Estimating the Magnitude of the Poor Households in Metro Manila Using the Poisson Regression Model[1]

Bernadette B. Balamban, Anna Jean C. Pascasio, Driesch Lucien R. Cortel[2] and Maxine R. Ridulme

Poverty and Human Development Statistics Division, Social Sector Statistics Service, Philippine Statistics Authority

ABSTRACT

In the Official Poverty Statistics, Metro Manila, also known as the National Capital Region (NCR), is one of the areas that belong to the least poor cluster – a cluster that has relatively low poverty incidences. The Philippine Statistics Authority released the 2018 Municipal and City Level Poverty Estimates using the Elbers, Lajouw, and Lanjouw (ELL) methodology. The 2018 Small Area Estimates (SAE) of Poverty also released estimates for the 14 sub-areas in Metro Manila, ranging from 1.5 percent to 6.5 percent. The city-level poverty estimates were released as official statistics using direct estimation technique.

Given the relatively low poverty incidences of the region, this paper aims to estimate the 2018 poverty incidence for the legislative districts in NCR, including the 14 sub-areas of the City of Manila, using the Poisson regression methodology and to compare with the results of the ELL methodology. Data sources include the 2015 Census of Population, and the merged 2018 Family Income and Expenditure Survey and January 2019 round of the Labor Force Survey. A total of 5 significant indicators were included in the final model. Results show that the Poisson model produced more reliable estimates for NCR than the ELL methodology. These SAE techniques allow for generating more granular poverty statistics useful for targeting poor beneficiaries. Furthermore, information may provide an opportunity for the LGU to act swiftly and provide appropriate subsidies for areas within Metro Manila.

Keywords: poverty statistics, small area estimation, survey and census data, time invariance

1. INTRODUCTION

At the beginning of the previous administration in 2016, the Philippine government committed to the zero-to-ten-point socioeconomic agenda to achieve the *Ambisyon Natin* 2040, a vision that reflects the Filipino people's collective aspiration to end extreme poverty by 2040 [NEDA, 2021]. While the availability of suitable employment opportunities contributed to higher incomes and improved the living condition of low-income earners, various social protection programs have become drivers of economic activities, especially in the National Capital Region (NCR). The region has been the epicenter of economic movement and employment opportunities in the country. However, the region experienced the biggest economic shock caused by the global pandemic. Like most urban areas, NCR has suffered consistently high cases of patients with Covid-19, which led to strict area-wide lockdowns. Amidst the effort of the national government to sustain the economy, job losses and business closures have affected many Filipinos residing in the region.

In recent years, the country has been successful in finally starting to reduce the proportion of the poor. However, since the pandemic, the national poverty incidence among the population based on the

---

[1] Presented at the 15th National Convention on Statistics, 03-5 October, 2022, Crowne Plaza Manila Galleria, Ortigas Center, Quezon City and published in this issue as a *Technical Paper*.

[2] Corresponding Author: *d.cortel@psa.gov.ph*

national poverty line has significantly increased from 16.7 percent in 2018 to 18.1 percent in 2021, based on the recent release of the Philippine Statistics Authority [PSA, 2022], and this shows that the effect of the global pandemic has taken a toll on the poverty reduction efforts of the country from the previous years. Further, the 2021 poverty incidence among Filipinos in NCR was only 3.5 percent in 2021. However, this has already translated to almost half a million poor Filipinos, a significant increase compared to the 302 thousand poor Filipinos in 2018.

As supported by the findings of the Philippine Institute for Development Studies (PIDS) [Reyes et al., 2010], the increase in the number of poor caused by the pandemic may become a part of the chronic poor if they do not recover quickly. The social safety nets implemented by the government, in coordination with the national and local government agencies, are only temporary. Implementing effective local programs and their assistance are needed to help those who lost their jobs or closed their businesses bounce back from the adverse impact of the pandemic. With this, national government agencies, local government units, and civil society organizations demanded data availability at the lower-disaggregated levels towards data-driven decision-making and program implementation. Thus, this paper seeks to present the results of a study that generates the legislative district-level poverty statistics of Metro Manila. It also aims to compare the model-based estimates and official poverty releases.

Small-area methods attempt to solve the low representativeness of surveys within areas or the lack of data for specific areas/subpopulations, which are done by incorporating information from other sources that are becoming increasingly available such as geospatial information produced using remote sensing. The strength of these target datasets is their granularity on the subpopulations of interest. The resulting estimates can be used as information support for both the government and the private sector to guide them in allocating national and local government funds and planning for national and local government. Specifically, the paper will present the results of district-level poverty statistics generated through the Small Area Estimation (SAE) technique using the Poisson modeling approach and evaluate the poverty estimates of the districts in NCR.

