ETHIOPIAN POPULATION HEALTH (and analytical lessons from California)

Presentation for Addis Ababa University, College of Social Sciences

October 24, 2022

By Aynalem Adugna, PhD Research Scientist III, Health Equity Research and Statistics Section (HERSS) California Department of Public Health

CONTENTS

- Objectives
- A spatial statistical tool in focus: the Geographically Weighted Regression (GWR)
- The Ethio-demography and health website
- Income Inequality and Health (lessons from California)

OBJECTIVES

- To highlight a spatial statistics tool and show examples of its application using Ethiopia's 2016 Demographic and Health Survey (DHS) data
- 2. To introduce a demography and health website and solicit contributions from demographers, epidemiologists, sociologists, and geographers
- To share experiences and best practices from my employment at the California Department of Public Health

PART I. SPATIAL STATISTICS

The research question: Are the unweighted proportions of women's responses to DHS question V744c related to educational status?

Q. V744c: It is the respondent's opinion that a husband is justified in hitting or beating his wife when she argues with him

- Yes
- No

METHODOLOGY

Data source: Eth. DHS 2016

- Data processing using R-programing language
- Spatial statistics using geographically weighted regression (GWR) - an ArcGIS-Pro application

R-PROGRAMMING

DHS files

- > #Couples' recode
- > DHS2016_Couple <- read.dta("D:/DHS2016/ETCR71DT/ETCR71FL.DTA")
- > nrow(DHS2016_Couple)
- [1] <mark>6141</mark>
- > #Household recode
- > DHS2016_HHold <- read.dta("D:/DHS2016/ETHR71DT/ETHR71FL.DTA")
- > nrow(DHS2016_HHold)
- [1] <mark>16650</mark>
- > #Individual recode
- >DHS2016_Wom <- read.dta("D:/DHS2016/ETIR71DT/ETIR71FL.DTA")</p>
- > nrow(DHS2016_Wom)
- [1] <mark>15683</mark>
- > #Births recode
- > DHS2016_Births <- read.dta("D:/DHS2016/ETBR71DT/ETBR71FL.DTA")
- > nrow(DHS2016_Births)##41392
- [1] <mark>41392</mark>
- >

Women's Attitude Toward Wife-beating (Eth. DHS 2016)

It is the respondent's opinion that a husband is justified in hitting or beating his wife when:

#V744A She goes out without telling him
#V744B She neglects the children
#V744C She argues with him
#V744D She refuses to have sex with him
#V744E She burns the food

#V744C She argues with him (National)

R-programing (National) DHS2016_Wom <- read.dta("D:/DHS2016/ETIR71DT/ETIR71FL.DTA") > nrow(DHS2016_Wom)

[1<mark>] 15683 (number of women <mark>- UNWeighted</mark>)</mark>

> summary(DHS2016	Wom	<mark>\$v744c</mark>

no 9839 (62.9%)

5707 (36.4%)

don't know 137 (0.7%)

R-programing (Regional)

#Regional proportions of yes/no answers to beatings if wife argues with husband

WomArgReg <- table(DHS2016_Wom\$v101,DHS2016_Wom\$v744c)# WomArgReg # print table

#V744C She argues with him (Regional)

Regional percentage of women who answered yes/no to a question on wife-beating over spousal arguments

Region	No	% NO	Yes	% YES		Don't l	%
Afar	531	47.1	582		51.6	15	<u>1.3</u>
Oromia	965	51.0	905		47. 0	22	2.0
Tigray	894	53.2	783		<mark>46.</mark> 4	5	0.4
SNNPR	1014	54.8	823		44 .1	12	1.1
Amhara	980	57.0	725		<mark>4</mark> 1.7	14	1.2
Gambela	614	59.3	414		4 0.1	7	0.6
Benishang	725	64.4	400		35.5	1	0.1
Somali	913	65.6	446		31.5	32	2.8
Harari	654	72.2	243		27.0	9	0.8
Dire Dawa	879	77.7	239		21.1	13	1.2
Addis Aba	1670	91.6	147		7.8	7	0.6



