PREDICTIVE MODEL OF TOTAL INCOME FROM SALARIES/WAGES IN THE CONTEXT OF PASAY CITY

Wilson Cordova
wilson.cordova@cksc.edu.ph
Chiang Kai Shek College, Philippines

Abstract

There are varied sources of income (e.g. entrepreneurial, salary/wages, agricultural, aquatic, manufacturing, etc.). Based from the CBMS data of Pasay City, about 30% are engaged or dependent on salaries/wages compared to other income generating activities.

The objective of this study was to establish a predictive model of total income from salaries/wages based on some indicators. The basis of selection of the predictors was derived from prior knowledge of the proponent since most of the studies available were focused on demographic relations to income.

Keywords: Entrepreneurship, Salary and Wages.

Hypotheses

1. There is a significant relationship between the total income from salaries/wages and predictors such as:
   a. Average water consumption;
   b. Electric bill;
   c. Type of house;
   d. Tenure status of house/lot;
   e. Type of toilet facility;
   f. Type of water facility;
   g. Construction materials used in roof; and
   h. Construction materials used in walls;

2. There is no significant relationship that could be established between the total income from salaries/wages and the above-mentioned predictors.

Method

The data used in this study were obtained from the Community-Based Monitoring System (CBMS) database. The Community-Based Monitoring System is one of the tools developed in the early 1990s under the Micro Impacts of Macroeconomic Adjustment Policies (MIMAP) Project -Philippines to provide policymakers and program implementers with a good information base for tracking the impacts of macroeconomic reforms and various policy shocks.

The CBMS work in the Philippines evolved after it was observed that there were no disaggregated data in place for planning, program formulation, policy impact, and poverty monitoring. There was also a need for support mechanisms for the implementation of the decentralization policy. The CBMS is envisioned to serve as a tool for local governance and complement the national poverty monitoring system. The CBMS can likewise facilitate the implementation of targeted poverty reduction programs with its household and individual level data, as well as the monitoring and evaluation of these poverty reduction programs.

Today, the CBMS has been adopted by the National Anti-Poverty Commission (NAPC) and the Department of Interior and Local Government (DILG) as the local poverty monitoring system and as a tool for localizing the Millennium Development Goals (MDGs) in the country. Efforts are underway with the help of national government agencies, local government units (LGUs), non-government organizations (NGOs) and donor agencies to scale up the implementation of the CBMS ("CBMS Philippines," 2012, par. 3-5).

For this study, the CBMS data for 2011 of the City of Pasay were used. The City of Pasay is one of the cities and municipalities that make up Metro Manila in the Philippines.

Pasay City covers a total land area of 18.50 square kilometers making it the third smallest political subdivision in the National Capital Region. It borders City of Manila to the north, Parañaque to the south, Makati and Taguig to the northeast, and Manila Bay to the west. The city can be divided in 3 distinct areas: the city's urban area with an area of 5.505 square kilometers; the Civil Aeronautics Administration (CAA) complex, which includes the Ninoy Aquino International Airport (NAIA) and the Villamor Airbase, with an area of 9.5 square kilometers; and the reclaimed land from Manila Bay with an area of 4.00 square kilometers.

Pasay is composed of seven (7) districts, subdivided into twenty (20) zones, with a total of 201 barangays. The barangays do not have names but are only designated with sequential numbers. The largest zone with an area of 5.10 square kilometers is Zone 19, which covers barangays 178 and 191. The smallest zone with an area of 10 hectares is Zone 1, covering Barangays 1 to 3 and 14 to 17 ("Pasay," 2012, "Geography," para. 1-3).

The CBMS data were conducted among the residents of the city. For this study, only data of the respondents who successfully answered the chosen variables were used. Thus, the number of respondents was reduced to 67,622.