2. METHODOLOGY

2.1 DATA SOURCES

This study utilized the survey data sets that came from the Family Income and Expenditure Survey (FIES), Labor Force Survey (LFS), and Census of Population (POPCEN), which were all conducted by the Philippines Statistics Authority (PSA). Specifically, the 2018 FIES and LFS January 2019 Round survey data sets and 2015 POPCEN for NCR were used. The official poverty estimates of the highly urbanized cities and one municipality in NCR were also utilized in the study.

The LFS is a nationwide survey conducted every quarter of the year, which is the country's major source of official employment data. It also collects data on the demographic and socio-economic characteristics of the population [PSA, 2023]. On the other hand, FIES, a rider to the LFS, is conducted every three years. Aside from family income and expenditure, FIES also included the consumption levels of items which led to the identification of spending patterns and standard of living of Filipino families [PSA, 2016]. Since the FIES data set is only available every three years and is considered the main source of poverty statistics, the poverty statistics presented in this paper are also in three-year intervals.

The POPCEN is a complete enumeration of the population in the country and is conducted every ten years, between the conduct of the Census of Population and Housing (CPH). It gives information on the distribution of the population and the characteristics of their housing units.

2.2 DATA ANALYSIS

Numerous SAE techniques can be used to generate statistics in the local area. One of these techniques is the Poisson modeling approach, which is used to model count data to determine the number of poor households at the small areas in NCR.

First, the auxiliary variables from the POPCEN and barangay listing were correlated to the obtained direct estimates to determine the variables significantly associated with poverty counts. The identified auxiliary variables were utilized in developing the Poisson regression model. The maximum likelihood estimators of the regression coefficients were obtained by regressing the log of the expected weighted counts of poor households on the auxiliary variables. The exponentiated predicted values served as the model-based estimates. Estimates of standard error and coefficient of variation were computed to evaluate the predicted values.

Also, the model was assessed through its obtained pseudo-R-square, which is a good criterion for prediction purposes, wherein pseudo-R-square values of around or greater than 0.5 are considered ideal. The parsimony, significance of the predictors and the model itself, and the underlying economic theory of the chosen predictors were also considered. The fit of the model was assessed using the Pearson chi-square goodness-of-fit test. The procedure tested the hypothesis that the total number of poor households follows the Poisson distribution. Rejection of the null hypothesis would mean that the data is not Poisson distributed; hence, the model does not fit the given data set well.

The assessment of candidate models for a region involved a comparison of the similarity (a subset of) parameter estimates and similarity of small area estimates, in addition to basic statistical criteria such as adjusted R squares, among others. The model should also be simple and possess logical parameter estimates.

3. RESULTS AND DISCUSSION

Metro Manila is the central figure of the Philippines and is composed of 16 cities, which include 14 sub-areas, one municipality, and 32 legislative districts. Based on the 2015 POPCEN, the Region has a population of 12,787,669 with 3,095,766 households. From these households, a sample of 17,977 households was used in the 2018 household surveys like the Family Income and Expenditure Survey (FIES) and Labor Force Survey (LFS). The City of Quezon is the most populated in the Region, with a population of 2,919,657, while the Municipality of Pateros is the least populated, having only a population of 63,643.

The characteristics of the residents of NCR as well as the variables on the region's economic condition, were used to estimate the poverty incidences of the 42 legislative districts and sub-areas of NCR in 2018. The Poisson modeling approach was specifically utilized to obtain model-based estimates of the number of poor households in NCR.

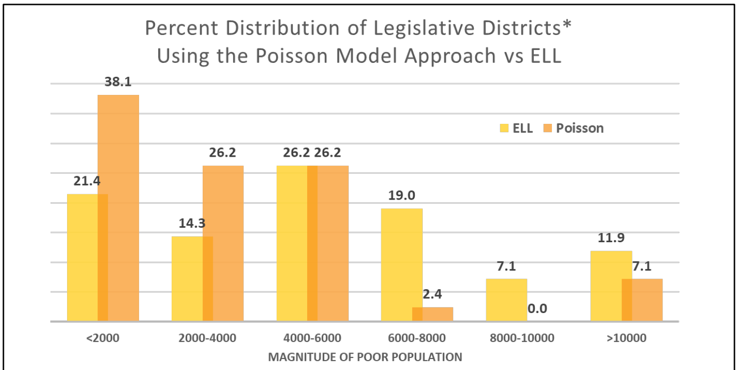
The Poisson modeling approach was specifically utilized to obtain model-based estimates of the number of poor households in NCR. Considering the assumptions of the model, five predictors entered the Poisson regression model, which can be observed in Table 1. One predictor was the average proportion of household members aged 25 and up. The other four predictors deal with the average number of family size in the district, barangays with market areas in the district, household members who are services workers and shop and market sales workers in the district, and household members who are officials of government and special interest organizations, corporate executives, managers, managing proprietors and supervisors in the district. It can also be observed that all predictors are significant at a 5% level, and even the whole model was noted to be significant. The model had a pseudo-R^2 of 60.42%. The Pearson chi-square goodness-of-fit test indicated that the model is not fitted to the data set and must be used with caution. With less than a 0.0001 probability value, the predicting model was used to estimate the total number of poor for each legislative district.