SPATIAL Autocorrelation Global Moran's I

50	23	58	38
19	84	→ 16	55
78	13	83	27
38	75	26	41

Dissimilar values Negative

-				
	84	83	58	38
Ļ	78	75	50	27
888888	55	41	26	19
	38	23	16	13
88				

Similar values Positive

19	83¢	84	>13
38	55	58	26
50	41	38	75
16	78	23	75

Random pattern No autocorrelation



Geographical HOT-SPOT Analysis



Source: https://pro.arcgis.com/en/pro-app/2.7/tool-reference/spatialstatistics/hot-spot-analysis.htm

Results of the Geographical HOT-SPOTS Analysis



Geographical OPTIMIZED HOT-SPOT Analysis

"Similar to the way that the automatic setting on a digital camera will use lighting and subject versus ground readings to determine an appropriate aperture, shutter speed, and focus, the Optimized Hot Spot Analysis tool interrogates your data to obtain the settings that will yield optimal hot spot results."

Source: <u>https://pro.arcgis.com/en/pro-app/2.7/tool-reference/spatial-statistics/how-optimized-hot-spot-analysis-works.htm</u>

Making sure there are enough weighted features for analysis.... - There are 643 valid input features. Evaluating the Analysis Field values.... - PCTWFBT YE Properties: 0.0000 Min: 96.0000 Max: 37.0251 Mean: Std. Dev.: 22.7485 Looking for locational outliers.... - There was 1 outlier location; it will not be used to compute the optimal fixed distance band. Looking for an optimal scale of analysis by assessing the intensity of clustering at increasing distances.... - The optimal fixed distance band is based on peak clustering found at 166072.1750 Meters Finding statistically significant clusters of high and low PCTWFBT YE values.... - There are 497 output features statistically significant based on an FDR correction for multiple testing and spatial dependence. - 2% of features had less than 8 neighbors based on the distance band of 166072.1750 Meters Creating output feature class: D:\EthGIS\EthioDHS 2016\Default.gdb\DHS2016 WomV744C OptimizedHotSpotAna lvsis - Red output features represent hot spots where high PCTWFBT YE values cluster. - Blue output features represent cold spots where low PCTWFBT YE values cluster. Succeeded at Sunday, September 25, 2022 9:26:59 AM (Elapsed Time: 6.80 seconds)

Results of the Geographical **OPTIMIZED HOT-SPOT** Analysis



Looking for an explanation of the spatial pattern

What explains women's "yes" or "no" answers to the below question?

It is the respondent's opinion that a husband is justified in hitting or beating his wife when: #V744c She argues with him

- Urbanism appears to be a factor: case in point are the low "yes" percentages in Dire Dawa (21.1), and Harari (27.0) and the single-digit percentage in Addis Ababa (7.8)
- Urbanism could be a proxy for women's level of education
- The low "yes" percentages in the primarily rural Somali Region where less than a third (31.5%) of interviewees gave such a response, shows that such occurrences are not limited to primarily urban locations

Looking for an explanation...contd. Terminology refresher (Multivariate and bi-variate regression)





Source: ESRI, https://www.youtube.com/watch?v=ob18SuuAJGI

Looking for an explanation...contd.

Problems with bi/multi-variate regression between V744c (y-variable) and any explanatory variables (x):

- The Morans-I statistic showed the presence of spatial autocorrelation for the dependent variable V744c – women's response to a question on wife-beating
- □ The presence of spatial autocorrelation creates problems with using x variables to explain y variables as it violates the assumption of the independence of y residuals*
- The main effect of such a violation is to underestimate the Error Sum of Squares (SS) thus inflating the value of the test statistic*
- This means that V744c (y) is not suitable for the ordinary least squares (OLS) regression analysis.

