USA retired citizens. This study attempts to show how household spending along with its different components changes with age. Household spending data is used from a nationally representative survey of Americans 50 years or older. The key finding is household spending is lower for older households, but the decrease is not uniform across the spending distribution. Between age groups 50–64 and 90 or older, the median drops nearly 55 percent, but the 90th percentile drops only 35 percent (Benerjee, 2014).

3. Methodology

The impact of this study should be understood by government administrates and financial companies. Therefore analysis part did not perform intensely since it should be clear to understand by the public (non-statisticians). The data was collected using a structured questionnaire and systematically selected household head was taken as respondent of the study (Babalola et al. 2014). The structured questionnaire was translated in Tamil language also for making efficient answers. A Tamil to Sinhala translator was also in field for accuracy of collected data. Name and their corresponding address were not collected since confidentiality is highly accepted throughout study.

The structured questionnaire covered details in living area, income, financial assistance and costing for basic needs (Mor et al. 2014). Then collected data was entered into excel sheets and they were imported to SPSS statistical software (Allen et al. 2014) for data analysis. At the first step of analysis, pie charts and bar charts were made to take a fundamental idea about behavior of data. Then summary tables and tables with descriptive measurements were tabulated to identify numerical behavior of counts and descriptive statistics across categorical variables.

The data was checked for normality before performing regression analysis. Anderson Darling test was used to check normality of the independent and dependent variables. The equation of Anderson Darling Test is given by equation 1.

$$S - \sum_{i=1}^N \frac{(2i-1)}{N} [\ln F(Y_i) + \ln(1 - F(Y_{N+1-i}))] \quad (1)$$

At the complementary analysis, multiple linear regression was used to predict total income and feasible amount of loan (Kumar et al. 2014). The format of used regression model is given by equation 2,

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_{10} X_{10} + \varepsilon_i \quad (2)$$

Where, Y_1 is total amount of income and Y_2 is feasible amount of loan. X_1, X_2, \dots, X_{10} are cost for foods, cost for electricity bills, cost for water bills, cost for education, savings, cost for health, cost for social activities, cost for transport, cost for clothing and cost for paying interest of loans respectively. $\beta_0, \beta_1, \beta_2, \dots, \beta_{10}$ represent parameter coefficients of constant, X_1, X_2, \dots, X_9 and X_{10} respectively. ε_i is the error term of the i^{th} regression model ($i=1,2$).

ANOVA table was used to check model significance. Then significance parameters were selected using summary table of coefficients and p-values were used to check significance (Benerjee, 2014). ANOVA table was also used in mean comparisons. Mean costs for basic needs were compared by living areas (urban, rural, estate). The analyzed results are discussed under section 4.

4. Results and Discussion

4.1. Major source of income

It is necessary to understand about what are the major occupations that people are involving in Badulla. The jobs were characterized as government, farming, laboring, business and pension. Majority of study are involving in laboring as their major income (53.57%). Those who are engaging in business and farming are equal by percentage as second highest that is 14.29%. 10.71% of respondents are in government sector and about 7.14% of people have only pension as the income.

Majority of Badulla district are involving other source of income also. Most significant combination of major and other source of income is Labor and farming. In some families both of family members are in laboring. There were some farmers who engage in business while cultivating. For an instance, a farmer was met that he sells his milk farming products in a small stall in Meegahakivula (Badulla-Kandy road).

People who are in each and every level of occupation are expecting financial assistance by government sector or private companies. Table 4.1.1 gives 95% confidence interval for average paying interest in rupees.

Table 4.1.1 shows that highest amount of interest is paid by people who occupy in government jobs. They can easily afford for a loan since they have fixed income to pay interest monthly. Labors are tending to get more loans than those who are in business and farming. Labors are in a huge debt because they are struggling with micro finance loans.

4.2. Effect of sub regions (divisional secretariats)

Proportion of expenditure for a basic need will not only depend on their occupations. Infrastructure facilities, easy access to job market, good standard of education, quality of health service and condition of transport system will influence for discrepancy of expenditure from region to region. It was convenient to recognize influence of sub regions by analyzing data across divisional secretariats.

Table 4.2.1 implies how average income is varying across divisional secretariats. The results reflect the dissimilarity of living conditions due to poor infrastructure facilities and public services.

There is a significant dissimilarity of income distribution when compare Badulla and Passara. A person who are live in Passara are not earning even a half of amount earn that by person in Badulla region. Meegahakivula and Uva-Paranagama are competing to income of Badulla as second and third by average income. Soranathota, Haldummulla and Passara do not exceed average income limit of Rs 40,000. The administrative should take immediate actions for improving living conditions of these three divisional secretariats.