The variables used in this study include the following:

Variable Name	**Variable Label**
Dependent wagcsh	 Total income from salaries/wages
Predictors ave_water	 Average water consumption
elec_bill	Electric bill
house_type1 • _Ihouse_ty~2 • _Ihouse_ty~3	Type of house (Dummy variable) (Dummy variable)
tenur1 • _Itenur1_2 • _Itenur1_3	Tenure status of house/lot (Dummy variable) (Dummy variable)

ISBN: 978 0 994 3656 68
www.apiar.org.au

toil1 • _Itoil1_2 • _Itoil1_3	Type of toilet facility (Dummy variable) (Dummy variable)
water1 • _Iwater1_2 • _Iwater1_3	Type of water facility (Dummy variable) (Dummy variable)
roof1 • _Iroof1_2 • _Iroof1_3	Construction materials used in roof (Dummy variable) (Dummy variable)
wall1 • _Iwall1_2 • _Iwall1_2	Construction materials used in walls (Dummy variable) (Dummy variable)

Categorical predictors were grouped and recoded as follows:

Typed of Water Facility (Grouped)	Freq.	Percent	Cum.
Piped Water	29,155	43.11	43.11
Bottled/Mineral Water	37,345	55.23	98.34
Others	1,122	1.66	100.00
Total	67,622	100.00	

Construction Materials of Walls (Grouped)	Freq.	Percent	Cum.
Strong materials	46,017	68.05	68.05
Light/Salvaged materials	6,035	8.92	76.97
Mixed materials	15,570	23.03	100.00
Total	67,622	100.00	

Construction Materials of Roof (Grouped)	Freq.	Percent	Cum.
Strong materials	45,345	67.06	67.06
Light/Salvaged materials	5,663	8.37	75.43
Mixed materials	16,614	24.57	100.00
Total	67,622	100.00	

Type of Toilet Facility (Grouped)	Freq.	Percent	Cum.
Water sealed flush- owned	44,895	66.39	66.39
Water sealed flush- shared	20,410	30.18	96.57
Others	2,317	3.43	100.00
Total	67,622	100.00	

House Type (Grouped)	Freq.	Percent	Cum.
Single/Duplex	48,928	72.36	72.36
Apartment/Condominium	15,724	23.25	95.61
Others	2,970	4.39	100.00
Total	67,622	100.00	

Tenure Status of House/Lot (Grouped)	Freq.	Percent	Cum.
Owned	16,900	24.99	24.99
Rent	27,087	40.06	65.05
Others	23,635	34.95	100.00
Total	67,622	100.00	

Alongside, dummy variables were introduced on these categorical predictors. Each of the first categories among these predictors was naturally omitted. The omitted category becomes the reference category against which the effects of the other categories are assessed. Then, results can be interpreted as the difference between each category and this omitted category.

In the data analysis, the study used multivariate approach particularly multiple linear regression analysis. It is a method for measuring effect of predictors to explain the variation in the dependent variable. A multiple linear regression model has been estimated to determine the best linear combination for predicting the total income from salaries/wages. The data analysis was carried out using Stata. Stata is a general-purpose statistical software package created in 1985 by StataCorp. It is used by many businesses and academic institutions around the world. Most of its users work in research, especially in the fields of economics, sociology, political science, biomedicine, and epidemiology. Stata's capabilities include data management, statistical analysis, graphics, simulations, and custom programming ("Stata," 2012, "Definition," par.1).

Results and discussion

Based on the output of the regression analysis, one can assess if the overall model is significant. With the p-value of the F-test equal to zero (up to four decimal places), it can be thought that the model is statistically significant. The R-squared is 0.0981, meaning that approximately 9.81% of the variability of the dependent variable, ***wagcsh*** is accounted for by the variables in the model. In this case, the adjusted R-squared indicates that about 9.79% of the variability of ***wagcsh*** is accounted for by the model; even after taking into account the number of predictor variables in the model. Compared to the regression model without dummies, the R-squared is higher by 0.3 percentage points.

```
      Source |       SS       df       MS              Number of obs =   67622
-------------+------------------------------           F( 14, 67607) =  525.03
       Model |  1.6305e+14    14  1.1646e+13           Prob > F      =  0.0000
    Residual |  1.4996e+15 67607  2.2182e+10           R-squared     =  0.0981
-------------+------------------------------           Adj R-squared =  0.0979
       Total |  1.6627e+15 67621  2.4588e+10           Root MSE      =  1.5e+05
```