Solution

The solution is the use of the Geographically Weighted Regression method which takes into account the presence of spatial autocorrelation and makes corrections to overcome its effects**

*Source:

https://ibis.geog.ubc.ca/courses/geob479/notes/spatial_analysis/spatial_autocorrelation.htm#:~:text=lf%20spatial%20autocorrelation%20is%20present,the% 20value%20of%20test%20statistic

**Source: https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1538-4632.1996.tb00936.x

Looking for an explanation...contd. Is the (y) variable suitable for ordinary least squared (OLS) regression (example: (y) versus education (x)?



The answer is no. The above visualization proves that it would be wrong to try and apply a single regression equation or model in the hope that it would be valid for the entire country

Looking for an explanation...contd. The dependent variable, V744c is not suitable for ordinary least squares regression as the residuals are not randomly distributed; they are clustered



Given the z-score of 57.666378, there is a less than 1% likelihood that this clustered pattern could be the result of random chance.

The solution is a Geographically Weighted Regression - GWR

https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1538-4632.1996.tb00936.x

 With GWR, each cluster is grouped with its immediate neighbors and is given a separate equation*



*Source: ESRI https://www.youtube.com/watch?v=ob18SuuAJGI

GWRcontd. Defining local



*Software finds an optimum distance by increasing distances by thirds (golden search) until the best local minima are found

Source: ESRI https://www.youtube.com/watch?v=ob18SuuAJGI



Histogram of the Percentage of Women Responding "yes" to a Question on Wife-beating



Bi-variate B₀, **B**₁ plots for Hypothetical Sampling Clusters A through E



Geoprocessing	- □ ×
(€) Geographically Weighted Regre	ssion (GWR) (+)
() j j	
Parameters Environments	?
Input Features	
DHS2016_Wom\$V744C	
Dependent Variable	Y
PCTWFBT_YE	-
Model Type	
Continuous (Gaussian)	-
Explanatory Variable(s)	Select All 🤣
PCTWFBT_YE	1
NO_EDUC	
PRIMARY	
SECONDARY	
HIGHER	
PCNT_NOED	
PIPED_WTER	
NO_RADIO	
THRE_PLSU5	
Output Features	
MASTER_CLOSTER_GWR2	
Neighborhood lype	
Number of neighbors	
Golden search	
Minimum Number of Neighbors	
	¥

🕟 Run



WARNING 110306: The final model didn't have the lowest AICc encountered in the Golden Search Results. WARNING 110259: At least one local regression had very limited variation after applying the weights. Use caution when interpreting the results. ----- Analysis Details -------- Golden Search Results Number of Features: 643 ----Dependent Variable: PCTWFBT YE **#Neighbors** AICc Explanatory Variables: PCNT NOED 30 5514.0940 Number of Neighbors: 72 643 5667.8465 264 5548.8335 409 5569.4515 ----- Model Diagnostics ------175 5524,7560 R2: 0.5123 119 5500.4482 85 5491.6356 AdiR2: 0.4585 64 AICc: 5489.8314 98 Sigma-Squared: 280.1931 77 Sigma-Squared MLE: 252.3868

Effective Degrees of Freedom: 579.1888

5492.1195 5493.9578 5489.8220 72 5489.8314

Succeeded at Friday, September 23, 2022 7:26:15 AM (Elapsed Time: 6.19 seconds)

Did the Geographically Weighted Regression method correct the ordinary least squared (OLS) regression problems associated with spatial autocorrelation of the (y) variable?



Yes. Geographically Weighted Regression corrected the ordinary least squared (OLS) regression problems associated with spatial autocorrelation of the (y) variable

GWRcontd. Proof that the assumption of independent residuals was not violated



Standardized Residuals (&), Geographically Weighted Regression with Women's Response on Wife-beating as (y) variable and education as (x) variable



Intercept (\mathcal{B}_0), Geographically Weighted Regression with Women's Response on Wife-beating as (y) variable and education as (x) variable



Beta Coefficient (\mathcal{B}_{r}), Geographically Weighted Regression with Women's Response on Wife-beating as (y) variable and education as (x) variable