The amount of received loan was analyzed by categorizing under divisional secretariats. The same order could be seen as in average income. Those who earn highly are tending to get higher amount of loan. But people in Passara are in 4th rank by amount of loan though they are 6th by average income.

4.3. Costs for basic needs

Living areas of Badulla were characterized under three categories. They are estate, rural and urban living areas. Costing of basic needs have been analyzed by comparing living areas. Table 4.3.1. presents 95% confidence intervals for average cost of foods, electricity, education, health and transport respectively.

Table 4.3.1 shows how people spend their income for fulfilling their basic needs and results has been tabulated by living areas. It is obvious that major amount of income is spent on foods which it is higher in Urban and lower in estate. Electricity, education, health and transport costs are also highest in Urban while lowest in estate. People who are in estates are not spending considerable amount for health needs with respect to rural and urban area. People who are in rural area are spending significant amount for their children's education which can compete with urban people. Mean cost for basic needs (example: mean cost for foods) was compared by living areas. The One-way ANOVA test was used and all p-values were less than 0.05(α).

Cost for paying interest for loans could be seen as a main share of expenditure of people in Badulla district. Therefore 95% confidence intervals were calculated for mean of interest payment by considering living area. Table 4.3.2. gives confidence intervals for mean interest payment for a received loan.

Table 4.3.2 implies people who are in urban area are paying higher amount for interest monthly. The rural are paying approximately half of amount which is waged by the urban. The estate people are paying approximately half of amount which is paid by the rural.

4.4. Predicting income and feasible loan amount

Analysis of income and distribution of expenditure were major concerns of the study. Financial assistantship makes a major impact for expenditure patterns in Badulla. Therefore study was extended to analyze received loan and interest payments. Initially major objective of the study was to predict income by cost of basic needs of people. Moreover, the feasible loan amount that can be taken with cost of basic needs was predicted at the extended analysis. Persons who are flexible with interest payment were chosen to fit the model to predict affordable loan amount.

4.3.1. Predicting income

The ten explanatory variables were selected to fit model to predict income and three variables were significant among of them. The coefficient of determination was (R^2) 0.709 for fitted model. It means the 70.9% of total

variation can be explained by fitted model. Table 4.3.1.1 illustrates Analysis of Variance (ANOVA) table for income prediction model.

According to table 4.3.1.1 p-value of ANOVA test is 0.005 which is less than confidence level ($\alpha=0.05$). It implies that model is significant which at least one parameter coefficient is significant. The significant parameter coefficients were identified using sample test statistics and p-values were used to recognize significant parameter coefficients. Cost for foods, cost for transport and cost for clothing were significant response variables. The fitted model is given by equation 3,

$$\text{Total Income} = 1.595 * \text{Cost for foods} + 5.068 * \text{Cost for transport} + 3.819 * \text{Cost for clothing} \quad (3)$$

4.3.2. Predicting feasible loan amount

As in fitted model for income, same ten explanatory variables were selected to fit model to predict feasible loan amount and five variables were significant. The coefficient of determination (R^2) is 0.870 of fitted model. It means the 87.0% of total variation can be explained by fitted model. Table 4.3.2.1 illustrates Analysis of Variance (ANOVA) table for feasible loan amount prediction model.

P-value of table 4.3.2.1 is 0.000 which is less than confidence level ($\alpha=0.05$). It can be conclude that fitted model is significant and need to be recognized what are the significant model parameters. Each parameter was tested and p-values of test statistics were used to identify significant parameters. Cost for electricity, cost for water bills, cost for education, cost for health and cost for social activities were significant model parameters. Fitted model to predict feasible loan amount is given by equation 4,

$$\text{Total loan amount} = 242.707 * \text{Cost for electricity} - 1287.248 * \text{cost for water bills} - 32.351 * \text{cost for education} - 95.573 * \text{cost for health} + 233.200 * \text{cost for social activities} \quad (4)$$

5. Conclusion

There are two major conclusions of this study. First one is multiple regression model can be fitted to predict total income with 70.9% model accuracy. The response variables of fitted model were cost for foods, cost for transport and cost for health.

Second is fitted multiple regression to predict feasible loan amount can explain 87.0% of total variation. Cost for electricity, cost for water, cost for education, cost for health and cost for social activities are significant response variables of fitted model. It can be concluded that loan amount can be predicted more accurately than income per a household unit.

