A Bayesian Hierarchical Modelling of Small Area Variation in Youth Unemployment in Namibia

Petrus Tweuthigilwa Iiyambo

Department of Computing, Mathematical and Statistical Sciences

University of Namibia, Windhoek, Namibia, 9000

Email: piiyambo@unam.na

Linda Vute Shitenga

Manager: Social Statistics

Namibia Statistics Agency

Department of Demographic and Social Statistics

Division of Social Statistics

Windhoek, Namibia, 9000

Email: vutelinda@gmail.com

1. Introduction

Youth unemployment is considered a barrier to the Southern African region's development (Devlin, 2013). The government frequently finds it difficult to implement employment creation measures despite Namibia's high youth unemployment rate because there is a lack of up-to-date small-area labour market statistics. Using the 2018 Namibia Labour Force Survey data, Bayesian hierarchical modelling was used to estimate the risk of youth unemployment in Namibia at the constituency level for targeted intervention. The findings demonstrated that youth unemployment varies greatly across constituencies, and rural youths are more likely to be unemployed than those in urban areas. In urban constituencies, male youths had a much higher chance of unemployment than female youths. Furthermore, youth without formal education are more likely to be unemployed as compared to those with formal education. The study recommends that more priority be given to integrating the youth into the labour market by improving their educational levels and implementing labour market policies to facilitate entry into the labour market, especially for rural youths.

2. Methodology

The data used in this study were extracted from the Namibia Labour Force Survey (NLFS), administered by the Namibia Statistics Agency (NSA) in 2018. The NLFS is conducted yearly from 2012 to 2018, to provide labour force information on the employment, socio-demographic and educational characteristics of all persons aged 8 years and above living in households in Namibia. However, the survey was not conducted in 2015 and 2017 due to a lack financial resources. All NLFS reports are freely available online at www.nsa.org.na. For more information about the 2018 survey, refer to the NLFS report of 2018.

The inclusion criteria for this study were all youth aged 15-34 years living in households during the reference period of the survey, as documented in the NLFS report of 2018, while persons who were above the age of 34 years were excluded from this study. The data used for this study was obtained freely from the NSA website and treated with confidentiality. All revealing information about the identities of the youth was already excluded from the data by the NSA before the data were made available online.

Hierarchical Bayesian (HB) approach has recently been proposed for SAE because of the following advantages as stipulated by (Trevisani and Torelli, 2007);

- Their specification is straightforward and allows them to take into account the different sources of variation and;
- Inferences are clear-cut and computationally feasible in most cases by using standard MCMC techniques.
- One can use this approach when the variable of interest is counted or a proportion, alternative model specifications can be considered.

The evaluation of the uncertainty in the unobserved random effects that are also contributing to the variation in the average probability of regional and constituency unemployment rate through the full posterior inference approach was achieved by incorporating the unobservable random effects of interest into the full hierarchical Bayesian model.

The outcome variable for this study is the unemployed youth counts data aggregated to the region and constituency level to carry out HB analysis. In addition, socioeconomic variables (head of household status, school attendance, highest level of education completed, marital status and citizenship, urban/rural) are also considered as covariates for modelling. The outcome variable for this study is the employed and unemployed youth counts data aggregated to the region and constituency level to carry out HB analysis.

The outcome variables for this study are binary; hence the binary logistic regression was used which is of the general form

$$\text{logit}(\pi) = \log(\frac{\pi}{1-\pi}) = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k$$

Taking into consideration that classical Generalized Linear Models (GLMs) make no, or limited use of the spatial structure of the data, neither do they consider possible nonlinear effects of the risk factors (Mwahi, 2014), a Structured Additive Regression Model (STAR) was fitted followed by jointed spatial models which were done through the shared component latent variables approach.

The STAR model that was used in this study is of the form

$$\eta = \gamma_0 + \gamma_1 noedu + \gamma_2 tedu + \ldots + \gamma_k citizen + f^{str}{}_{(const)} + f^{unstr}{}_{(const)}$$

where η is the predictor and γ_i are the fixed effects parameters. Spatial effects of the constituency are split into spatially and correlated part $f^{str}{}_{(const)}$ and an uncorrelated part $f^{unstr}{}_{(const)}$ where the correlated part is modelled by a Markov random field prior where the neighbourhood matrix and possible weights associated with the neighbours are obtained from the constituency map.