The coefficients for each of the variables indicate the amount of change one could expect in the dependent variable ***wagcsh*** given a one-unit change in the value of that variable, assuming that all other variables in the model are held constant. Taking for example the predictor ***ave_water***, assuming that all other variables in the model are held constant, there would be an increase by 45.17 in the total income from salaries/wages for every one unit increase in this variable. Likewise for the variable ***elec_bill***, for every one unit in this variable, there will be an increase of 24.71 in the total income from salaries/wages.

wagcsh	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
ave_water	45.17007	2.087749	21.64	0.000	41.07808	49.26205
elec_bill	24.71173	.6533384	37.82	0.000	23.43119	25.9922

To compare the strength of one coefficient obtained from the output of regression analysis to the coefficient for another variable, the results included the computed beta coefficients which will give the standardized regression coefficients. The beta coefficients are used by some researchers to compare the relative strength of the various predictors within the model. Because the beta coefficients are all measured in standard deviations, instead of the units of the variables, they can be compared to one another. In other words, the beta coefficients are the coefficients that one could obtain if the outcome and predictor variables were all transformed in standard scores, also called z-scores, before running the regression. Because the coefficients in the Beta column are all in the same standardized units, one can compare these coefficients to assess the relative strength of each of the predictors. The predictor ***elec_bill*** showed the largest beta coefficient which is 0.1633. This means that a one standard deviation increase in ***elec_bill*** would yield a 0.1633 increase in the predicted ***wagcsh***. Likewise for the predictor ***ave_water***, one standard deviation increase in this variable would produce a 0.0895 increase in the predicted ***wagcsh*** assuming all the other variables to be held constant.

wagcsh	Coef.	Std. Err.	t	P>|t|	Beta
ave_water	45.17007	2.087749	21.64	0.000	.0895356
elec_bill	24.71173	.6533384	37.82	0.000	.1632717

The Beta values for the ***house_type1*** dummy variable indicated that respondents with a house type of ***_Ihouse_ty~2* (Apartment/Condominium)**, have incomes that were about 0.04 more than respondents with a house type of ***Single/Duplex* (reference category)**. On the other hand, ***_Ihouse_ty~3* (Commercial/Industrial/Agricultural building, etc.)** have incomes that were about 0.02 less than respondents with a house type of ***Single/Duplex***. For the ***roof1*** dummy variables showed that respondents with **light/salvaged materials used for the construction of roof (*_Iroof1_2*)** and **mixed materials used for the construction of roof (*_Iroof1_3*)** earned incomes that were about 0.009 less than respondents with ***strong materials used for the construction of roof* (reference category).** The T-test results revealed that the difference between these dummy variables from our reference category were insignificant.

wagcsh	Coef.	Std. Err.	t	P>|t|	Beta
_Ihouse_ty~2	14397.2	1392.818	10.34	0.000	.0387869
_Ihouse_ty~3	-14816.84	3244.952	-4.57	0.000	-.0193631
_Iroof1_2	-5128.36	2948.4	-1.74	0.082	-.0090595
_Iroof1_3	-3220.48	2043.058	-1.58	0.115	-.0088415

Taking all the results into account, the predictive model for the dependent variable ***wagcsh* (Total income from salaries/wages)** expressing in terms of the predictors would be:

wagcsh = 84,244.76 + 0.0895(***ave_water***) + 0.1633(***elec_bill***) + 0.0388(***_Ihouse_ty~2***) – 0.0194(***_Ihouse_ty~3***) + 0.0421(***_Itenur1_2***) – 0.0228(***_Itenur1_3***) – 0.0796(***_Itoil1_2***) – 0.0222(***_Itoil1_3***) + 0.0832(***_Iwater1_2***) – 0.0005(***_Iwater1_3***) – 0.0091(***_Iroof1_2***) – 0.0088(***_Iroof1_3***) – 0.0238(***_Iwall1_2***) – 0.0295(***_Iwall1_3***).

Conclusion

The study has presented a predictive model regarding total income from salaries/wages, taking into account some of the social indicators. The results of the study found out that average water consumption, electric bill, type of house, tenure status of house/lot, type of toilet facility, type of water facility, construction materials used in roof, and construction materials used in walls were statistically significant.