Correlating the direct estimates with the auxiliary variables obtained from the census data source, the number of barangays with market areas obtained the highest correlate (0.4772) with a positive relationship to the direct estimates. The effect of their coefficients on the counts of the poor was also shown in the table. In addition, the average family size in the district also showed a positive effect, implying that increases in family size and barangays with market areas may also increase the poverty count in the district. Meanwhile, the following predictors contributed a negative effect on the total count of poor: the average proportion of household members aged 25 and up in the district, the average number of members who are services workers and shop and market sales workers in the district, and the average number of members who are officials of gov't and special interest organizations, corporate executives, managers, managing proprietors and supervisors in the district. All this suggests that these predictors drive down the number of poor in the district.

Table 1. Predictors of the Poisson model with their corresponding effect on the dependent variable and computed p-value.

Predictor	Effect	p-value
Average proportion of household members aged 25 and up in the district	-	< 0.01
Average family size in the district	+	< 0.01
Average number of barangays with market areas in the district	+	< 0.01
Average number of members who are services workers and shop and market sales workers in the district	-	< 0.01
Average number of members who are officials of government and special interest organizations, corporate executives, managers, managing proprietors and supervisors in the district	-	< 0.01
Constant		< 0.01

The distribution of the legislative districts based on their estimates from the Poisson modeling approach compared with the ELL Approach is shown in Figure 1. The Poisson Model Approach generated lower poverty counts than the ELL approach, which shows that most districts had poverty counts lower than 2000 households for the Poisson Model Approach, which is at 38%. Meanwhile, the ELL approach generated higher incidence of poor households where 64% have poor population greater than 4000 households. There are only a few cities and municipalities falling on the extreme values of estimates specifically on the high values. The model-based estimates obtained from the Poisson Model Approach and the ELL approach are found in Appendix Table 1.



* Including all the sub-areas of the City of Manila and excluding the newly created 3rd District of Caloocan City.

Figure 1. Distribution of legislative districts using the ELL and Poisson model approach.

As observed in the heatmap (Figure 2), it may be noted that the yellow areas are the least poor while those in dark green have the highest magnitude of poor. And those are both of the Districts of Caloocan City. Followed by the first congressional district of Manila which is the West Side of Tondo.

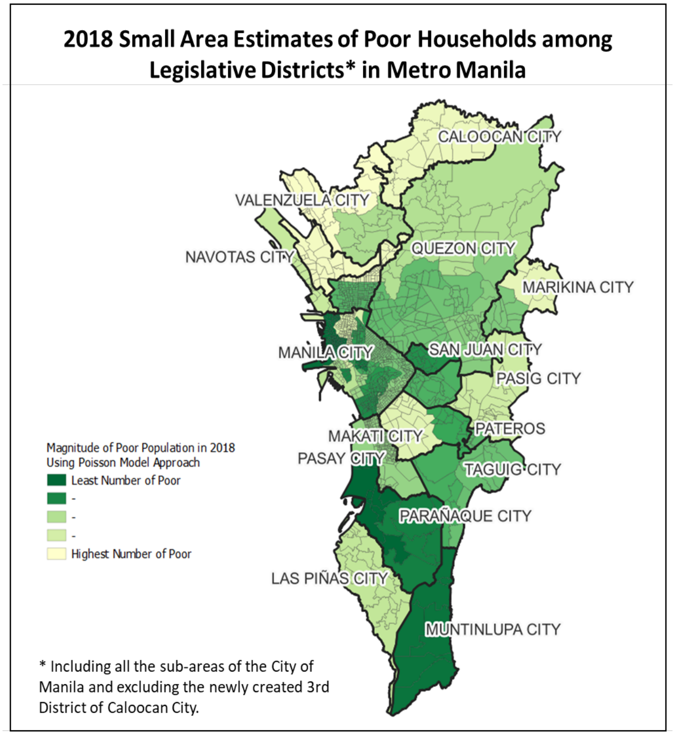


Figure 2. 2018 Small area estimates of poor households among legislative districts in Metro Manila