3. Findings

3.1 Model comparison

Table 1. Model comparison in respect of male and female youth unemployment

	Male				Female			
Model	Deviance	pD	DIC	ΔDIC	Deviance	pD	DIC	ΔDIC
M_0	4945.85	6.77	4959.39		5788.08	7.02	5802.12	
M_1	5441.23	61.42	5564.08	663.18	6155.55	54.19	6263.94	544.46
M_2	5597.53	10.08	5617.70	716.79	6270.55	12.04	6294.63	575.14
M_3	5402.96	71.16	5545.29	644.38	6134.13	63.98	6262.09	542.61
M_4	5596.73	10.86	5618.45	717.55	6269.34	12.30	6293.95	574.47

M_5	5447.30	58.47	5564.24	663.34	6161.48	50.01	6261.49	542.01
M_6	5404.79	70.66	5546.11	645.20	6142.69	57.94	6258.57	539.09
M_7	4829.92	44.84	4919.61	18.71	5607.90	59.31	5726.52	7.04
M_8	4890.67	16.71	4924.10	23.19	5708.63	17.95	5744.53	25.04
M_9	4837.70	36.43	4910.56	9.65	5624.21	50.22	5724.64	5.16
M_{10}	4783.67	61.58	4906.83	5.93	5590.23	67.31	5724.85	5.37
M_{11}	4889.50	17.39	4924.29	23.39	5707.62	18.36	5744.34	24.85
M_{12}	4789.16	55.87	4900.90	0.00	5603.16	58.16	**5719.48**	0.00

Model 12 with a smallest DIC for both males (4900) and females (5719.48) was found to be the best model taking into effect the regional and constituency effects. The table show that model M9 and M10 were weakly differentiated as indicated by the DIC differences from the best model.

3.2 Fixed effects results

It can be observed that female youths with no formal education are more likely to be unemployed compared to the male youths with a formal education, OR=1.11 (0.85, 1.46) and OR=0.48 (0.38, 0.61) respectively. However, the probability of being unemployed is very high among male youths than female youths in urban areas with OR=1.35 (1.10, 1.66) and OR=0.79 (0.65, 0.96) respectively. The table also illustrates that the risk of a youth being unemployed is low for those who are regarded as head of household amongst males and females.

Table 2: Odds ratios and their 95% confidence intervals (C.I.'s) for fixed effects summaries of the best STAR models.

	Male		**Female**	
Covariates	**OR**	**(95% C.I.)**	**OR**	**(95% C.I)**
Is youth Head of Household?				
No	1.00		1.00	
Yes	0.19	(0.16, 0.22)	0.23	(0.20,0.27)
Is youth attending school?				
No	1.00		1.00	

Yes	0.61	(0.47, 0.76)	0.47	(0.37,0.61)
Youth has no formal education				
No	1.00		1.00	
Yes	0.48	(0.38, 0.61)	1.11	(0.85,1.46)
Youth has tertiary education				
No	1.00		1.00	
Yes	0.51	(0.39, 0.66)	0.48	(0.38,0.60)
Youth's marital status				
Never married	1.00		1.00	
Married	0.47	(0.34, 0.67)	0.48	(0.39,0.59)
Type of residence				
Rural	1.00		1.00	
Urban	1.35	(1.10, 1.66)	0.79	(0.65,0.96)

3.3 Spatial effects

The study shows that most of the constituencies in some regions have either a low or moderate probability of unemployment except Gibeon constituency in the Hardap region. The study also shows that both males and females were at high risk of being unemployed mostly in constituencies such as the Kabbe South (in Zambezi region), Keetmanshoop rural (in !Karas region) and Katima Mulilo rural (in Zambezi) for males while for females it was observed in the Epembe (in Ohangwena region) and Epukiro constituencies (in Omaheke region).

4. Conclusion

This study had several limitations due to the data and the methods used. This study used the 2018 Namibia Labour Force Survey (NLFS) data. This is the latest national labour force survey conducted in Namibia to date; hence the youth unemployment situation may have changed over the past years given the negative impacts of Covid-19 pandemic that has taken place from 2020 to date. Based on the study findings, it may not be obvious which specification is better when estimating counts, but pertinent quantities should be accurately characterized. In conclusion, it is clear that future research in the subject of SAE should concentrate on defining appropriate devices for model determination. Additionally, model structure must be complicated to take into account more realistic situations such as adding auxiliary information, more complex sampling designs must be experienced, and various real population phenomena must be investigated in order to magnify differences between models at comparison in simulation studies “small areas at different level of territorial aggregation, small area population with different distributional characteristics, etc.” (Trevisani and Torelli, 2017).

5. References

Mwahi, E. M. (2014). *Shared-component model with application to mapping gender specific pattern in HIV testing and condom use in Namibia (Doctoral dissertation).*

Namibia Statistics Agency. (2019). *The Namibia Labour Force Survey 2018 Report.* Namibia Statistics Agency. Windhoek

Trevisani, M. and Torelli, N. (2017) A Comparison of Hierarchical Bayesian Models for Small Area Estimation of Counts. *Open Journal of Statistics*, 7, 521-550.

Trevisani, M., and Torelli, N. (2007). Hierarchical Bayesian models for small area estimation with count data. *Dipartimento di Scienze Economiche e Statistiche, Universita Degli studi di Trieste*.

Gulsuner, S., Walsh, T., Watts, A. C., Lee, M. K., Thornton, A. M., Casadei, S., ... & McClellan, J. M. (2013). *Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network*. *Cell*, *154*(3), 518-529.