Badulla divisional secretariat has highest average by income and received loan amount. Passara divisional secretariat has lowest average income. Highest proportion of income is allowed for cost for foods. Those who are in laboring as major income are more likely to do farming as other income method. The mean cost for basic needs were differed from living conditions and it implies that living area makes an influence on expenditure patterns. Cost for health in Meegahakivula divisional secretariat is comparatively lower than expenditures for other needs. It is necessary to take actions to improve their health condition and need some programs to inform about benefit of having good health. The overall study implies that Passara divisional secretariat is a significant spot since poor economic status. It is lower by each measuring factor. Therefore this region should be paid more attention by government administrates.

The current study was conducted across randomly selected 6 divisional secretariats and it is expected to expand as it covers all divisional secretariats and it will be lead for more accurate summary about Badulla district.

References

- Allen, P., Bennett, K., & Heritage, B. (2014), “SPSS statistics version 22: A practical guide”, Cengage Learning Australia.
- Babalola, D.A. & Isitor, S.U. (2014), “Analysis of the determinants of food expenditure patterns among urban households in Nigeria: Evidence from Lagos State”, *IOSR Journal of Agriculture and Veterinary Science*, 7(5), 71-75.
- Breuer, C., Hallmann, K., Wicker, P. & Feiler, S. (2010), “Socio-economic patterns of sport demand and ageing”, *European Review of Aging and Physical Activity*, 7(2), 61.
- Mor, K. & Setia, S. (2014), “Changing Consumption expenditure pattern of Haryana: A case study of Ambala District, India”, *Acme Intellectuals International Journal of Research in Management*, 7(7), 1-12.
- Banerjee, S. (2014), “How Does Household Expenditure Change with Age for Older Americans?”.
- Kumar, P., Singhal B. & Jalal, R.S. (2014), “Change in household expenditure pattern due to women’s participation in income augmentation activities (A cased study of District U.S. Nagar, Uttarakhand)”, *The International Journal of Humanities and Social Studies*, 2(12), 186-191.



Figure 1.1. Images of estate housing schemes in Soranathota (Keenakele), Passara (Ury estate) and Haldummulla (Wiharagala estate)

Table 4.1.1. 95% Confidence Intervals for average paying interest.

Major Source of Income	95% Confidence Interval for average paying interest
Business	(5750± 1.96*4750.00)
Farming	(5833.33±1.96*1301.78)
Government	(42500.00±1.96*14500.00)
Labour	(7155.50±1.96*1872.91)
Pension	(32000.00±1.96*102.50)

Table 4.2.1. 95% confidence intervals for average income

Divisional Secretariat	95% confidence interval for average income (main income)
Badulla	(46600.00±1.96*8249.85)
Meegahakivula	(43333.33±1.96*9276.61)
Uva-Paranagama	(41600.00±1.96*14868.09)
Haldummulla	(35450.00±1.96*8495.73)
Soranathota	(30000.00±1.96*5244.04)
Passara	(21100.00±1.96*3906.405)

Table 4.3.1. 95% confidence intervals for mean cost of basic needs (based on living areas)

	Estate	Rural	Urban	P-value(ANOVA)
Foods	(14785.71±1.96*1950.426)	(16666.67±1.96*3574.602)	(20625.00±1.96*2203.386)	0.019
Electricity	(419.29±1.96*100.628)	(875.00±1.96*325.256)	(1325.00±1.96*245.495)	0.006
Education	(3357.14±1.96*1963.095)	(6500.00±1.96*2247.221)	(7400.00±1.96*2293.469)	0.004
Health	(1425.00±1.96*257.354)	(3625.00±1.96*554.339)	(3812.50±1.96*2403.378)	0.009
Transport	(1535.38±1.96*3574.602)	(2125.00±1.96*707.784)	(4812.50±1.96*954.021)	0.003

Table 4.3.2. 95% confidence intervals for average paying interest

Living area	95% confidence intervals for average paying interest
Estate	(5600.00 ± 1.96*1016.764)
Rural	(10166.67 ± 1.96*5449.261)
Urban	(20100.00 ± 1.96*7582.279)

Table 4.3.1.1. ANOVA table for fitted model to predict income

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	13991215087.9	10	1399121508.7	46.1422	.005
Residual	5730845179.93	189	30321932.17		
Total	19722060267.8	199			

Table 4.3.2.1. ANOVA table for fitted model to predict feasible loan amount

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	13504535665241.352	10	1350453566524.135	211.4599	.000 ^b