The reliability of each estimate was determined by obtaining the coefficients of variation (CV) for each of the estimated magnitudes of poor households. Table 2 shows the percentage distribution of the coefficients of variation of the poverty counts obtained. In the Poisson Model Approach, almost all estimates have CVs below 20%, which are deemed reliable. Meanwhile, the ELL approach obtained more than 50% of estimates with CVs greater than 20, which are considered unreliable. There were 33 districts that had a coefficient of variation at most 10 using the Poisson Model Approach, while only one district was observed from the ELL approach. The Poisson Model approach generated more reliable estimates for NCR compared to the ELL methodology. The rest with unreliable estimates should be used with much caution.

Table 2. Distribution of the coefficient of variation of the estimates obtained ELL and Poisson model approach.

Coefficient of Variation	**ELL**		**Poisson**	
	Frequency	**Percent**	**Frequency**	**Percent**
At most 10.0	1	2.4	33	78.6
10.1 – 20.0	16	38.1	8	19.0
Greater than 20.0	25	59.5	1	2.4
Total	**42**	**100**	**42**	**100**

To further compare the differences between the ELL and the Poisson Model approach, Table 3 exhibits the discrepancy between the two techniques used for small are estimation. Although the two SAE techniques ultimately generate poverty measures, the two model-based approaches predicted different dependent variables based on the foundation of their assumptions wherein the ELL methodology was initially crafted by the World Bank to predict log transformed expenditure data for poverty estimation while the Poisson regression model estimates log transformed count dependent variables. It can also be observed that the occupation-related and household composition variables were the significant predictors that entered the two models. Comparing the adjusted R-squared and root mean square error (RMSE) of the two models, it can be observed that the Poisson model approach yield a better predicting power with 60.42% adjusted R-squared against the 48.83% adjusted R-squared of the ELL methodology. Further, the Poisson model approach generated a smaller RMSE of 0.1814 against the 0.4662 of the ELL methodology.

Table 3. Comparison between the ELL and Poisson Model in terms of the dependent variable, predictors and sign of the estimated coefficients.

	ELL	Poisson
Dependent Variable	LN (PER CAPITA INCOME)	LN (POOR POPULATION)
Predictors	(+) Members who have clerical jobs in the city	(+) Barangays with market areas in the district
	(-) Members who are laborers in the city	(-) Members who work as managers in the district
	(+) Proportion of members aged 25 to 60 in the city	(-) Proportion of members aged 25 and up in the district
	(-) Dependency ratio (members aged 18below and 60up / 18-60) in the city	(+) Average family size in the district
	(-) Average squared family size in the city	(-) Members who are services workers in the district
	(-) Proportion of married household head in the city	
	(+) Male members who are at least college graduate in the city	
	(+) Female members who are at least college graduate in the city	
Adjusted R-squared (%)	49.83	60.42
RMSE	0.4662	0.1814

4. CONCLUSION

For estimating the poverty counts in NCR, the 2018 Merged FIES-LFS and 2015 POPCEN were used to generate estimates on the poor households at the legislative district levels of the region. Forty-two estimates were obtained. Almost 98%, which was 41 out of 42, of the generated district and sub-area level estimates have acceptable measures of reliability when using the Poisson model approach, which is a significant improvement on the ELL approach, which is the methodology used on the regularly conducted project on small area estimation of the PSA. It was also observed that most of the districts and sub-areas that obtained high counts of poor households are also among those that are observed to have slower economic activities.

Both ELL and Poisson Model Approach estimates showed that the First Legislative District of Caloocan City remained the highest number of poor households among the districts in the region. Further, based on the obtained standard errors and coefficients of variation, the Poisson model approach was determined to produce more reliable estimates than the other estimation method. Considering common correlates of poverty, it would be best to explore and define the strong determinants of poverty. Since the ELL was initially created for poverty mapping, consultation with the experts on the Poisson model approach methodology may be done to explore other indicators that the methodology could measure specifically for highly urbanized areas such as the NCR.

In conclusion, using the Poisson modeling approach, small area estimation can be used successfully to produce poverty counts at the legislative district levels. The generation of poverty counts is advantageous for policymakers since identifying areas with a high magnitude of poor population aids the government in establishing programs for the right beneficiaries. Since the technique generates poverty counts, this approach could be explored to produce poverty incidences to be compared with municipal and city-level estimates generated from the SAE of Poverty Project of PSA.