References

[1] CBMS Philippines. Retrieved December 8, 2012, from http://cbmsphilippines.webs.com/

[2] Pasay. (n.d.). In *Wikipedia*. Retrieved December 8, 2012, from http://en.wikipedia.org/wiki/Pasay

[3] Stata. (n.d.). In *Wikipedia*. Retrieved December 8, 2012, from http://en.wikipedia.org/wiki/Stata

[4] Regression with Stata. UCLA: Statistical Consulting Group. From http://www.ats.ucla.edu/stat/stata/webbooks/reg/chapter1/statareg1/ (accessed December 8, 2012).

[5] Anderson, R. E., Black, W. C., Hair, J. F. Jr., Tatham, R. L. (1998). *Multivariate data analysis* (5th ed.). Upper Saddle, New Jersey: Prentice Hall.

Appendixes

Table 1: Summary of Data

variable name	storage type	display format	value label	variable label
ave_water	double	%10.0g		Average Water Consumption
elec_bill	double	%10.0g		Electric Bill
house_type1	float	%21.0g	house_type1	House Type (Grouped)
tenur1	float	%9.0g	tenur1	Tenure Status of House/Lot (Grouped)
toil1	float	%26.0g	toil1	Type of Toilet Facility (Grouped)
water1	float	%21.0g	water1	Typed of Water Facility (Grouped)
roof1	float	%24.0g	roof1	Construction Materials of Roof (Grouped)
wall1	float	%24.0g	wall1	Construction Materials of Walls (Grouped)
wagcsh	double	%10.0g		Total income from salaries/wages (cash)

Table 2: Summary Statistics of Data

Variable	Obs	Mean	Std. Dev.	Min	Max
ave_water	67622	363.1653	310.8211	0	1900
elec_bill	67622	1033.623	1036.032	0	6000
house_type1	67622	1.320369	.5527915	1	3
tenur1	67622	2.099598	.7678047	1	3
toil1	67622	1.370353	.5492942	1	3
water1	67622	1.585446	.52525	1	3
roof1	67622	1.575123	.8577563	1	3
wall1	67622	1.549747	.8414492	1	3
wagcsh	67622	132364.2	156807	0	1182000

Table 3: Regression Results without Dummy Variables

Source	SS	df	MS
Model	1.5098e+14	8	1.8873e+13
Residual	1.5117e+15	67613	2.2358e+10
Total	1.6627e+15	67621	2.4588e+10

Number of obs	= 67622
F(8, 67613)	= 844.09
Prob > F	= 0.0000
R-squared	= 0.0908
Adj R-squared	= 0.0907
Root MSE	= 1.5e+05

wagcsh	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ave_water	44.43169	2.092413	21.23	0.000	40.33057	48.53282
elec_bill	24.76442	.6507863	38.05	0.000	23.48888	26.03996
house_type1	7806.657	1066.535	7.32	0.000	5716.25	9897.063
tenur1	-6027.509	783.6311	-7.69	0.000	-7563.425	-4491.592
toil1	-25055.01	1155.936	-21.68	0.000	-27320.65	-22789.38
water1	22912.43	1112.981	20.59	0.000	20730.99	25093.87
roof1	-2185.286	998.5327	-2.19	0.029	-4142.409	-228.1633
wall1	-6906.706	1026.07	-6.73	0.000	-8917.801	-4895.61
_cons	105132.3	3278.07	32.07	0.000	98707.27	111557.3

Table 4: Regression Results with Dummy Variables

```
i.house_type1     _Ihouse_typ_1-3     (naturally coded; _Ihouse_typ_1 omitted)
i.tenur1          _Itenur1_1-3        (naturally coded; _Itenur1_1 omitted)
i.toil1           _Itoil1_1-3         (naturally coded; _Itoil1_1 omitted)
i.water1          _Iwater1_1-3        (naturally coded; _Iwater1_1 omitted)
i.roof1           _Iroof1_1-3         (naturally coded; _Iroof1_1 omitted)
i.wall1           _Iwall1_1-3         (naturally coded; _Iwall1_1 omitted)
```

Source	SS	df	MS
Model	1.6305e+14	14	1.1646e+13
Residual	1.4996e+15	67607	2.2182e+10
Total	1.6627e+15	67621	2.4588e+10