Local R-Squared, Geographically Weighted Regression with Women's Response on Wife-beating as (y) variable and education as (x) variable



PART I: SUMMARY

- Place-based ordinary least squared (OLS) regression analyses produce <u>global</u> results for all locations in the study by neglecting the role of geographical distance and the resulting autocorrelation of values from nearby locations
- As near things are more related than distance things, such negligence results in biased statistical estimations
- Using the geographically weighted regression method (GWR) corrects the traditional methodological flaws by producing <u>local</u> results for all locations in the study
- This presentation has provided evidence by testing for the presence of spatial autocorrelation in unweighted proportions of women responding to a question on wife-beating
- Nationally, a "no education" status explained 46% of the variability in the percentage of women responding "yes" to a question on wife-beating (V744c) but this varied locally from 0 to a high of 71%.
- Other explanatory variables, for example, the DHS Wealth Index, could have been included in the analysis but were omitted due to the risk of multicollinearity
- Results obtained in this analysis call for policy interventions aimed at protecting women against spousal abuse to be framed with local (not national or regional) sets of preventative measures in mind

PART II. ETHIO-DEMOGRAPHY-AND-HEALTH

Visits					RM	Day Week Month
#	Date/Time	Visitor	Referrer	Page Views	Earlier Page Views	Image: Weight of the sector of the
1	10:52 AM	5 0	···· 👌	1		4 5 6 7 8 9 10 11 12 13 14 15 16 17
2	10:47 AM	0	tigray.html	2		18 19 20 21 22 23 24 25 26 27 28 29 30 1
3	10:45 AM	6	www.ethiodemographyandhealth.org/	1		Quick select:
4	10:17 AM	0	···· 🔧	1		Select a period V
5	10:15 AM	6	www.ethiodemographyandhealth.org/	1	2	Prom: Io: 9/16/2022 9/30/2022
6	8:59 AM	2	amhara.html	3		
7	8:54 AM	2 0	www.ethiodemographyandhealth.org/	1		
8	8:50 AM	2 0	···· 3	2		
9	8:33 AM	0	www.ethiodemographyandhealth.org/	1		
10	7:58 AM	0	www.ethiodemographyandhealth.org/	1		
11	7:38 AM		···· 🚷	1		
12	7:24 AM	2	···· 👌	1		
13	6:36 AM	0	www.ethiodemographyandhealth.org/	2		
14	6:35 AM	0	More_faces.html	1		
15	6:26 AM	6	Benishangul.html	3	9	
16	6:24 AM	0	.Benishangul.html	> 1		
17	6:22 AM	0	oromia.html) 1		
18	6.22 AM	-	www.ethiodemographyandhealth.org/	1		

...www.ethiodemographyandhealth.org/

1

Reports
Settings
Traffic
Referrers
Visitor Profiles
Technology
Behavior
Visitors
Current Visitors
Last 100 Visits

19

6:16 AM

0

PART II. DATA DASHBOARDS AND MAPS ON ETHIO-DEMOGRAPHY-AND-HEALTH WEBSITE

Demography and health

https://ethiodemography.maps.arcgis.com/apps/opsdashboard/index.html#/e485f88663af 4304992f9980324bd69a

Health care institutions

https://ethiodemography.maps.arcgis.com/home/webmap/viewer.html?webmap=8e3f3e3 a6c6c47acbc5da3192d1ef041&extent=33.5072,4.7111,41.6206,9.2193

COVID-19

https://ethiodemography.maps.arcgis.com/apps/dashboards/a612bb3e59174cc8920b4b4 16705d3f3

http://www.ethiodemographyandhealth.org/Aynalem_Adugna_COVID_19_in_Ethiopia_P DF.pdf

My proposal for an Ethiopian population journal

Ethiopian Journal of Population Studies (ethiojps.org)

PART III. INCOME INEQUALITY AND HEALTH Lessons from California

Income inequality and health: Story Map and dashboard

https://storymaps.arcgis.com/stories/84fcd9cce18f498b8ac5861b95ad d64a