Furthermore, the technique could be further explored for other relevant indicators to address the data gaps in the country's official statistics to support the monitoring of the Philippine Development Plan as well as the Sustainable Development Goals.

5. LITERATURE CITED

NATIONAL ECONOMIC AND DEVELOPMENT AUTHORITY (NEDA), 2021, Statement on the 2021 First Semester Official Poverty Statistics. Available at: https://neda.gov.ph/statement-on-the-2021-first-semester-official-poverty-statistics

PHILIPPINE STATISTICS AUTHORITY (PSA) 2016 Labor Force Survey (LFS). Available at: https://psa.gov.ph/statistics/survey/labor-and-employment/labor-force-survey/tech-notes

____, 2022, Full Year 2021 Official Poverty Statistics Authority. Available at: https://psa.gov.ph/poverty-press-releases

____, 2023 Family Income and Expenditure Survey (FIES). Available at: https://psa.gov.ph/income-expenditure/fies-technical-notes

REYES, C., ASIS, R., ARBONEDA, A., and VARGAS, AR., 2020, Mitigating the Impact of COVID-19 Pandemic on Poverty. Philippine Institute for Development Studies. Available at http://hdl.handle.net/11540/13038.

APPENDIX: 2018 Model-based estimates for the total number of poor households in NCR using the Poisson Model Approach and Elbers, Lajouw and Lanjouw (ELL) Methodology.

PSGC Code	Legislative District/ Sub-Area	ELL		Poisson	
		Estimate	CV (%)	Estimate	CV (%)
133900302	Binondo	236	27.0	492	12.8
133900303	Quiapo	1,073	22.4	1,256	13.6
133900304	San Nicolas	1,724	26.4	3,655	10.3
133900305	Santa Cruz	2,537	14.5	1,722	12.0
133900406	Sampaloc	6,725	9.5	2,999	10.4
133900508	Ermita	793	18.4	342	22.5
133900509	Intramuros	240	22.1	1,310	16.6
133900510	Malate	232	33.0	1,314	10.7
133900511	Paco	2,084	13.9	1,705	7.6
133900513	Port Area	1,379	20.0	4,218	10.1
133900607	San Miguel	1,651	20.5	1,879	11.0
133900612	Pandacan	4,122	29.4	1,746	9.1
133900614	Santa Ana	4,221	15.2	3,192	6.7
133901101	Tondo 1	14,967	12.5	11,716	7.7
133901201	Tondo 2	4,248	13.3	4,939	5.6
137400001	Mandaluyong City, Lone District	9,147	13.3	2,273	10.1
137400003	Pasig City, Lone District	12,806	25.0	4,860	6.5
137400005	San Juan City, Lone District	2,458	17.1	1,679	6.4
137402102	City of Marikina, 1st District	3,372	31.0	3,002	5.8
137402202	City of Marikina, 2nd District	5,506	24.3	5,008	5.9
137404104	Quezon City, 1st District	5,634	14.6	2,661	7.4
137404204	Quezon City, 2nd District	18,453	20.8	4,207	6.5
137404304	Quezon City, 3rd District	4,218	15.1	3,005	7.2
137404404	Quezon City, 4th District	5,685	20.6	2,842	8.5
137404504	Quezon City, 5th District	10,999	14.7	6,662	5.3
137404604	Quezon City, 6th District	8,547	19.8	4,240	6.3
137500002	Malabon City, Lone District	7,461	27.4	5,781	6.3
137500003	Navotas City, Lone District	6,368	26.4	4,775	7.1
137501101	Caloocan City, 1st District	29,453	22.4	51,773	5.4
137501201	Caloocan City, 2nd District	6,699	17.1	22,860	6.8
137504104	Valenzuela City, 1st District	4,839	30.5	5,977	6.6
137504204	Valenzuela City, 2nd District	3,804	32.2	4,034	6.6
137600001	Las Piñas City, Lone District	9,684	24.1	4,526	8.1
137600003	Muntinlupa City, Lone District	6,242	47.0	1,195	13
137600005	Pasay City, Lone District	6,684	11.9	3,343	10.6
137602102	Makati City, 1st District	2,731	19.0	501	12.2
137602202	Makati City, 2nd District	4,116	17.5	1,787	8.9
137604104	Parañaque City, 1st District	4,351	26.7	1,102	11.3
137604204	Parañaque City, 2nd District	5,179	26.7	1,291	9.3
137667106	Pateros	1,273	33.8	2,482	7.7
137667107	Taguig, 2nd District	6,498	28.5	2,750	9.8
137677207	Taguig (w/o Pateros), 1st District	6,413	24.9	1,997	11.7

Note: For computational exercises only. Not for official release.