Number of obs	=	67622
F(14, 67607)	=	525.03
Prob > F	=	0.0000
R-squared	=	0.0981
Adj R-squared	=	0.0979
Root MSE	=	1.5e+05

wagcsh	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ave_water	45.17007	2.087749	21.64	0.000	41.07808	49.26205
elec_bill	24.71173	.6533384	37.82	0.000	23.43119	25.99227
_Ihouse_ty~2	14397.2	1392.818	10.34	0.000	11667.28	17127.12
_Ihouse_ty~3	-14816.84	3244.952	-4.57	0.000	-21176.94	-8456.733
_Itenur1_2	13466.77	1530.541	8.80	0.000	10466.91	16466.63
_Itenur1_3	-7493.75	1578.559	-4.75	0.000	-10587.72	-4399.776
_Itoil1_2	-27200.31	1338.063	-20.33	0.000	-29822.91	-24577.71
_Itoil1_3	-19132.81	3767.777	-5.08	0.000	-26517.65	-11747.97
_Iwater1_2	26236.77	1204.172	21.79	0.000	23876.59	28596.95
_Iwater1_3	-591.5475	4700.021	-0.13	0.900	-9803.585	8620.49
_Iroof1_2	-5128.36	2948.4	-1.74	0.082	-10907.22	650.5016
_Iroof1_3	-3220.48	2043.058	-1.58	0.115	-7224.873	783.9124
_Iwall1_2	-13063.09	2866.589	-4.56	0.000	-18681.6	-7444.572
_Iwall1_3	-10976.67	2106.153	-5.21	0.000	-15104.73	-6848.616
_cons	84244.76	1767.98	47.65	0.000	80779.52	87710

Table 5: Regression Results with Dummy Variables Showing Beta Coefficients

```
i.house_type1     _Ihouse_typ_1-3     (naturally coded; _Ihouse_typ_1 omitted)
i.tenur1          _Itenur1_1-3        (naturally coded; _Itenur1_1 omitted)
i.toil1           _Itoil1_1-3         (naturally coded; _Itoil1_1 omitted)
i.water1          _Iwater1_1-3        (naturally coded; _Iwater1_1 omitted)
i.roof1           _Iroof1_1-3         (naturally coded; _Iroof1_1 omitted)
i.wall1           _Iwall1_1-3         (naturally coded; _Iwall1_1 omitted)
```

Source	SS	df	MS
Model	1.6305e+14	14	1.1646e+13
Residual	1.4996e+15	67607	2.2182e+10
Total	1.6627e+15	67621	2.4588e+10

Number of obs	=	67622
F(14, 67607)	=	525.03
Prob > F	=	0.0000
R-squared	=	0.0981
Adj R-squared	=	0.0979
Root MSE	=	1.5e+05

wagcsh	Coef.	Std. Err.	t	P>|t|	Beta
ave_water	45.17007	2.087749	21.64	0.000	.0895356
elec_bill	24.71173	.6533384	37.82	0.000	.1632717
_Ihouse_ty~2	14397.2	1392.818	10.34	0.000	.0387869
_Ihouse_ty~3	-14816.84	3244.952	-4.57	0.000	-.0193631
_Itenur1_2	13466.77	1530.541	8.80	0.000	.0420832
_Itenur1_3	-7493.75	1578.559	-4.75	0.000	-.0227871
_Itoil1_2	-27200.31	1338.063	-20.33	0.000	-.079629
_Itoil1_3	-19132.81	3767.777	-5.08	0.000	-.0221955
_Iwater1_2	26236.77	1204.172	21.79	0.000	.0832018
_Iwater1_3	-591.5475	4700.021	-0.13	0.900	-.0004819
_Iroof1_2	-5128.36	2948.4	-1.74	0.082	-.0090595
_Iroof1_3	-3220.48	2043.058	-1.58	0.115	-.0088415
_Iwall1_2	-13063.09	2866.589	-4.56	0.000	-.0237508
_Iwall1_3	-10976.67	2106.153	-5.21	0.000	-.0294702
_cons	84244.76	1767.98	47.65	0.000	.